



*at*  
***Queens College***  
*of the*  
***City University of***  
***New York***  
**October 13-15, 2023**

---

# **Variation in the World's Languages**

Queens College and the Program Committee welcome you to Queens, New York, “the world’s borough,” the most linguistically and ethnically diverse county in the US with a population of almost 2.5 million mostly first or second generation immigrants. Queens College has a mostly commuter student body of 16,500 that reflects the borough it is named for. Be sure to take the opportunity to talk to one of our undergraduate volunteers, and you’ll meet the future of New York and the US. Also, take a break from talks to walk around campus. You’ll see the skyline we use in our logo by Rosenthal Library, and be sure to go onto the streets and sample the food near campus. There is Chinese, of course, across the Long Island Expressway, typical NY Pizza at Gino’s across Kissena Blvd, or going south, away from the Expressway, you’ll find a Dominican Diner and an excellent Afghan restaurant. There are also food carts of various types on and off campus. Halal carts—essentially meaning a kind of Afghan-Arab combination—have led a recent NY coinage, the count noun “a halal” as in something you get from them, a plate of rice and lamb, for example.

Also, get to know Downtown Flushing. There are food courts like the New World Mall, which would not be out of place in Singapore or Hong Kong. If you get a chance, take the Number 7 train to Manhattan, or get off at 74<sup>th</sup> Street, Jackson Heights. The stop is at the intersection of a South Asian neighborhood to the north, a Latin American one to the east, a Chinese, Thai, and Vietnamese one to the Northeast, and a Filipino one to the West.

<b>PLENARIES</b> .....	<b>1</b>
REVISITING CONTACT AND SWITCHING IN BILINGUAL NEW YORK FROM A SPEAKER-CENTERED TRANSLANGUAGING PERSPECTIVE .....	1
CROSSLINGUISTIC VARIATION IN ARTICLE SYSTEMS .....	1
IS VARIATIONIST SOCIOLINGUISTICS FIRST-WORLD PROBLEM-SOLVING? QUANTITATIVE THEORY/METHODOLOGY AND AFRICAN LANGUAGES.....	2
VARIATION IN ENGLISH NEGATION REVISITED .....	2
<b>TALKS AND POSTERS</b> .....	<b>4</b>
DIASPORA IDENTITY AND LANGUAGE IN THE MALAYALEE-AMERICAN COMMUNITY.....	4
SOCIOSYNTACTIC VARIATION IN SINGULAR THEY REFLEXIVES .....	5
AN INTERDISCIPLINARY APPROACH TO INVESTIGATE LANGUAGE CHOICES IN A TRILINGUAL SETTING.....	6
ALL OF (US/WE) ARE INTERESTED IN TYNESIDE ENGLISH FIRST-PERSON PLURAL OBJECTS.....	7
ETHNICITY-SPECIFIC AND -INDEPENDENT LEARNING: EFFECTS OF GUISE ON ADAPTATION TO NOVEL FOREIGN-ACCENTED TALKERS .....	8
DISCRETENESS IN A CREOLE-STANDARD CONTINUUM? A PERCEPTUAL AND ATTITUDINAL APPROACH TO THE ‘NAMED LANGUAGES’ DEBATE.....	9
FROM DIGLOSSIA TO CREOLE REVIVAL AND CHANGING INDEXICALITIES: A PERCEPTUAL AND ATTITUDINAL TESTING OF THE MARTINICAN (SEMI-)CONTINUUM .....	10
A STUDY ON BARE PREPOSITIONS: A POTENTIAL INNOVATION IN BRAZILIAN PORTUGUESE.....	11
NAVIGATING VARIETIES OF ARABIC IN LITERACY ASSESSMENT .....	12
A POTENTIAL PHONOLOGICAL CHANGE IN AN UNDER-STUDIED SPEECH COMMUNITY: TUNISIAN TAMAZIGHT VARIETY OF FAHMINE.....	12
NOT ALL SOUND CHANGES IN PROGRESS ARE USED IN EARLY-ADOLESCENT STANCETAKING .....	13
CORPUS DATA MEET SURVEY DATA: A CASE STUDY FROM SINGAPORE ENGLISH MODALITY.....	14
A RESPONSE TO CHAMBERS AND TRUDGILL (1980): THE STYLISTIC IMPLICATIONS OF (ING) ACROSS THE LIFESPAN OF THE INDIVIDUAL	15
INTENSIFIER VARIATION AND USE: THE EFFECTS OF FORMALITY ON THE INTERNET.....	16
A WHO YUH TINK SEH MEET TINA: A REANALYSIS OF “SAY” COMPLEMENTIZERS IN JAMAICAN PATOIS.....	16
INNOVATION AND STANDARDIZATION IN CROSS-SOCIOLINGUISTIC CORPUS DEVELOPMENT: A CASE STUDY WITH SUBJECT PRONOUN EXPRESSION .....	17
LEXICAL VARIATION IN TERMS FOR ROMANTIC PARTNERS IN QUÉBEC FRENCH: A DIACHRONIC STUDY .....	18
AUTOMATIC TRANSCRIPTION OF LANGUAGE DOCUMENTATION RECORDINGS FOR SOCIOLINGUISTIC ANALYSIS: SPEECH RECOGNITION AND FORCED-ALIGNMENT FOR NORTHERN PRINMI.....	19
OPERATIONALIZING GENDER: METHODS FOR STATISTICAL MODELING OF OPEN-RESPONSE DEMOGRAPHIC DATA.....	20
INTENSIFIERS NEVER GO OUT OF STYLE: QUANTIFYING STYLE AND ITS EFFECTS ON LEXICAL VARIATION.....	21
UNDERSPECIFICATION AND GENDER/SEXUALITY-BASED MEANINGS OF NOUN-PHRASE NUMBER AGREEMENT.....	21
LITERAL OR DETAIL-ORIENTED? EXPLORING THE SOURCE OF NERDY PERSONA EFFECTS ON PRAGMATIC REASONING .....	22
OPERATIONALIZING GENDER: IMPLICATIONS OF PARTICIPANT-DRIVEN MEASUREMENTS .....	23
/T/-RELEASE AND BURST DURATION IN THE CONSTRUCTION OF JEWISH MASCULINITY .....	24
RHYTHM AS A CONTESTED MARKER OF ETHNICITY IN MODERN HEBREW .....	25
THE TRIGGERING ROLE OF AGE-GRADING: THE RISE OF SIMPLE PRONOUNS OVER 40 YEARS IN LAURENTIAN FRENCH.....	26
INDIVIDUAL AND GROUP-VARIATION: THE PRAGMATIC FUNCTIONS OF RISING TONAL ACCENTS ON SOUTH FUNEN. ....	27
THE USE OF THE UNSTRESSED PAST TENSE PREFIX AMONG GREEK-CANADIANS .....	28
THE UNIVERSALITY OF POLARITY SENSITIVITY.....	28
SALIENCE AND SOUTHERNNESS: A PERCEPTION STUDY ON [HW].....	29
“WE WAS ONE BIG HAPPY FAMILY THERE”: TRACKING THE EVOLUTION OF PAST BE IN FALKLAND ISLAND ENGLISH .....	30
A SOCIOPHONETIC STUDY OF CREAKY VOICE IN CANADIAN ENGLISH AND FRENCH .....	31
SPANISH GRAMMATICAL GENDER ASSIGNMENT OF BASQUE LEXICAL ITEMS: EXPLORING THE ROLE OF THE DOUBLE DEFINITE ARTICLE CONSTRUCTION .....	32
LANGUAGE ATTITUDES AND STYLISTIC VARIATION OVER TIME: RESULTS OF A PANEL STUDY FROM ULRICHSBERG (AUSTRIA) .....	33
REPLENISHING AN IMPOVERISHED INVENTORY: EVIDENCE OF ATTENUATIVE TYPOGRAPHICAL VARIATION FROM TEN EUROPEAN LANGUAGES.....	34
THE SYSTEMATICITY BEHIND US SPANISH VARIABLE GRAMMAR: THE CASE OF MOOD VARIATION IN SOUTHERN ARIZONA .....	35
EXPLAINING THE VARIANCE IN US SPANISH VARIABLE PATTERNS: A COMPARISON OF DIFFERENT OPERATIONALIZATIONS OF BILINGUAL EXPERIENCES.....	36

SPANISH LEXICAL BORROWINGS IN PORTUGUESE AMONG THE VENEZUELAN DIASPORA IN BRAZIL: A VARIATIONIST PERSPECTIVE .....	37
SPLIT ERGATIVITY AND LOSS OF RICH VERBAL AGREEMENT.....	37
PERCEIVING PROSODY: DIALECT CONTRIBUTIONS TO PERCEPTIONS OF PITCH AND SPEECH RATE IN MIAMI .....	39
SLUR RECLAMATION AND METALINGUISTIC DISCOURSE AMONG LGBTQ SPEAKERS.....	40
EFFECTS OF PHONOLOGICAL LENGTH AND MORPHOLOGICAL VARIATION IN CODE-SWITCHING PROCESSING AND REPRESENTATION .....	41
ORIENTATION TOWARDS THE VERNACULAR AND STYLE-SHIFTING AS LANGUAGE BEHAVIOURS IN SPEECH OF FIRST-GENERATION POLISH MIGRANT COMMUNITIES SPEAKING NORWEGIAN IN NORWAY.....	41
INVESTIGATING COMMUNICATION ACCOMMODATION AND ENDONORMATIVITY: THE CASE OF TH-STOPPING AND POSTVOCALIC /R/ IN MALAYSIAN ENGLISH .....	42
ASSESSING LITERACY IN THE HOME LANGUAGES OF IMMIGRANT STUDENTS .....	43
TALKING SUSHI: CROSS-LINGUISTIC VARIATIONS AND THE CONSTRUCTION OF A COSMOPOLITAN IDENTITY IN CHANGSHA, CHINA.....	44
WITHIN-GROUP TALKER VARIATION AND MEMORY ASYMMETRIES .....	44
VARIATION IN DIFFERENTIAL OBJECT MARKING IN COPALA TRIQUI .....	45
WHO SPEAKS TAIWANESE? LANGUAGE VARIATION, STANDARDIZATION, AND CREATION .....	47
THE ACTUATION OF UNSTRESSED /A/-RAISING IN MODERN HEBREW.....	48
WHEN ZE IS HIR GENDER: MORPHOSYNTACTIC APPROACHES TO NEOPRONOUNS IN ENGLISH .....	48
“THE GIRLS IN THE VALLEY CONCUR:” STANCETAKING AND LANGUAGE IDEOLOGIES REGARDING THE SPEECH OF JULIA FOX .....	50
PHONETIC VARIATION OF DURATION MEASURES FOR PALATAL PLOSIVES IN AUSTRALIAN LANGUAGES.....	51
THE NEW YORK CITY METRO AREA SURVEY: LANGUAGE ATTITUDES AND THE LOW BACK MERGER .....	52
IS Y’ALL FOR ALL? THE SOCIOLINGUISTIC PERCEPTION OF Y’ALL.....	53
VARIABLE ENREGISTERMENT AND VARIABLE STIGMA: SHIFTING INDEXES OF LOCALITY IN GALICIAN .....	53
PERCEIVED SOCIAL CONGRUENCE BETWEEN SOURCES OF SOCIAL INFORMATION ABOUT SPEAKERS .....	54
VARIATION IN HOMORGANIC CONSONANTS ACROSS ALBANIAN DIALECTS AND ITS IMPLICATIONS FOR THE HISTORY OF THE LANGUAGE: EVIDENCE FROM THE MALSIA MADHE DIALECT .....	55
ACTING DIFFERENT(-LY): BRINGING DERIVATIONAL MORPHOLOGY INTO VARIATIONIST LINGUISTICS.....	56
COMPARISON OF THE GRAMMATICALIZATION OF GENRE IN HEXAGONAL FRENCH AND SWISS FRENCH.....	56
NASALITY VARIATION IN GREATER MANCHESTER: A MULTI-DIMENSIONAL APPROACH TO QUANTIFYING SOCIOPHONETIC NASALITY ....	57
HERITAGE SPEAKERS CREATE SYNCHRONIC VARIATION AND FURTHER DIACHRONIC CHANGE: THE CASE OF SPANISH <i>SER/ESTAR</i> .....	58
LOCAL SOCIAL MEANINGS AND LANGUAGE USE IN A CHICAGO BILINGUAL COMMUNITY.....	59
ETHNIC BOUNDARIES AND SPEAKER IDENTITY: TH-STOPPING IN THE MONTREAL ANGLO-ITALIAN COMMUNITY .....	60
COMPLEMENTARY MEASURES IN THE ASSESSMENT OF CHANGE: A CASE STUDY OF FRENCH POLAR INTERROGATIVES.....	61
LEXICAL FREQUENCY AND IDIOSYNCRASY IN MORPHOSYNTACTIC VARIATION - EVIDENCE FROM SPANISH SUBJECT PRONOUNS .....	62
NEUTRAL TONE IN BEIJING: REEXAMINING THE UTILIZATION OF THE NEUTRAL TONE IN BEIJING PROFESSIONALS .....	63
VARIATION IN CASE MARKING ACROSS NUBRI VALLEY .....	64
VISIBLE ARTICULATORY VARIATION AS A CUE TO SOUND CHANGE: LIP ROUNDING AND LIP PROTRUSION VARIABILITY IN THE MANDARIN SIBILANT MERGER .....	64
METAPHOR AND METONYMY AS MARKERS OF REGISTER VARIATION .....	65
EXPLORING THE EFFECTS OF CROSS-CULTURAL VARIATION AND TOURISM ON THE PERCEPTION OF UTAH ENGLISH.....	66
COGNITION AND THE HUMAN SPEECH COMMUNITY: LARGE-SCALE EVIDENCE OF CROSSLINGUISTIC SIMILARITIES IN VARIABLE SUBJECT PRONOUN EXPRESSION .....	67
WHO BENEFITS? BILINGUAL EDUCATION PROGRAM ACCESS FOR IMMIGRANT COMMUNITIES IN NEW YORK CITY .....	68
CHANGING STRUCTURES OF COVARIATION.....	68
SARCASM SOCIALLY CONTEXTUALIZED: HOW SOCIO-INDEXICAL INFORMATION MODULATES THE EVALUATION OF PRAGMATIC BEHAVIOR.....	69
MOVING BEYOND PROSODIC TRANSCRIPTION IN THE ANALYSIS AFRICAN AMERICAN ENGLISH INTONATION: A CASE STUDY WITH BIN UTTERANCES.....	70
BIDIALECTAL 7-9-YEAR-OLDS’ EVALUATIVE LABELS FOR THE STANDARD AND LOCAL VARIETIES OF HUNGARIAN .....	71
THE SOCIAL MEANINGS OF DEFINITE ARTICLES WITH PROPER NAMES IN SPANISH .....	72
HERITAGE SPEAKERS’ CATEGORIZATION OF VARIABLE PHONETIC INPUT IN SPANISH.....	73
CROSS-LINGUISTIC VARIATION IN THE DOMAIN OF INTENTIONALITY AND ITS CONSEQUENCES FOR REAL-LIFE LEGAL OUTCOMES .....	73
HEDGES AND APOLOGIES IN CHATGPT RESPONSES TO AFRICAN-AMERICAN ENGLISH AND MAINSTREAM U.S. ENGLISH .....	74
SOME EARLY OBSERVATIONS OF AN INCIPIENT MERGER OF /I/ AND /Y/ IN GOTHENBURG SWEDISH .....	75

SEEKING A LIFE WITHOUT SHAME IN THE HAITIAN LINGUISTIC MARKET: UNDERSTANDING POPULAR RESISTANCE TO KREYÒL MOTHER TONGUE EDUCATION .....	76
GETTING SOCIALIZED: VARIATION AND CHANGE IN THE PASSIVE IN CANADA.....	77
INDIVIDUAL VARIATION IN THE PROCESSING OF FILLER-GAP DEPENDENCIES IN L2 ITALIAN .....	77
WORD-GENDER AND VOICE-GENDER PERCEPTIONS IN LGBTQ+ VS. CISHET LISTENERS.....	78
A NEW WAY OF ANALYZING VOWEL SPACE AREA .....	79
WORD-INITIAL /P/ IN KHALKHA MONGOLIAN: VARIATION IN CONNECTED SPEECH.....	80
RAPID AND INTROSPECTIVE PROCESSING OF SOCIOLINGUISTIC ASSOCIATIONS OF (ING) IN CONTEXT .....	81
TRANSMASCULINE ENTRAPPED VOCALITY: IDENTIFICATION AND IDENTITY .....	81
HERITAGE SPANISH WRITING COMPLEXITY AND ITS RELATIONSHIP TO PROFICIENCY AND WRITING GENRE .....	82
VARIABILITY OF LANGUAGE EVALUATION IN THE AMERICAN NORTH & SOUTH AS DESCRIBED BY THE OTHER REGION .....	83
REGIONAL VARIATION OF PHONOLOGICAL FEATURES IN AFRICAN AMERICAN LANGUAGE.....	84
INVESTIGATING LEXICAL VARIATION IN A MIXED LANGUAGE: A VARIATIONIST CORPUS-BASED ANALYSIS OF LÁNNANG-UÈ CONJUNCTIONS AND PREPOSITIONS .....	85
AGEING IN STYLE: TOWARDS DISENTANGLING STYLE SHIFTING AND LIFESPAN CHANGE.....	86
VOWELS, AFFECT, AND (DIS)AFFILIATION IN A NON-BINARY COMMUNITY.....	87
A LAYERING APPROACH TO DP2 AGREEMENT.....	88
CAREGIVER NARRATIVES: VARIATION IN THE INPUT AND CHILD AFRICAN AMERICAN LANGUAGE .....	90
DEVELOPING A CHILD-CENTRED APPROACH TO UNDERSTAND CHILDREN’S EXPLICIT SOCIOLINGUISTIC AWARENESS: HOW 9-11 YEAR OLD CHILDREN IN A LONDON SUBURB PERCEIVE SOCIAL MEANING IN TALK.....	90
EXPLORING THE DEFICIENCIES OF LARGE LANGUAGE MODELS IN USING AAL: AN INTERDISCIPLINARY STUDY .....	91
/O/! THEY'RE I/U/ST ABOUT THE SAME!: VOWEL SHIFT IN HERITAGE AND HOMELAND SEOUL KOREAN .....	92
EXPLORING THE VIRTUAL LINGUISTIC LANDSCAPE: AN ANALYSIS OF PRONOUNS OF ADDRESS IN HISPANIC AND LUSOPHONE UNIVERSITIES.....	93
SUBJECT PRONOUNS AND COMMUNICATIVE FUNCTION IN FIVE LANGUAGES .....	94
ACQUISITION OF VARIATION: ACQUIRING CONSTRAINTS ON VARIABLE CLITIC CLIMBING IN CHILDREN .....	95
STYLIZED PERFORMANCE OF “MOCK BERBER” IN A MOROCCAN STAND-UP COMEDY TALENT SHOW.....	96
CONVERSATIONAL COMFORT AND LINGUISTIC INSECURITY WITH RESPECT TO GENDER IDENTITY, SEXUAL ORIENTATION, AND TRANSNESS .....	96
INDIVIDUAL LANGUAGE VARIATION ACROSS DISCOURSE TYPES: A STUDY OF 112 IDIOLECTS .....	97
HOW DO YOU SAY PASTA?: FOREIGN (A) IN CANADA.....	98
(DH)ERE’S MORE THAN ONE WAY TO STRENGTHEN A FRICATIVE: VOICED INTERDENTALS IN A DIVERSE SAMPLE OF CALIFORNIANS .....	99
THE PRESENT INDICATIVE AND CONFIDENCE ABOUT THE FUTURE: PRESUMED SETTLEDNESS IN FRENCH AND ACROSS ROMANCE .....	99
ATTENTION TO SPEECH (VARIETIES) OR ATTENTION TO TASK? LOCATING STYLES IN ARABIC AND FRANCOPROVENÇAL .....	100
PERCEPTIONS OF "SOUTHERN" UTAH ENGLISH .....	101
WIGGLY LIFESPAN CHANGE IN A CRISIS – CONTRASTING REACTIVE AND PROACTIVE IDENTITY CONSTRUCTION .....	102
AN APPARENT-TIME STUDY OF DAEJEON KOREAN STOP LARYNGEAL CONTRASTS .....	103
VOT MERGER IN PROGRESS AND SPEECH RATE ACCOMMODATION IN PERCEPTION: A CASE STUDY OF DAEJEON KOREAN.....	104
“I COULD PROBABLY ASSUME WHAT YOU DID”: CHARACTEROLOGICAL CREATION IN IMPROVISED COMEDY.....	104
NE-OMISSION IN IVORIAN FRENCH CELEBRITY INTERVIEWS ON YOUTUBE .....	105
THE LAST L1 JUDEO-SPANISH SPEAKERS OF NEW YORK CITY.....	106
ANALYZING THE PERFORMANCE OF MALE SEXUALITY IN SEVILLIAN SPANISH THROUGH REMOTE SPEECH DATA COLLECTION .....	107
INTER- AND INTRA-SPEAKER /ai~ei/ VARIATION IN ‘ŌLELO HAWAĪI: ACOUSTIC AND HISTORICAL-COMPARATIVE EVIDENCE.....	108
MYRIAD ORIENTATIONS: SEXUALITY AND LOCALITY IN MANX ENGLISH.....	110
“WE WORK HARD AND GET NOTHEN FOR IT”[1]: REGIONAL VARIATION IN PIN/PEN MERGER IN A POST-RECONSTRUCTION CORPUS [1]	
REVEREND JOHN JACOBS, ESCAMBIA COUNTY, ALABAMA, OCTOBER 7, 1879.....	111
DOUBLE NEGATION IN RUSSIAN: VARIATION IN THE FUNCTIONAL STRUCTURE OF PPS .....	112
AN ANALYSIS OF UKRAINIANS’ LANGUAGE ATTITUDES AND IDEOLOGY: A CONFLICT-CATALYZED IDENTITY SHIFT.....	113
LONGITUDINAL CHANGE IN THE ADJECTIVE INTENSIFIER SYSTEM OF HEXAGONAL FRENCH .....	114
I NEED TO, FOR PERSONAL REASONS: SHIFTING FORMS AND FUNCTIONS IN AMERICAN ENGLISH OBLIGATION MODALS .....	115
SPANISH VOWEL AND CONSONANT CONTRIBUTIONS TO TALKER IDENTIFICATION AND LEXICAL CONTRAST .....	116
CONSTRUCTIONS OF WHITENESS IN COLLEGE STUDENTS’ CONVERSATIONS .....	117
LEXICAL EFFECTS IN THE CHANGE OF CANTONESE NEUTRAL QUESTION FORMS .....	118

SETTING {STRAIGHT} THE RECORD {STRAIGHT}: ACCEPTABILITY OF ALTERNATIVE WORD ORDERS IN RESULTATIVES BY HERITAGE CANTONESE SPEAKERS.....	119
PLACE, IDEOLOGY, AND IDENTITY: THE CONSTRUCTION OF AFRICAN AMERICAN LINGUISTIC IDENTITY IN JACKSON, MS .....	120
THE EFFECTS OF STYLISTIC COVARIATION IN PHONOLOGICAL VARIANT CHOICES .....	122
THE TAISHANESE BELTED L: ACOUSTIC FEATURES, LATERALIZATION FROM CANTONESE /s/, AND HERITAGE VS. NATIVE SPEAKERS..	123
FRENCH CLITICS: A PARADOX FROM VARIATION .....	123
YOU ARE NOT A NATIVE SPEAKER, BUT I WANT TO SOUND LIKE YOU: UNVEILING ALTERNATIVE VOT CONVERGENCE PATTERNS IN L2 ENGLISH SPEECH.....	125
VARIABLE ADAPTATION OF ENGLISH /ə/ IN QUEBEC FRENCH .....	126
“READ A NEWS REPORT AND MAKE A BLESSING”: ACCESSING THE SOCIAL MEANING OF THE PHARYNGEALS VARIANTS IN HEBREW.....	126
LANGUAGE-SPECIFIC PROSODY, CONVERGENCE, AND HYPERARTICULATION IN STATEMENTS OF PALENQUERO/SPANISH BILINGUALS .	127
AGE EFFECTS AND LANGUAGE DIFFERENTIATION FOR YES/NO QUESTION INTONATION IN TWO GROUPS OF PALENQUERO/SPANISH BILINGUALS .....	128
PRONOMINAL EXPRESSION IN CHILD HERITAGE GEORGIAN.....	129
A DYNAMIC NEURAL MODEL OF THE INTERACTION BETWEEN SOCIAL AND LEXICAL INFLUENCES ON SPEECH PRODUCTION: THE CASE OF RETROFLEX SIBILANTS IN TAIWAN MANDARIN. ....	130
REAL-TIME INTERPRETATION OF MORPHOSYNTACTIC VARIATION ACROSS VARIETIES AND LISTENERS .....	131
RACE IN THE PERFORMANCE OF A BALTIMORE PERSONA.....	132
COMMUNITY EMBEDDEDNESS, SOCIAL PRACTICE AND COGNITION: EXPLAINING WAS/WERE VARIATION IN AN ECONOMICALLY-DEPRIVED COMMUNITY.....	133
DISCOURSE/PRAGMATIC VARIATION IN THE TAG SYSTEM OF SOUTHERN ARIZONA SPANISH .....	134
SOCIAL PERCEPTIONS OF TRADITIONAL NEW MEXICAN SPANISH PARAGOGE ACROSS COMMUNITIES.....	135
SOCIOLINGUISTIC STUDY OF PARTIAL QUESTION STRUCTURES IN HEXAGONAL FRENCH: CONCLUSIONS FROM EXPERIMENTAL DATA ...	136
PEDRO CASTILLO, LA SIERRA, AND LA COSTA: INTRA-SPEAKER VARIATION AND DYNAMIC BRICOLAGE IN PERUVIAN VOWEL SPACES ..	137
AGE AND GENDER-MEDIATED PATTERNS IN SOCIOLINGUISTIC PERCEPTION – COMPARING US AND UK LISTENERS’ EVALUATIVE REACTIONS TO TYNESIDE VOICES.....	137
VOICING EXPERTISE: EXPLORING STRATEGIC USE OF VOWEL VARIANTS IN THE ENGLISH PRONUNCIATION OF CHINESE LANGUAGE INSTRUCTORS .....	138
LINGUISTIC PRIDE OR ASSIMILATION? AN EXAMINATION OF LEXICAL CHANGE, LINGUISTIC PRACTICES, AND CONSTRUCTION OF ETHNIC IDENTITY AMONG CHAOUI FRIENDSHIP NETWORKS IN BATNA CITY, ALGERIA.....	139
LEARNING TO BE (UN)HIP IN PANEL DATA - EXPLORING QUOTATIVE BE LIKE ACROSS THE ADULT LIFESPAN.....	140
EXPLORING VARIATION IN HERITAGE TAMIL RETROFLEX PERCEPTION AND PRODUCTION.....	141
“MY SAFE WORD WILL BE [m]ISKEY!” – AN ACOUSTIC APPROACH TO THE WHINE-WINE SPLIT.....	143
ONE BOOK WATCHES THE OTHER: USING MULES AND MEN TO DETERMINE DISTINCTIVE FEATURES OF BARRACOOON.....	144
IN SEARCH OF “LITERARY” TAIWANESE: NEGATION IN A TAIWANESE HAND PUPPET THEATER PERFORMANCE .....	144
LINGUISTIC INNOVATION OR ANCESTRAL FEATURE? THE CASE OF TSUZAMEN AND TSAM IN HASIDIC YIDDISH.....	145
D1 FEATURE MAINTENANCE IN A D2 REGION: OLDER NEW YORKERS IN TORONTO .....	146
A SAN DIEGAN DEBATE: TAKE 8 OR TAKE THE 8, INVESTIGATING THE USE OF THE DETERMINER “THE” BEFORE NUMERIC FREEWAYS IN SAN DIEGO.....	147
IT’S GIFING GAY: REACTION IMAGES AND GIFS ON GAY TWITTER .....	147
FINDING ORDER (OR NOT): HIERARCHICAL DENSITY BASED CLUSTERING OF NEW YORKERS’ VOWELS. ....	148
THE SOCIAL EMBEDDING OF LECTAL COHERENCE AMONG MOBILE SPEAKERS .....	149
INVESTIGATING THE LINK BETWEEN INTELLIGIBILITY AND ATTITUDES TOWARDS VARIETIES OF FRENCH.....	150
NORTHERN MAINE AS A TRANSITION ZONE: EVIDENCE FROM RHOTICITY IN SOUTHERN AROOSTOOK COUNTY ENGLISH.....	151
REGRESSIVE VOICING ASSIMILATION IN MODERN GREEK DIALECTS .....	151
SOCIAL AND PHONOLOGICAL FACTORS IN (R) VARIATION IN 1930S INLAND MASSACHUSETTS .....	152
IS MAXENT CUMULATIVITY TOO RESTRICTIVE? EVIDENCE FROM A VARIETY OF NYC AAE.....	153
CODE-SWITCHING DATASETS FOR NLP: USING INTONATION FOR ELEMENTARY UNITS .....	154
SOCIOLINGUISTIC VARIATION IN AMERICAN SIGN LANGUAGE IN WASHINGTON STATE.....	155
SPANISH RHOTIC PRODUCTION AND DIALECTAL VARIATION IN MEXICAN HERITAGE SPEAKERS.....	156
VARIATION AT THE LEXICON-SYNTAX INTERFACE IN SPANISH: DIALECTAL DIFFERENCES IN COPULAR SENTENCES WITH ADJECTIVAL PREDICATES .....	157
VARIABLE REALIZATIONS OF CORONAL PLOSIVES IN A FRENCH URBAN YOUTH VERNACULAR: A CASE STUDY.....	158

(HERITAGE) RUSSIAN CASE-MARKING: VARIATION AND PATHS OF CHANGE .....	159
“MISSION DIYARI”: POLYSEMY & CONTACT-INDUCED SEMANTIC CHANGE .....	160
AN OCEAN APART: PRONOUN CHOICES IN TURKISH HERITAGE SPEAKERS .....	161
THE SURPRISING PATHWAYS OF CHANGE: LESSONS FROM THE VERNACULAR .....	161
CODESWITCHING VARIATION IN CHANKA QUECHUA: PALINDROMIC SWITCHES .....	162
OLD WINE IN NEW BOTTLES: TOKYO VOWEL DEVOICING .....	163
ARE THERE REALLY PIRATES IN BEIJING? PERCEPTUAL DIALECTOLOGY OF RHOTICIZATION IN MANDARIN .....	164
THE INTERSECTION OF ETHNICITY AND SOCIAL CLASS IN LANGUAGE VARIATION AND CHANGE .....	164
THE IMPACT OF SOCIAL INFORMATION ON VOT SHADOWING BY NONBINARY SPEAKERS .....	165
YOU BEST BUTTON [bʌrən] IT UP: AN AMERICAN ENGLISH CHANGE-IN-PROGRESS WITH PEOPLE IDENTIFIED AS BLACK TAKING THE LEAD .....	166
MORPHOLOGICALLY AND TOPICALLY CONDITIONED [t] – [k] VARIATION IN OLELO NIIHAU .....	167
CO-OCCURRENCE AND ABSENCE OF FEATURE: INVESTIGATING COMBINATORIAL EFFECTS OF STEREOTYPICAL SOUTHERN FRENCH FEATURES ON SPEECH PERCEPTION .....	167
SOCIOPHONETIC VARIATION IN BRETON: ANALYZING THE EFFECTS OF SOCIAL FACTORS, LANGUAGE CONTACT, AND SPEAKER ATTITUDES .....	168
YOU KNOW I’M FROM THE CHI: IDENTITY CONSTRUCTION, NEGOTIATION, AND PLACEMAKING AMONG BLACK CHICAGOANS .....	169
EXAMINING VARIATION AND GRAMMATICALIZATION OF THE ENGLISH EXPRESSION A COUPLE OF IN THE CANADIAN PRAIRIES.....	170
“¡A[ʎ]Á NO, PERO AQUÍ USAMOS LA ELLE!”: ARTICULATION AS A SOCIOLINGUISTIC SIGN OF IDENTITY IN BOLIVIA.....	171
FROM ‘CUCK’ TO ‘PROSTITUTE’: CROSSLINGUISTIC VARIATION OF DEKASEGI IN BRAZILIAN PORTUGUESE AND JAPANESE ONLINE DISCOURSE.....	172
AGE AND OTHER SOCIAL FACTORS IMPACTING GENERIC THIRD PERSON PRONOUN USE IN NONNATIVE ENGLISH SPEAKERS .....	173
A CULTURE OF LABOUR: INDEXING ‘BLUE-COLLAR’ THROUGH A LOWER /s/ COG IN MANITOBA, CANADA.....	174
EMBODYING ENERGY: CORRELATING PITCH AND GESTURE IN PERFORMANCES OF SOCIAL PERSONAE.....	175
DECODING SOCIAL INEQUALITY IN HISTORICAL DOCUMENTS. UNRAVELING TRACES IN THE EARLY MODERN ERA.....	176
“I DON’T EVEN TAKE NOTE OF THE BAD ENGLISH LYRICS”: THE LISTENING SUBJECTS OF INTERNATIONAL K-POP .....	177
WITCHCRAFT AND THE HISTORICAL PARADOX: ANALYSING SENTIMENT IN SALEM WITCH TRIAL PROTOCOLS.....	178
SISSY THAT TALK! A VENEZUELAN DRAG QUEEN’S SPANISH RHOTICS.....	178
"OH SURE." THE FINAL PERIOD IN TEXTING AS AN ‘ANTI-EXCLAMATION’ MARK .....	179
VARIATION IN /st/ CLUSTER PRODUCTION OF WESTERN ANDALUSIAN SPANISH .....	180
RESISTING THE FALL OF DEMOCRACY: THE LINGUISTIC LANDSCAPES OF PROTEST SPACES IN ISRAEL, 2023 .....	181
FACTORS CONDITIONING VARIABLE ASSIMILATION IN KOREAN NASAL-LIQUID SEQUENCES .....	181
THE MEANING AND FUNCTION OF CODE-SWITCHING IN A MIXED LANGUAGE: LANGUAGE CONTACT AND VARIATION IN MOLISE CROATIAN (ITALY) .....	182
IT’S NOT THE THOUGHT THAT COUNTS FOR A LOT OF MICHIGAN ENGLISH SOUND CHANGES .....	183
SOCIAL VARIATION IN THE USE OF TRAER AND TENER AS VERBS OF POSSESSION IN MEXICAN SPANISH .....	184
REGIONAL SOUND CHANGE IN A MINORITIZED COMMUNITY: SOCIOPHONETICS OF AFRICAN AMERICAN ENGLISH IN ALBANY, NY...	185
LANGUAGE EXPERIENCE IMPACTS L2 ENGLISH SCOPE COMPUTATION .....	186
DO WOMEN REALLY USE INTENSIFIERS MORE FREQUENTLY THAN MEN? A CROSS-LINGUISTIC AND CROSS-VARIETY STUDY .....	187
PREVELAR RAISING AND MERGER IN CALIFORNIA: A SOUND CHANGE IN PROGRESS.....	187
MONOPHTHONGAL /ow/ AMONG NORDIC AMERICANS IN PUGET SOUND: LANGUAGE CONTACT IN THE DEVELOPMENT OF REGIONAL DIALECT FEATURES .....	189
THE ‘1984’ OF LINGUISTIC CHANGE: A SOCIOLINGUISTIC SHOCK POINT IN THE LATE 20TH CENTURY .....	190
PITCH VARIABILITY CUES PERCEPTIONS OF SINGLISH: A PERCEPTUALLY-GUIDED APPROACH TO SOCIOPHONETIC VARIATION .....	190
MULTIPLE WAYS TO DO AUTHENTICITY: A CASE STUDY OF A DIASPORIC SPEAKER.....	191
VARIABLE EQUIVALENCE FOR BILINGUAL CODE-SWITCHING: SPANISH-ENGLISH CLAUSE COMBINING .....	192
DETERMINER DROP IN VERNACULAR ENGLISH.....	193
BETTER LATE THAN NEVER: BILINGUAL LANGUAGE PROCESSING OF EVIDENTIALITY IN A TURKISH-ENGLISH TRANSLATION TASK .....	194
REVISITING THE CONCEPT OF “TRIGGERING” OF CODE-SWITCHING.....	195
CAN THE VARIABLE USE OF ENGLISH AND CANTONESE IN A DIGITAL MEDIUM BE PREDICTED? ANALYZING WHATSAPP CODE-SWITCHING IN HONG KONG COMPUTATIONALLY.....	196
AM/P~OM/P MERGER IN HONG KONG VS. TORONTO CANTONESE: AN UNDER-DOCUMENTED HOMELAND SOUND CHANGE IN A HERITAGE LANGUAGE CONTEXT .....	197



THE DECLINE OF RHOTICITY IN LANCASHIRE, NORTHERN ENGLAND: DATA FROM ULTRASOUND AND SOCIOLINGUISTIC INTERVIEWS ...	198
STABILITY IN THE FACE OF CONTACT: EVIDENCE FROM TORONTO HERITAGE TAGALOG /U/ .....	198
VARIATION IN THE ACCEPTABILITY OF SINGULAR THEY IN SINGAPORE ENGLISH .....	199
NEGATIVE MARKING AS A SOCIOLINGUISTIC VARIABLE AMONG PALESTINIAN SPEAKERS IN BEIRUT .....	200
QUANTIFYING GRAMMATICALIZATION VIA TRANSLATION: THE POSITION OF LO CUNTO DE LI CUNTI'S NEAPOLITAN ON THE ROMANCE GRAMMATICALIZATION CLINE .....	201
EXPLORING THE INTERPLAY OF LITERACY AND SOCIOECONOMIC STATUS IN LANGUAGE PROCESSING: A FIELD-BASED STUDY .....	202
REGION > AGE > GENDER: ON THE ROLE OF GEOGRAPHICAL AND SOCIAL FACTORS ON (MORPHO-)SYNTACTIC VARIATION IN THE GERMAN DIALECTS OF AUSTRIA.....	203
MAKING A SPLASH :THE SALIENCE OF LIQUID VARIATION AMONG BOSTON SPANISH SPEAKERS.....	204
QUANTITATIVE PERSPECTIVES ON REGIONAL VARIATION FROM THE MAP-BASED ACCENT-RECOGNITION TASK: THE CASES OF DUTCH AND FRISIAN .....	204
MACRO AND MESO APPROACHES TO QUEER LINGUISTICS: VARIATION IN NON-BINARY PRONOUNS IN DUTCH.....	205
SOCIOLINGUISTIC RESEARCH PROJECTS AS BRANDS .....	206
"IN MY REAL VOICE?" THE ROLE OF ELICITATION TASK ON THE SPEECH OF AMERICAN AND BRITISH TRANSATLANTIC MIGRANTS.....	207
ON THE MORPHOPHONOLOGICAL PRODUCTIVITY OF /v/ GLIDE INSERTION FOLLOWING UNDERLYING FINAL /U/ IN DATIVE CASE SUFFIXATION OF MODERN KANNADA NOUNS .....	208
PREACHING TO THE CHOIR: WHAT ALL LINGUISTS NEED TO KNOW ABOUT DEFINING AND OPERATIONALIZING ETHNICITY AND RACE IN RESEARCH .....	208
EN LA ZONA DONDE ELLO[H] [B][V]IAN: DIALECTAL AND LANGUAGE CONTACT OUTCOMES ON LIMITED SPIRANTIZATION AND LABIODENTALIZATION OF /B/ IN THE PRODUCTION OF US SALVADORANS .....	209
WHY ARE -ING AND -IN' PROCESSED DIFFERENTLY? EXPLORING A PROCESSING ASYMMETRY BETWEEN CANONICAL AND NONCANONICAL SOCIOLINGUISTIC VARIANTS .....	210
SOCIAL IDENTITY AND SELF-PERCEPTION IN THE LANGUAGE CLASSROOM: A PRACTICAL APPROACH.....	211
MULTIRACIAL IDENTITIES AS LOCALLY CONSTRUCTED STYLES.....	211
TAKING CENTER STAGE: MEASURING VARIATION AND CHANGE ACROSS THE CENTER AND PERIPHERY OF LEXICAL FIELDS.....	212
THE FLOATING BUBBLE: LINGUISTIC INNOVATIONS OF THE HIGHLY MOBILE THIRD-CULTURE KIDS (TCKs).....	213
LANGUAGE SHIFT AND MORPHOSYNTACTIC CHANGE IN TGDAYA SEEDIQ.....	214
THREE-WAY VARIANCE IN SCOPE JUDGMENTS IN MANDARIN .....	215
"THE BUREAUCRAT", "THE NEIGHBORHOOD GRANDMA", AND "THE CHILDISH GIRL" IN BEIJING: UNCOVERING THE SOCIAL MEANING OF STYLISTIC TIMING VARIATION.....	217
NASALIZATION CHANGE OVER TIME IN MICHIGAN ENGLISH.....	217
MODELING SOCIOPHONETIC VARIATION IN L2 VARIETIES OF ENGLISH: CONSIDERATIONS OF PROFICIENCY AND NORM ORIENTATION .	218
SOCIAL PERCEPTION AND CATEGORIZATION OF SOUTHERN MANDARIN ACCENTS .....	219
SOCIAL ATTRIBUTE RATINGS OF MANDARIN VARIETIES IN DIFFERENT COUNTRIES .....	220
EXAMINING RHOTICITY PATTERNS IN L2 ENGLISH ACQUISITION: A COMPARATIVE STUDY OF CHINESE SPEAKERS FROM NORTHERN AND SOUTHERN REGIONS OF CHINA .....	221
ARTIFICIAL ACCENTS: ASSESSING PHONOLOGICAL VARIATION IN VOICE CLONING SOFTWARE .....	222

## Plenaries

---

### Revisiting contact and switching in bilingual New York from a speaker-centered translanguaging perspective

Ricardo Otheguy  
Graduate Center, City University of New York  
NWAV / RISLUS

The Latino population of New York City (NYC) provides the linguist with an opportunity to study in its formative stages the grammatical consequences of people contact (n.b., not the opportunity to study language contact). These consequences can involve differences between grammars of impacted bilinguals in NYC and grammars in Latin America. But there has not been sufficient attention paid to the questions of how exactly one determines where and when these differences arise, and especially of why they arise in the first place.

A study is sketched of subjunctive and progressive forms in the grammars of Latinos in NYC. The study stresses the distinction between cultural and grammatical differences, and advances quantitative predictions regarding sub-populations where one can observe greater cultural, and consequently larger grammatical, distance between speakers in NYC and Latin America.

The correct analysis requires a recognition of cross-cultural differences of conceptualization of the same or equivalent referent in the different precursor (pre-contact) settings. These cross-cultural differences of conceptualization hold even in cases where the precursor grammars have congruous meanings of lexical and morphosyntactic forms.

The different conceptualizations of equivalent referents constitute the trigger for, and therefore the explanation of, grammatical differences, even though they do not constitute themselves a grammatical difference. When the different conceptualizations can be expressively wrangled out of precursor meanings successfully, no grammatical innovation obtains in impacted bilinguals. When precursor meanings are insufficient to the task, the meanings of precursor forms can be observed to be different in speakers in the people contact setting in NYC. The overall picture is one where the notion of “language contact” has led to considerable overstatement of the grammatical differences between Latin America and NYC.

A connection is offered between the explanation based on cross-cultural differences of conceptualization of referents and the supporting translanguaging view of a single representation for the two languages of bilinguals. The translanguaging position de-licenses the articulation of such familiar notions as interference, convergence, transfer, and code switching. For the interest of applied linguistics, the analysis offers a picture of NYC bilinguals, including the people who have come to be known as heritage speakers, that is not entirely equivalent to the prevailing account.

---

### Crosslinguistic Variation in Article Systems

Veneeta Dayal  
Yale University

A question that continues to intrigue linguists is this: how can the article system of languages that have them be so complex, when so many languages manage easily without them? One plausible line of explanation is to think of languages without articles as having phonologically null counterparts of definite and indefinite articles. Another is to think of nouns without articles as being underspecified and therefore usable in contexts that allow the definite as well as in contexts that allow the indefinite. On either view, lack of articles is predicted not to lead to lack of expressive power. But if there is no functional pressure to do so, why do languages develop articles at all (the definite from the demonstrative, the indefinite from the numeral ‘one’)?

Taking reference to kinds as a central piece of the puzzle, I focus on the fact that languages vary along two dimensions, number as well as (in)definiteness marking. A particularly dramatic demonstration of this comes from English (I stay neutral on whether bare nouns have a null D):

*Tigers are an endangered species.*                      Pl Number;              No overt D (\*Def D/\*Indef D)

*The tiger is an endangered species*                      Sing Number;              Def D (\*Indef D, \*No overt D)

Object-level statements also show variation in article use in a way that interacts with number. Illustrating again from English – the plural licenses KIND  $\rightarrow \exists$  reading, the singular does not:

*Tigers* are roaming outside.     $\exists$  reading = some tigers

*The tiger* is roaming outside.    \* $\exists$  reading  $\neq$  a/some tiger

Looking at languages without articles from the perspective of kind-reference leads to questions that I have explored in previous work: how are kind-reference and (in)definiteness in the ordinary domain related? is there cross-linguistic variation wrt  $KIND \rightarrow \exists$  reading along the dimension of number? are there cross-linguistic differences in the  $\exists$  readings of bare nominals?

Based on a recent study of seven typologically unrelated languages, spoken in geographically diverse regions, I argue against the standard view that bare nouns in article-less languages cover the same semantic space as definite and indefinite nouns in articulated languages. In a sizeable majority of article-less languages, bare nominals are sensitive to the size of their N-set (in the singular case, they require it to be a singleton) though not to its discourse status (the referent may be novel or familiar). This means that bare nouns cannot function like specific indefinites. In the absence of an indefinite article, languages draw on the numeral ‘one’ (or ‘some’) for this purpose, just like specific indefinites in articulated languages require the indefinite article or the numeral ‘one’ (or ‘some’).

Against this cross-linguistic backdrop about the form-meaning map for bare nouns, and using reference to kinds as a lens, I look at what has been said about article use in New Englishes.

---

## Is variationist sociolinguistics first-world problem-solving? Quantitative theory/methodology and African languages

John Victor Singler  
New York University

William Labov’s foundational work in establishing variationist sociolinguistics is grounded in industrialized society. There is, however, nothing a priori about that work or the field that it launched that requires a comparable social setting in order for its insights to emerge. Nonetheless, the application of quantitative methodology to African languages is rare. In Stanford and Preston’s *Variation in indigenous minority languages* (2009), only one of the 21 languages in the collection is African.

I begin my talk by reviewing (1) existing work on African languages with a variationist orientation, (2) studies that are variationist in direction yet statistically limited in scope, and (3) work that, after carefully considering internal and external factors, stopped too soon, thereby missing out on what quantitative analysis can provide.

My central question is why linguists who study African languages in society undertake variationist studies so rarely. Building on Salami (1986), I consider the practical reasons for this. However, I also consider obstacles that are located within theory. Here I return to the questions that Nagy and Meyerhoff raise (2008) about variationist sociolinguistics vis-à-vis multilingualism, what they term “the curious monolingual bias of sociolinguistics” (p. 2). Moving forward, I draw on Meyerhoff’s (2022) and Mesthrie’s (to appear) recent work. I see a fundamental parallel to Aboh’s (2023) challenge to the centrality of monolingualism in formal theory.

- Aboh, Enoch O. 2023 *Universal multilingualism: Why we are born to code-mix*. The Sapir Lecture. LSA Institute, University of Massachusetts.
- Mesthrie, Rajend. to appear. *Variation in multilingual communities: An exploration*. In Yoshiyuki Asahi, Alexandra D’Arcy and Paul Kerswill (eds.), *Routledge handbook of variationist sociolinguistics*. London: Routledge.
- Meyerhoff, Miriam. 2022. Baby steps in decolonizing linguistics: *Urban language research*. In Paul Kerswill and Heike Wiese (eds.), *Urban contact dialects and language change: Insights from the global North and South*, 145-157. Milton, England: Taylor and Francis.
- Nagy, Naomi, and Miriam Meyerhoff. 2008. Introduction. In Miriam Meyerhoff and Naomi Nagy (eds.), *Social lives in language—sociolinguistics and multilingual speech communities: celebrating the work of Gillian Sankoff*, 1-16. Amsterdam: John Benjamins.
- Salami, L. Oladipo. 1986. Prospects and problems of urban sociolinguistic survey in Africa: Notes from Ile-Ife, Nigeria. *Anthropological Linguistics* 28 (4): 473-483.
- Stanford, James N., and Dennis R. Preston (eds.). 2009. *Variation in indigenous minority languages*. Amsterdam: John Benjamins.

---

## SYNTAX Workshop Plenary

### Variation in English negation revisited

Gary Thoms, New York University

An important way in which syntactic analysis feeds into sociolinguistics is to determine which syntactic elements count as “the same” in the grammatical system. Although this might at first appear to be a simple matter of getting the basic analysis right, it becomes increasingly important as we take steps towards formulating a concrete and psychologically plausible theory of how variation is regulated by speakers. Consider for example the variation

between adverbial and contracted negation in mainstream varieties of English, which is known to be subject to stylistic constraints:

- a. The dog was not barking.  
The dog wasn't barking.

One analysis, these are simply two pronunciations of a single syntactic element, which we might identify as an adverb in the specifier of a NegP projection. On an alternative, they are distinct syntactically: the full *not* is an adverb, but the contracted version is the head of the NegP projection. These two analyses locate the variation in different places in the grammatical system (variation in phonological realization vs variation in syntactic features), and as we build up our understanding of how one kind of variation might differ from the other more broadly, these analytical decisions matter. To sketch one possible state of affairs, if we were to find in case after case that variables which show strong sociolinguistic conditioning (e.g. style shifting) were best analysed as variation in phonological realization, then that might steer us in the direction of a pronunciation-based analysis of the variable in (1). And yet sometimes seemingly simple syntactic data might discourage such an analysis, such as contrasts of the following kind.

- a. \*Was not the dog barking?  
Wasn't the dog barking?

A typical analysis of this contrast is that in (2b) contracted negation forms a complex head with the finite auxiliary and this complex head formation feeds subject-auxiliary inversion, while there is no such complex head formation in (2a), since the negative is not a NegP head.

In this talk (drawing upon joint work with David Adger, Caroline Heycock, E Jamieson and Jennifer Smith), I make the case for a more holistic and micro-comparative approach to framing analyses of morphosyntactic variables, with English negation as a case study. My focus is a superficially similar facet of variation in Scots, where the mainstream *-n't* in (1b) varies with the Scots-specific negative element *-nae*, illustrated in (3). An interesting aspect of this negative element, which is described as a contracted negation in traditional descriptions, is that it largely patterns with *not* rather than *-n't* with respect to inversion, in that it doesn't participate in inversion, (4).

- (3) The dog wasnae barking. (4) \*Wasnae the dog barking?

Despite this superficial disparity, I show, drawing upon data from the *Scots Syntax Atlas*, that *-n't* and *-nae* pattern together in many other ways, and I argue that a more complete understanding of the variation we find across Scots varieties weighs in favour of a unifying analyses of (3) and (1b). I thus argue that defining the envelope of variation cannot be done without consideration to the broader microcomparative picture.

## Talks and Posters

---

### Diaspora identity and language in the Malayalee-American community

Maya Ravindranath Abtahian<sup>1</sup>, Savithry Namboodiripad<sup>2</sup>

<sup>1</sup>University of Rochester, Rochester, USA. <sup>2</sup>University of Michigan, Ann Arbor, USA

Malayalee-Americans have heritage in Kerala, India (where Malayalam is the official language). Linguistically and culturally Malayalee-Americans differ from an image of “Indian” as popularly imagined in US culture, including the fact that many Malayalees orient towards English rather than the national language Hindi (Namboodiripad 2021). At the same time alignment with the imagined community (Anderson 1983) of the Indian/South Asian diaspora may facilitate use of the lingua franca Hindi. We investigate the nuances of Malayalee-Americans’ alignment with/against this homogenized diasporic identity in their language choice and use of ethnic identifiers (e.g. terms like *desi*, ABCD, and Indian-American vs Malayalee, Malayalee-American, Mallu) using a survey, modeled on ethnic-orientation (Hoffman and Walker 2010) and language-use questionnaires (Cohn et al. 2013). The survey includes questions about respondents’ social networks, participation in South Asian social groups; use of ethnic identifiers (above); use of English, Malayalam, and other languages; and choice of labels for images of “homey” food items like rice, yogurt, and tea. The project speaks to generational and social differences in the linguistic construction of diasporic identity, with the understanding that minoritized groups have access to multiple linguistic means to index ethnic identity (Dubois and Horvath 1998; Newmark et al 2016). Here we focus on use of the ethnic identifier *desi* and investigate how participants’ association with *desi* as an ethnic identity marker predicts the language origin of their choice for eight food terms.

Respondents include immigrant (1gen) and US-born (2gen) Malayalees (N=148, ages 18-81). We find evidence of language shift, with Malayalam being used more often with older generations than younger (81% with grandparents vs 31% with children). We also find evidence of the emergence in apparent time of certain ethnic identifiers: for instance, participants who said they self-identified as *desi* were more likely to be younger than participants who did not (Fig. 1). For food terms, there was considerable item-specific variation: 4/8 items had a clear community-wide preference, but not in the same direction (“Malayalam”-origin *chaaya* was overwhelmingly preferred to *chai*: 74% vs 10%, while “English”-origin vegetables was preferred by 75% of participants). Focusing on the four variable lexical items, we hypothesized that “Hindi”-origin terms would be more likely to be chosen by participants who identified as *desi*. This was borne out for yogurt, water, and pickle (Fig. 2); those who chose “Hindi”-origin terms for ‘yogurt’ and ‘water’ were in the minority and uniformly identified as *desi*. These results speak to the importance of a focus on linguistic variables beyond language choice in the construction of ethnic identity.

Previous work in this area has largely dealt with more long-standing communities of 3rd and 4th generation Americans (Chun 2001, Cashman 2018, Bucholtz 2004, Hinrichs and Tseng 2021, Wagner 2013); none has considered South Asians. This “newer” community offers an opportunity to examine the generational processes at play in the construction of ethnic identity as hyphenated Americans, with the ultimate goal of characterizing these different imagined identities in the context of the diaspora, immigration, and language shift.

Anderson, B. 1983. *Imagined communities*. London: Verso.

Benor, S.B.. 2010. Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics* 14(2): 159-183.

Bucholtz, M. 2004. Styles and stereotypes: The linguistic negotiation of identity among Laotian American youth. *Pragmatics* 14: 127–147.

Chun, E. 2001. The construction of white, black, and Korean-American identities through African American Vernacular English. *Journal of Linguistic Anthropology* 11: 52–64.

Cohn, A. C., et al. 2013. Kuesioner Penggunaan Bahasa Sehari-hari (Daily language use questionnaire) <http://lingweb.eva.mpg.de/jakarta/kuesioner.php>

Dubois, S. and B. Horvath. 1998. Let’s tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change* 10: 246-261.

Hinrichs, L. and A. Tseng. 2021. Mobility, polylingualism, and change: Toward an updated sociolinguistics of diaspora. *Journal of Sociolinguistics* 25: 649-661.

Hoffman, M. and J. Walker. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22: 37-67.

- Namboodiripad, S. 2021. Non-Dravidian elements and (non)diasystematic change in Malayalam. Chapter in *Constructions in Contact 2: Language change, multilingual practices, and additional language acquisition*. Benjamins. Höder, S. & Boas, H., eds. <https://doi.org/10.1075/cal.30.07nam>
- Newmark, K., N. Walker, and J. Stanford 2016. 'The rez accent knows no borders': Native American ethnic identity expressed through English prosody. *Language in Society* 45(5): 633-664.
- Wagner, S. 2013. We act like girls and we don't act like men: Ethnicity and local language change in a Philadelphia high school. *Language in Society* 42(4): 361-383

---

## Sociosyntactic Variation in Singular They Reflexives

Kirby Conrod<sup>1</sup>, Byron Ahn<sup>2</sup>

<sup>1</sup>Swarthmore College, Swarthmore, PA, USA. <sup>2</sup>Princeton University, Princeton, NJ, USA

There is socially and syntactically determined variation in the acceptability of singular *they* (henceforth “ST”) pronouns, especially in the context of nominals that are definite and specific (“dST”; [6], [10], [11], a.o.). This paper explores variation of the reflexive forms of ST: *themselves* and *themselves*. Using two large-scale internet-based acceptability judgment tasks, we uncover four patterns of acceptability, which can broadly be described in terms of a microparametric typology (with respect to  $\phi$ -features in reflexive anaphors in dST contexts) while also showing effects of linguistic and social ideologies for this change-in-progress.

**Background/Motivation:** English number is marked through a single privative [SG] feature, and a given nominal lacking [SG] (e.g., *you*) can be compatible with singular or non-singular antecedents while agreeing with verbs as “plural” (e.g., [4]). English *they* lacks a [SG] #-feature (e.g., [6], [10], [11]), allowing there to be a single *they* (that can be “singular” or “plural” and uniformly exhibit plural verb agreement). If, as presumed by many (e.g. [3]), reflexive anaphors had a single bundle of  $\phi$ -features, the word *themselves* would suggest a [SG] *them* and *themselves* would suggest a non-[SG] *them*. However, the premise of this statement is faulty: English reflexive anaphors consist of two nominals ([1], [2], [7]) each with  $\phi$ -features that can be distinct from each other and/or the antecedent ([9]). Thus we expect ‘*themselves*’ (non-[SG] *them* + [SG] *self*) and ‘*themselves*’ (matching in non-[SG]) as both in-principle possible with [SG] antecedents. This leads to our research question about ST reflexive anaphors: when do speakers accept *themselves* and when do they accept *themselves*?

**Methods/Results:** Study 1 (n=1,123) gathered acceptability data for hypothesis generation, while Study 2 (n=222) tested those hypotheses with a rigorous experimental design. To test for effects of linguistic variables, stimuli varied by antecedent-type (e.g., quantified NP, proper name, etc.) and anaphor form (*themselves* or *themselves*). To test for social variation, participants also completed social surveys about their macrosocial characteristics as well as linguistic and social ideologies. Additionally, to test for distinct patterns of ratings behaviors (“dialects”), we ran k-means cluster analyses on the acceptability ratings (following methods of [12]).

The results of Study 1 suggested that a primary site of variation was in stimuli with ‘familiar’ antecedents (proper names, definite specific NPs): “dST contexts.” Ratings clusters differed with respect to whether *themselves*, *themselves* and *themselves*, or neither were acceptable in dST contexts. On the basis of these suggestive findings, we hypothesized that there would be a fourth type of behavior in dST contexts: a group preferring *themselves*. Results suggested social effects of prescriptivist ideologies about language and gender (lowering overall ratings of *themselves* and *themselves*) and being trans or nonbinary (improving overall ratings of *themselves* and *themselves*).

The results of the clustering analysis showed that ‘familiar’ antecedents (i.e., dST contexts) were the primary locus of variation (see figure) across clusters, and that the clusters differed on whether the participant preferred *themselves*, *themselves*, either, or neither in such contexts. With respect to social variation, across antecedent types, ratings of *themselves* improved with being more prescriptivist ( $p < 0.001$ ) or further to the political right ( $p = 0.006$ ). In models with social variables interacting with antecedent type and *-self/-selves*, the only effects found to be significant were in dST contexts. As suggested by Study 1, more transphobic participants rated dST *themselves* and dST *themselves* lower ( $p = 0.014$  &  $p < 0.001$ , respectively). Departing from Study 1’s suggestions, we found no effects of age or prescriptivism on dST *themselves*, and we found dST *themselves* was rated lower as participants got older ( $p = 0.006$ ) and higher as they got more prescriptivist ( $p < 0.001$ ).

**Discussion/Conclusions:** While we found social effects across the four clusters, social variables do not reliably predict membership in one cluster over another (using Random Forest modelling). This is suggestive that clusters reflect different grammars that are in competition during this change-in-progress for the form(s) of the reflexive anaphor used for a dST context (which is the primary locus of variation). The group with the largest difference between acceptability of the two forms is the group that prefers *themselves*, possibly suggesting that this group has a grammar that resembles the (settled) English grammar that governs usage of *yourself/yourselfs*, while others are less settled in which form is preferred (if either) in dST contexts; we expect this will stay in flux for quite some time, as all groups had young speakers in them, meaning this domain should be investigated further for years to come.

Notable for binding theories is that every participant accepted some type of  $\phi$ -feature mismatch between the three nominals involved in reflexive binding (antecedent, pronoun, *-self*). This means any analysis of English binding that assumes or derives complete matching of  $\phi$ -features must be revised (see also [5]). Instead, we argue that these “dialect” groups vary on which  $\phi$ -features must match, in which contexts. (In dST contexts, a dST.cons grammar requires the antecedent to #-match the pronoun; a Tself.better grammar requires *them* to #-match with *-selves*; a Tself.better grammar requires the antecedent to #-match with *-self*; an all.dST.Good grammar places no apparent #-matching constraints.) Additionally, this work raises new conceptual arguments against positing gender  $\phi$ -features on English nouns (see also [8], [13]). A key argument for having gender  $\phi$ -features on English common noun comes from contrasts like ‘*a girl ... herself/#himself*.’ However, this argument is built on the premise that English binding mandates  $\phi$ -feature matching — a premise that this study undermines.

Turning to sociolinguistic interpretations of this data, we note that this variation isn’t driven solely by macrosocial categories or social ideologies, but by an intersection of those social variables with one’s own grammatical configurations. This is witnessed by the effects of social variables being limited to dST contexts (where we identified the grammatical variation), and also by the fact that the effect sizes for all of this social variation were noticeably smaller than those for the linguistic variables (antecedent type and *-self/-selves*) alone. Second, the effects of social variables are complex and not as might be obviously expected, highlighting the need for deeper sociolinguistic investigations. We tentatively suggest that the choice between forms may be indexically linked to particular personas, and that that might be complicating the effects of the social variables. This seems especially likely, given that this is a change from above for many, but we leave this to be investigated in future studies.

- [1] Postal, P. M., 1966. On so-called ‘pronouns’ in English. In F. P. Dinneen, ed., *Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies*, 177–206. Georgetown University Press.
- [2] Helke, M., 1971. *The Grammar of English Reflexives*. Ph.D. thesis, MIT.
- [3] Hornstein, N., 2001. *Move! A Minimalist Theory of Construal*. Malden, MA: Blackwell Publishing.
- [4] Sauerland, U., J. Anderssen & K. Yatsushiro, 2005. The Plural Is Semantically Unmarked. In *Linguistic Evidence — Empirical, Theoretical, and Computational Perspectives*, 409–430. Mouton de Gruyter.
- [5] Hicks, G., 2009. The Derivation of Anaphoric Relations, *Linguistik Aktuell/Linguistics Today*, vol. 139.
- [6] Bjorkman, B., 2017. Singular *they* and the syntactic representation of gender in English. *Glossa*, 2(1).
- [7] Ahn, B. & L. Kalin, 2018. What’s in a (English) Reflexive? In *NELS 48*, vol. 1.
- [8] Ackerman, L., 2019. Syntactic and Cognitive Issues in Investigating Gendered Coreference. *Glossa*, 4(1).
- [9] Ahn, B., 2019. Features, Identity, and ‘Yourself’. In *NELS 49*, vol. 1.
- [10] Conrod, K., 2019. *Pronouns Raising and Emerging*. Ph.D. thesis, UW.
- [11] Konnelly, L. & E. Cowper, 2020. Gender diversity and morphosyntax. *Glossa*, 5(1).
- [12] Camilliere, S., A. Izes, O. Leventhal & D. Grodner, 2021. *They* is Changing: Pragmatic and Grammatical Factors that License Singular *they*. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43.
- [13] Conrod, K., 2022. Abolishing Gender on D. *CJL*, 67(3).

---

## An Interdisciplinary Approach to Investigate Language Choices in a Trilingual Setting

Carina Ahrens

University of Duisburg-Essen, Essen, Germany. Ruhr University Bochum, Bochum, Germany

Psycholinguistic experiments are designed to provide insights into how the human brain processes language. However, speech situations in labs differ greatly from socially-situated everyday conversations, where many factors influence the linguistic choices speakers make (un-)consciously as they interact with different interlocutors in different situations while talking about a variety of topics (e.g., Trudgill 1974; Rickford and Price 2013).

Following Tucker and Ernestus’ (2016) call, I combine casual speech data collected in sociolinguistic interviews and careful speech data gathered in a controlled environment (i.e., in a laboratory) to assess the extent to which speakers’ linguistic choices pattern differently in controlled environments vs. naturalistic conversations.

A corpus consisting of sociolinguistic interviews and cognitive tests from multilingual speakers will be compiled to investigate how social factors interact with the processing of language across levels of linguistic representation in

multilingual brains. Participants will answer questionnaires about their language acquisition history (e.g., years spent learning, motivation, time abroad). The participants are L1-speakers of French who studied English and German at school for at least eight (English) and six years (German), and who currently live and study in Germany. They will take part in sociolinguistic interviews and cognitive tests twice a year over at least 2-3 years (for a total of 4-6 recordings).

I hypothesize that the influence of German at university and German and French in participants' day-to-day lives will influence their production of English and, thus, code-switching depending on language proficiency and amount of use/activation is expected (e.g., Hammarberg 2015). This phenomenon will be assessed on different linguistic levels. On the lexical level, I will test reaction time in a cognate test: Participants will be asked to click on the corresponding picture of an English cognate word presented on a screen (e.g., Elefant (German), éléphant (French)). These results will be compared to code-switching in the sociolinguistic interviews to see whether, and to which language participants switch while aiming to use their non-dominant language (English). The morphological and phonological levels will be investigated through a nonsense-word test (as used by Ulicheva et al. 2020) and again, results will be compared to the speakers' language production in the sociolinguistic interviews. Moreover, the study aims to track phonetic variation in a trilingual setting over time: A sociophonetic analysis will be conducted of sounds that are exclusive to the English sound inventory to see whether speakers are changing their pronunciation and whether it is influenced by the surrounding language German or their L1 French. Linear mixed models which take the social factors into consideration to determine their influence on language production in a multilingual setting will be fitted to each language level.

Comparing results from a psycholinguistic experiment to extemporaneous speech from sociolinguistic interviews and including social factors (age, gender, motivation to acquire the languages, intention to stay in Germany), allows me to explore the extent to which lab-based language production and perception are comparable to semi-naturally occurring speech. The analysis will thus add important insights into how the English language is processed in multilingual brains.

Hammarberg, B. (2012), 'Activation of L1 and L2 during production in L3: A comparison of two case studies', *Processes in Third Language Acquisition* (Edinburgh, 2005; online edn, Edinburgh Scholarship Online, 22 Mar. 2012), <https://doi.org/10.3366/edinburgh/9780748635115.003.0006> [javascript%3A], accessed 31 May 2023.

Rickford, J. and Price, M. (2013), 'Girlz II women: Age-grading, language change and stylistic variation.' *J Sociolinguistics*, 17: 143-179. <https://doi.org/10.1111/josl.12017> [javascript%3A], accessed 31 May 2023.

Trudgill, P. (1974), 'The social differentiation of English in Norwich'. Cambridge: CUP.

Tucker, B. V. and Ernestus, M. (2016), 'Why we need to investigate casual speech to truly understand language production, processing and the mental lexicon.' *The Mental Lexicon* 11(3):375-400. <http://dx.doi.org/10.1075/ml.11.3.03tuc>, [javascript%3A] accessed 31 May 2023.

Ulicheva, A., Harvey, H., Aronoff, M., Rastle, K. (2020), 'Skilled readers' sensitivity to meaningful regularities in English writing.' *Cognition* 195. <https://doi.org/10.1016/j.cognition.2018.09.013>, [javascript%3A] accessed 31 May 2023.

---

**All of (us/we) are interested in Tyneside English first-person plural objects**  
Aziz Alenezi, Louis-Geoffrey Gousset, Jiaqi Liu, Maisie Woodcock, Daniel Duncan  
Newcastle University, United Kingdom

Several British English varieties use noncanonical object pronouns. While this is typically seen in limited contexts with the first-person singular, Tyneside English stands out for its robust usage of noncanonical objects in the first-person more generally, across a wide range of grammatical contexts (Beal 2004). Despite observable intraspeaker variation (1), previous studies often view this usage as categorical (Beal 2004) or focus more on variability in the first-person singular (Childs et al. 2021). In this paper, we turn to variability in the first-person plural object. Drawing on data from the Diachronic Electronic Corpus of Tyneside English (DECTE, Corrigan et al. 2012), we illustrate patterns in pronoun usage concerning linguistic and social factors, and show that what previously was a prototypical sociolinguistic variable has taken on a more localized indexical meaning in recent decades.

DECTE is an approximately 800,000-word corpus of dyadic and monadic sociolinguistic interviews collected with 161 residents of the North East of England. First-person plural objects (n=582) were collected using AntConc (Anthony 2020), and were manually cleaned to exclude instances of singular *us* and subject *we*. We coded for language-internal factors of grammatical context (2) and whether the object was accompanied by an overt quantifier to indicate plurality (3). Language-external factors included speakers' social class, education level, sex, and age. Age was operationalized as both age when recorded and period of recording, as DECTE was constructed to support both real- and apparent-time analyses.

The noncanonical object was quite uncommon overall (n=115, 19.3%). We used logistic mixed effects regression to determine which factors condition *us/we* usage (all reported results significant at  $p < 0.05$ ). Non-age fixed effects,



alongside speaker and person coding the token as random intercepts, were included in a base model, with age group, period of recording, and interactions between social factors added using AIC as a guide. Both language-internal predictors affect production; overt quantifiers favor we, while objects of prepositions disfavor it. Compared to the baseline of female, middle-class university students recorded in the 2000s, lower middle-class speakers and men disfavor we. However, interaction terms show that men in older recordings favor we, as do LMC speakers who did not obtain a university degree. Participants in dyadic interviews were more likely to use we when paired with another we user, suggesting accommodation effects.

That this variable is in fact variable and conditioned by language-internal predictors is itself a novel contribution to the field. More noteworthy perhaps is the social conditioning. In the two oldest subcorpora, Tyneside first-person plural objects pattern much like a prototypical sociolinguistic variable as, for example, male speakers favor the noncanonical we. However, in the most recently collected subcorpus this is turned on its head: highly educated, female, middle-class speakers favor the noncanonical variant. This suggests that like other Tyneside features such as stop glottalization (Buchstaller et al. 2022), what was an overtly stigmatized local variant is seeing its indexical meaning changing to be repurposed by groups who historically avoided it in order to perform localness and in-group solidarity in interactions between Tynesiders

Anthony, Laurence. 2020. *AntConc* (Version 3.5.9). Tokyo, Japan: Waseda University.  
<https://www.laurenceanthony.net/software>.

Beal, Joan. 2004. "English Dialects in the North of England: Morphology and Syntax". In Bernd Kortmann, and Edgar Schneider, eds. *A Handbook of Varieties of English*, 114–141. Berlin: Mouton de Gruyter.

Buchstaller, Isabelle, Adam Mearns, Anja Auer, and Anne Krause-Lerche. 2022. Exploring age-related changes in the realisation of (t) : Panel research from Tyneside. *English-World-Wide* 43(3): 297-329.

Childs, Claire, Carmen Llamas, and Dominic Watt. 2021. Pronoun exchange in North-Eastern English: Perception versus use. Paper presented at UKLVC13, University of Glasgow.

Corrigan, Karen P., Isabelle Buchstaller, Adam Mearns and Hermann Moisl. 2012. *The Diachronic Electronic Corpus of Tyneside English*. Newcastle University. <https://research.ncl.ac.uk/decte>.

---

## Ethnicity-specific and -independent learning: Effects of guise on adaptation to novel foreign-accented talkers

Nicholas Aoki, Georgia Zellou  
University of California, Davis, Davis, USA

**Background:** How do listeners adapt to foreign-accented talkers? Prior work shows that hearing multiple Mandarin-accented English speakers induces accent-specific learning: comprehension improves for a novel Mandarin-accented talker, not a Slovakian-accented talker [1]. However, other studies demonstrate that hearing multiple speakers from different language backgrounds increases comprehension for a novel-accented talker (accent-independent learning) [2]. Taken together, these mixed findings support theories that listeners attend to systematic (accent-specific or -independent) similarities across talkers and use these to generalize to new talkers (i.e., the ideal adapter framework [3]).

Prior work has focused on voice similarity across talkers. In the current study, we explore whether top-down, social cues impact cross-talker generalization. Extensive research demonstrates that social expectations can influence speech perception [4]. Comprehension of a Mandarin-accented English speaker is higher when paired with an East Asian face than a Caucasian face, ostensibly because matching acoustic and visual information enhances understanding [5]. By examining how apparent ethnicity affects cross-talker generalization, we investigate how listeners leverage social information in perceptual learning.

**Study and Hypotheses:** Listeners first transcribed sentences in noise [6] for three male Mandarin-accented talkers (exposure phase) and then for a novel male Mandarin-accented talker (test phase). Two competing accounts were compared for the effect of apparent ethnicity on transcription accuracy for the novel talker. A visual-acoustic congruency prediction is that intelligibility should increase when congruent bottom-up (voice) and top-down cues are presented (image). A top-down-only prediction is that generalization is only facilitated when talkers share top-down cues. When listeners view faces of the same ethnicity in exposure, learning should only generalize when the test talker shares the same ethnicity (ethnicity-specific learning). Additionally, viewing faces of different ethnicities should result in generalization for a talker with a novel ethnicity (ethnicity-independent learning).

**Methods and Results:** 236 native English speakers were assigned to either the “No Guise”, “Asian Exposure + Asian Test”, “Mixed Exposure + Asian Test”, or “Caucasian Exposure + Asian Test” conditions (Table 1). Transcription accuracy for the novel talker was assessed with a mixed-effects logistic regression model [7] containing a fixed effect of Condition with random intercepts for Listener and Sentence.

Figure 1 displays the results. Relative to the No Guise condition, accuracy improved in the Asian Exposure + Asian Test ( $p = 0.01$ ) and Mixed Exposure + Asian Test ( $p = 0.009$ ) conditions. The No Guise and Caucasian Exposure + Asian Test conditions did not differ ( $p = 0.24$ ).

Conclusion: Results support top-down only accounts. Viewing faces of the same ethnicity in exposure resulted in ethnicity-specific learning: generalization was facilitated when the test talker ethnicity matched (Asian Exposure + Asian Test) and blocked with a mismatch (Caucasian Exposure + Asian Test). Additionally, seeing faces with multiple different ethnicities led to ethnicity-independent learning, or improved accuracy for a talker with a novel ethnicity (Mixed Exposure + Asian Test). Contra visual-acoustic congruency accounts, accuracy was reduced in the Caucasian Exposure + Asian Test condition, even though an Asian face was also shown at test.

This study demonstrates that apparent ethnicity can affect cross-talkers generalization, highlighting the importance of social expectations in perceptual learning.

- [1] Bradlow, A. R., Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition* 106, 707-729.
- [2] Baese-Berk, M., Bradlow, A. R., Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *J. Acoust. Soc. Am.* 133, EL174-180.
- [3] Kleinschmidt, D. F., Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychol. Rev.* 122, 148–203.
- [4] Campbell-Kibler, K. (2010). Sociolinguistics and Perception. *Lang. Linguist. Compass* 4, 377-389.
- [5] McGowan, K. B. (2015). Social Expectation Improves Speech Perception in Noise. *Lang. Speech* 58, 502–521.
- [6] Kalikow, D. N., Stevens, K. N., Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J. Acoust. Soc. Am.* 61, 1337-1351.
- [7] Bates, D., Mächler, M., Bolker, B. M., Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67, 1-48.

---

## Discreteness in a Creole-Standard continuum? A Perceptual and Attitudinal Approach to the 'named languages' debate

Chiara Ardoino

Queen Mary University of London, London, United Kingdom

Bilingual communities are a well-known site of inter-speaker variation and fluid linguistic practices. When viewed through the prism of speaker-centred third-wave approaches, such inter-speaker variation can be so prominent to lead scholars to question the very existence of languages as psychological objects (Otheguy, García & Reid, 2015).

One can expect fluidity and inter-speaker variation – and, thus, language ‘porousness’ – to be particularly rife in minoritised languages that have undergone little/no standardisation. A case in point are Creole-Standard bilingual settings, many of which have been described as ‘continua’ because of the supposed absence of hard boundaries between the Creole and the Standard.

Bilingualism and lack of standardisation, however, do not necessarily exclude the psychological reality of languages and language boundaries, as even speakers of un(der)standardised and closely related languages have been shown to consistently detect language switching and mixing in speech (Lipski, 2019, 2020).

To bypass the rift between structural approaches, predicated on the notion of language boundaries, and qualitative third-wave approaches that question such boundaries, this paper tests (i) speakers’ perception of language boundaries and (ii) whether such perception is likely to be a by-product of standardisation and standard language ideology, as claimed by ‘language-sceptical’ approaches.

Adopting as a testing ground the French department of Martinique – where French is spoken alongside a still largely unstandardised French-lexicon Creole – I have used a questionnaire to investigate:

1. to what extent speakers are sensitive to the crossing of boundaries between French and Creole. Respondents have listened to sentences in 5 conditions (nominally French, mixed French, code-switched, mixed Creole and Creole) and rated them on a 5-point Likert scale ranging from fully French to fully Creole;
2. whether such sensitivity is stronger, and inter-speaker agreement higher, for French – the more standardised of the two languages, and
3. whether reporting of language boundaries crossing correlates with measures of speakers’ purism, collected through Likert-scale statements such as ‘When one speaks French/Creole, one should avoid using expressions that are clearly Creole/French’.

Results from 158 respondents paint a mixed picture. On the one hand, French appears to be a slightly more bounded code than Creole, with clearer differences between the ratings of nominally mixed and unmixed stimuli and higher inter-speaker agreement. On the other hand, it is actually Creole that triggers more purist responses from the sampled population, and reported purism is a better – albeit not strong – predictor of speakers’ perceptions for Creole than for French stimuli.

Through an innovative design that combines the investigation of perceptions and attitudes, this paper presents additional evidence on the question of language boundedness and its relation to standardisation/purism. First, purism can be even stronger for un(der)standardised than for fully standardised languages – perhaps because the former are perceived as more endangered. Secondly, even under-standardised varieties like Martinican Creole can be perceived as distinct and bounded codes – albeit with a higher degree of inter-speaker variation. In conclusion, degree of standardisation and purism need not be directly correlated, and neither is necessarily a strong predictor of perceived boundedness.

Lipski, John M. 2019. Field-Testing Code-Switching Constraints: A Report on a Strategic Languages Project. *Languages* 4,7.  
Lipski, John M. 2020. *Palenquero and Spanish in Contact: Exploring the Interface*. Amsterdam/Philadelphia: John Benjamins.  
Otheguy, Ricardo, Ofelia García, Wallis Reid. 2015. Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review* 6: 281-307.

---

## From Diglossia to Creole Revival and Changing Indexicalities: A Perceptual and Attitudinal Testing of the Martinican (Semi-)Continuum

Chiara Ardoino<sup>1</sup>, Noémie François-Haugrin<sup>2</sup>, Stéphane Térosier<sup>3</sup>

<sup>1</sup>Queen Mary University of London, London, United Kingdom. <sup>2</sup>Université des Antilles, Schoelcher, Martinique. <sup>3</sup>Leiden University, Leiden, Netherlands

The traditional Creole-standard continuum model rests on the assumption of an equivalence between the structural/linguistic and the social dimensions of language variation – whereby the term basilect, for instance, refers to the variety that is simultaneously (i) most distant from the Creole’s lexifier and (ii) used by the working classes in the least formal contexts (Souprayen-Cavery & Simonin, 2013).

In the French department of Martinique, the representation of French and Creole as poles of a socio-stylistic continuum does not factor in recent societal changes. While French has become most Martinicans’ mother tongue (Beck, 2017), Creole has gained access to formal domains once the preserve of French. Martinique’s transition to a less diglossic bilingualism raises the question of whether the continuum model still has heuristic value or if, on the contrary, Creole has developed a new formality continuum of its own.

More precisely, this paper investigates:

1. how activists’ basilect – purportedly ‘distanced’ from French to establish Creole as a fully-fledged language – is perceived by Martinicans. Is it interpreted as less formal (because of its structural proximity to old rural Creole) or more formal (because of its rarity and deliberate avoidance of French borrowings)?

2. how French and Creole compare in terms of speakers’ representations/attitudes. Has Creole developed indexicalities that go beyond the solidarity values typically associated with non-standard languages?

To address these questions, we have used a 2x2 questionnaire design that tests speakers’ perceptions of Creole sentences in two structural (+/- Creole) and two stylistic (+/- formal) conditions. We have also probed the relation between speakers’ evaluations and their sociolinguistic profiles (e.g. purism and exposure to activist Creole), and tapped into Creole indexicalities by gauging representations of and attitudes towards Creole.

Findings from 123 respondents paint a paradoxical situation. On the one hand, more basilectal Creole is often perceived as more formal than its French-sounding counterpart, and reported purism is higher for Creole than for French. On the other hand, even respondents reporting high Creole purism and positive status-related attitudes associate Creole with traditionally low-prestige, rural environments.

This paper makes both methodological and theoretical contributions. Methodologically, it proposes a new perceptual approach to the study of Creole/dialect-standard continua – one that taps into neglected dimensions of linguistic variation which corpus-based approaches can only touch on indirectly (Bellonie & Pustka, 2019).

Theoretically, it tells a cautionary tale about language-wide indexical change. While reports of high purism and positive attitudes to Creole on the status dimension expose the need to update traditional diglossic models, the persistence of traditional indexicalities warns us against equating Martinican Creole to a fully standardised language like French. Admittedly, the paradox outlined above could just be a temporary sign of ongoing indexical change or, else, stem from the fact that some questions trigger more stereotypical answers than others. However, it could also indicate the persistence of ‘authenticity’ values (traditional life, localness) in the face of ongoing language standardisation (Urla, Amorrortu, Ortega et al., 2016) and, thus, the need to accommodate greater indexical complexity in descriptions of minoritised language standardisation.

Beck, B. 2017. La transmission des langues en Martinique, en Guadeloupe et à La Réunion. *Cahiers québécois de démographie* 46(2) : 241-262.

Bellonie, J. D., E. Pustka. 2019. Représentations des « mélanges » linguistiques en Martinique : des créolismes au français régional. *Études créoles. Cultures, langues, sociétés* 36(1-2).

- Souprayen-Cavery L., J. Simonin. 2013. Continuum linguistique. In J. Simonin, S. Wharton (eds.) Sociolinguistique du contact. Dictionnaire des termes et concepts, 123-141. Lyon: ENS Éditions.
- Urla, J., E. Amorrortu, A. Ortega, J. Goirigolzarri. 2016. Authenticity and linguistic variety among new speakers of Basque. In V. Ferreira, P. Bouda (eds.) Language documentation and conservation in Europe, 1–12. Honolulu: University of Hawai'i Press.

---

## A study on bare prepositions: a potential innovation in Brazilian Portuguese

Julia Bahia Adams

University of Campinas, Campinas, Brazil

Most variationist works deal with cases of stable variation or changes in progress, while few have analyzed the beginning of a change. Bare preposition placement in Brazilian Portuguese, in structures as in (1a) and (2a), is a possible innovation in its earliest stages of implementation, which, to our knowledge, hasn't yet been thoroughly described. In analyzing variation in relative clauses, Tarallo (1983) doesn't mention the stranded variant. Conversely, Poplack et al.'s (2012) study on bare prepositions in Quebec French indicates that orphaning in Romance languages may not be idiosyncratic. Structures with bare prepositions, though still infrequent, can currently be observed in both spoken and written Brazilian Portuguese, in alternation with the standard variant (1b), the prepositional-phrase chopping variant (1c, 2b), and the resumptive pronoun variant (1d, 2c). Although the actuation problem (Weinreich; Labov; Herzog, 1968) is the most obvious question, a preliminary step to the definition of a variable is circumscribing the variable context (Buchstaller, 2009). This study, thus, sought to define the syntactic, lexical, and pragmatic constraints to bare prepositions.

Considering the issue of data scarcity in syntactic variation and written data on social media as a rich source for studying variation, the data consist of a primary sample of 20,000 Tweets (approximately 400,000 words), harvested through the rtweet package (Kearney, 2018), and secondary samples of 163 sociolinguistic interviews from the Projects SP2010 (Mendes; Oushiro, 2013) and LínguaPOA (UFRGS, 2015-2019), and the newspaper Folha de São Paulo.

The qualitative analysis focused on the prepositions *sobre* 'about', *de* 'of', *com* 'with', and *para* 'to/for'. Semantic equivalence was established based on the concept of speaker presupposition (Stalnaker, 2002, 2014), in contrast to the truth-value of utterances. The variable context was defined based on Oushiro's (2011) proposal: whether the structures factually alternate considering the concept of communicative competence (Hymes, 1979), that is, if they are possible, feasible, adequate, and employed in the same set of contexts.

We found 180 tokens of bare prepositions in the Twitter corpus and 13 in the LínguaPOA sample. The results suggest that the main criterion for their occurrence is the presence of common ground (Stalnaker, 2002, 2014): there must be shared background information among interlocutors so that they can retrieve the referent of the preposition's object. We find that bare prepositions are always possible in relative clauses, where common ground is constantly present, thus we conclude that preposition stranding is a fourth relativization strategy in Brazilian Portuguese. The data also reveal that these phenomena have occurred more in non-relative clauses (N=189) and with the prepositions *sobre* (N=191) and *com* (N=2). The semantic weight of prepositions (Poplack et al., 2012) appears to be a predictor variable, as *sobre* and *com* are context-independent whereas *de* and *para* are semantically weak and could be dropped.

This research thus contributes to discussions on syntactic variation and the actuation problem, as structures with bare prepositions in Brazilian Portuguese are a possible change in progress at its earliest stage which can be described from its inception.

- BUCHSTALLER, I. (2009). The Quantitative Analysis of Morphosyntactic Variation: Constructing and Quantifying the Denominator. *Language and Linguistics Compass* 3/4: 1010–1033.
- HYMES, D. (1979). On communicative competence. In: BRUMFIT, C. J.; JOHNSON, K. (Eds.). *The Communicative Approach to Language Teaching*. Oxford: Oxford University Press. 5–26.
- KEARNEY, M. (2018). rtweet: Collecting Twitter data. Comp. software. Version v0.6.8.9000. [S.l.]. Available from: <https://doi.org/10.5281/zenodo.2528481>.
- MENDES, R. B., OUSHIRO, L. (2013). Documentação do Projeto SP2010 – Construção de uma amostra da fala paulistana. Accessed on: 31 aug. 2022. Available from: <http://projetsp2010.fflch.usp.br/producao-bibliografica>.
- OUSHIRO, L. (2011). Uma análise variacionista para as interrogativas-Q. 2011. MA thesis – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo. DOI: 10.11606/d.8.2011.tde-14102011-135709.
- POPLACK, S., ZENTZ, L., DION, N. (2012). Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence. *Bilingualism: Language and Cognition*, 15(2), 203–225.
- STALNAKER, R. (2002). Common Ground. *Linguistics and Philosophy*, v. 25, 701–721.
- STALNAKER, R. (2014). *Context*. 1. ed. New York: Oxford University Press.
- TARALLO, F. (1983). Relativization strategies in Brazilian Portuguese. PhD thesis – University of Pennsylvania, Philadelphia. 288f. Available from: <https://repository.upenn.edu/dissertations/AAI8326337>.

---

## Navigating Varieties of Arabic in Literacy Assessment

Alkhodor Ayach, Nishta Trivedi  
Multilingual Literacy SIFE Screener

Arabic is a term that encompasses a large variety of spoken and written languages; there are 375 million Arabic-speakers in the world. This presentation describes written and spoken varieties of Arabic, from Classical and Modern Standard Arabic to the different varieties of present-day spoken Arabic across the Arabic-speaking world, touching upon their characteristics, use, and mutual intelligibility. We emphasize the differences between Modern Standard Arabic, a formal variety that is not a native language, and the varieties spoken throughout the Arabic-speaking world.

As part of our work developing online semi-adaptive home language skill assessments for Arabic speakers, one challenge is to make a single assessment that is appropriate for students from throughout the Arabic-speaking world. This presents a particular challenge because students are taught to read in Modern Standard Arabic but at home, speak varieties that may be very different from Modern Standard Arabic. Students therefore must learn vocabulary and grammar of Modern Standard Arabic in school as part of their early elementary education. This presentation describes some of the ways in which we have created assessments of early literacy and foundational reading skills, reading comprehension, and math to serve all students who report Arabic as their language of education.

One of the issues we explore is how textbooks of mathematics look different depending on which Arabic-speaking country students learn math in. Some Arabic-speaking countries use the “Eastern Arabic” or “Indo-Persian” mathematical writing system [٩٨٧٦٥٤٣٢١٠], while others use the “Western Arabic” or “Hindu-Arabic” system [9876543210]. Moreover, number pronunciation can vary across different spoken varieties of Arabic. Additionally, there are differences in mathematical notation, such as directionality and decimal notation. Equations are read from right to left in the Arabic-speaking countries using the “Eastern Arabic” system, and from left to right in the countries using the “Western Arabic” system. However, starting in 6th grade, equations in Western Arabic are printed from left to right. We have devised a way to show the math equations using a display that will be appropriate for all students.

---

## A potential phonological change in an under-studied speech community: Tunisian Tamazight variety of Fahmine.

Soubeika Bahri  
University of Colorado Denver, USA

The goal of this study is to describe the use of /f/ instead of /t/ as in uflay ‘speak’ (elsewhere in Tunisian Tamazight: uṭlay) and instead of /k/ in the personal pronouns of 2nd person prefixed to the verb, like in a f-d-asey “I will come to you” (elsewhere in Tunisian Tamazight a k-d-asey, ta fen-uflayey “I will speak to you” (elsewhere in Tunisian Tamazight ta ken-uṭlayey) in Fahmine village, a multilingual Amazigh speech community located in Djerba island in the Gulf of Gabes governorate off Tunisia’s southern coast (Brugnatelli, 2023). The study attempts to adopt a variationist sociolinguistic approach, which with a few exceptions (e.g., Sanford 2016, Ding et al. 2019, Abtahian 2020), has rarely engaged with endangered or minority languages in non-Western contexts (Guy 2022).

By focusing on the variations in Fahmine sub-variety of Tunisian Tamazight (mostly known as Tunisian Berber), the study raises three main questions: (i) what linguistic factors motivate the use of /f/ instead of /t/ and /k/; (ii) is there a style-shifting effect triggering the use of /f/ in the aforementioned situations; and (iii) how is this phonological variation socially meaningful? Data were gathered from ten Djerbi Amazigh folktales that were read and recorded on YouTube by a member of the Fahmine speech community as part of his documentation project on Tunisian Tamazight. All the folktales were transcribed and translated into English for analysis yielding 13 tokens of uflay ‘speak’ and 29 tokens of /f/ in the personal pronouns of the 2nd person prefixed to the verb. The same community member participated in a classic sociolinguistic interview; the data considered for this study is from the informal conversation portion. Two Tamazight speakers from the villages of Guellala and Sedouikech, also located on the island of Djerba, took part in informal interviews for comparison purposes.

The findings reveal evidence of phonological change in the Tamazight variety spoken by the Fahmine community suggesting a shift away from Tunisian Tamazight varieties towards a local-oriented one. The findings also show the effects of style on the distinction of the phonemic variant in the speech community in question, as well as suggest that Tamazight speakers in the Fahmine speech community are linguistically and creatively engaged in constructing a distinct Amazigh indigenous identity through a set of phonological features.

Adli, A., & Guy, G. R. (2022). Globalizing the study of language variation and change: A manifesto on cross-cultural sociolinguistics. *Language and Linguistics Compass*, 16(5-6), e12452.

Abtahian, M. R. (2020). Style, identity, and language shift. *Language Ecology*, 4(1), 17-38.

Brugnatelli, V. (2023). Fathi Ben Maammar, Tinfas seg Jerba - Hikāyāt amāzighiyya jarbiyya. *Journal of Amazigh Studies*, 1(1). Available at <https://scholarship.claremont.edu/jas/vol1/iss1/5>

Ding, S. L., Wu, C. C., & Goh, K. L. (2019). In quest of a new identity? Language variation in Sabah. *Lingua*, 227, 102703.

Stanford, J. N. (2016). A call for more diverse sources of data: Variationist approaches in non-English contexts. *Journal of Sociolinguistics*, 20(4), 525-541.

---

## Not all sound changes in progress are used in early-adolescent stancetaking

Adam Barnhardt

Michigan State University, East Lansing, USA

Because language change is hypothesized to be driven by adolescents (Weinreich, Labov & Herzog, 1968; Eckert, 1988; Labov, 2001) and spread through stylistic use (Eckert 2011, 2019; Sneller, 2020), determining exactly how adolescents stylistically employ changes in progress is vital to the understanding of language change itself.

In this paper, I take a close look at the linguistic behavior of one white, middle-class, early-adolescent female Michigander. I examine her pronunciation of two types of sound change in progress during her performance of child stances vs. teen stances, specifically asking whether she uses more progressive pronunciations during teen stances (competent, cool, innocent, weak) and more conservative pronunciations during child stances (incompetent, uncool, mature, tough). Data come from self-recorded "audio diaries" from my speaker (MCD-00261), submitted regularly between the ages of 11 and 13 as part of the MI Diaries project (Sneller, Wagner & Ye, 2022).

The first type of sound change examined is the rapid reversal of Northern-Cities-Shifted (NCS) LOT and pre-oral TRAP, which are both backing and lowering in the community (Wagner et al., 2016; Nesbitt et al., 2019; Nesbitt, 2021a). Despite evidence that NCS LOT and pre-oral TRAP are emergent sociolinguistic markers (Sneller, 2012; Nesbitt, 2021b), and despite previous work finding preadolescents using an ongoing vowel change to enact child vs. teen stances (Eckert, 2011), I surprisingly find that my speaker does not change her pronunciation of these two vowels while enacting child vs. teen stances. Following the discussion in Eckert (1994) about socially useful performances of "white womanhood" for white, middle-class girls, I hypothesize that the reason LOT and pre-oral TRAP are not leveraged in stancetaking is that neither NCS vowel pronunciations, nor the pronunciations they are changing to, carry relevant social meanings for MCD-00261.

To test this hypothesis, I also investigate MCD-00261's use of creak in teen vs. child stances. Creak is a highly remarked-on feature found to be indexically linked to a socially relevant feminine persona similar to that of the valley girl (Eckert, 2008; Villarreal, 2018; Yuasa, 2010). A series of one-sample t-tests suggests that creak is in fact leveraged by my speaker to index teen vs. child stances. She uses creak at a lower rate when performing incompetent vs. competent stance ( $T = -12.767$ ,  $P < 0.001$ ), mature vs. innocent ( $T = -3.7644$ ,  $P < 0.001$ ), and tough vs. weak stance ( $T = -15.566$ ,  $P < 0.001$ ). She does not vary her use of creak between uncool vs. cool stances ( $T = 0.9143$ ,  $P = 0.3606$ ).

Taken together, my findings indicate that in order for a change in progress to be used by early-adolescents in age-related stancetaking, it must be indexically connected to a persona that is socially relevant for that speaker. These findings also introduce a mystery: How can changes below the level of awareness or with irrelevant social meanings spread, as we know LOT and TRAP are in Michigan?

Eckert, P. (1988). Adolescent social structure and the spread of linguistic change. *Language in Society*, 17 (2), 183–207.

Eckert, P. (1994). Entering the heterosexual marketplace: Identities of subordination as a developmental imperative. In *Working Papers on Learning and Identity* (Vol. 2). Palo Alto, CA: Institute for Research on Learning.

Eckert, P. (2008). Where do ethnolects stop? *International Journal of Bilingualism*, 12 (1-2), 25–42.

Eckert, P. (2011). Language and power in the preadolescent heterosexual market. *American Speech*, 86 (1), 85–97.

Eckert, P. (2019). The individual in the semiotic landscape. *Glossa: a journal of general linguistics*, 4 (1).

Labov, W. (2001). *Principles of linguistic change, vol. 2: External factors*. Oxford: Blackwell Publishers.

Nesbitt, M. (2021a). The Rise and Fall of the Northern Cities Shift: Social and Linguistic Reorganization of TRAP in Twentieth-Century Lansing, Michigan. *American Speech*, 96 (3), 332–370.

Nesbitt, M. (2021b). From an indicator to a marker. *Urban Matters: Current Approaches in Variationist Sociolinguistics*, 27, 159–177.

- Nesbitt, M., Wagner, S. E., & Mason, A. (2019). A Tale of Two Shifts: Movement Toward the Low-Back Merger Shift in Lansing, Michigan. *Publication of the American Dialect Society*, 104(1), 144–165.
- Sneller, B. (2012). “Aw Man!”: The Effect of Hometown Affiliation on NCS Shifting in Holland, Michigan. [MA thesis]. University of Essex.
- Sneller, B. (2020). Phonological rule spreading across hostile lines: (TH)-fronting in Philadelphia. *Language Variation and Change*, 32 (1), 25–47.
- Sneller, B., Wagner, S. E., & Ye, Y. (2022). MI Diaries: ethical and practical challenges. *Linguistics Vanguard*, 8(s3), 307–319.
- Villarreal, D. (2018). The construction of social meaning: A matched-guise investigation of the California Vowel Shift. *Journal of English Linguistics*, 46 (1), 52–78.
- Wagner, S. E., Mason, A., Nesbitt, M., Pevan, E., & Savage, M. (2016). Reversal and re-organization of the Northern Cities Shift in Michigan. *University of Pennsylvania Working Papers in Linguistics*, 22 (2), 19.
- Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for Historical Linguistics* (pp. 97–195). Austin: University of Texas Press.
- Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85 (3), 315–337.

---

## Corpus data meet survey data: A case study from Singapore English modality

Carmelo Alessandro Basile

Université Paris Cité, Paris, France. Université Sorbonne Nouvelle, Paris, France

The contact variety of English spoken in Singapore (SgE) has undergone significant grammatical restructuring, influenced by other languages spoken on the island, namely Mandarin Chinese, Malay, and Tamil. One domain that has witnessed a distinctive development compared to inner-circle varieties of English is modality. At least two approaches have been proposed to explain SgE modality: the substratist and the contact-grammaticalization approaches. The substratist approach (Bao 2010) suggests that features from the lexifier language (English) may converge with equivalent features in the substratum language, adopting their functions. For example, the restriction to non-epistemic uses of *must* in SgE is attributed to the substratum influence of the Mandarin deontic particle *bixu*, which has not yet grammaticalized to epistemic uses. The second approach explores modality through the lens of contact-grammaticalization theories (Heine and Kuteva 2005). Building on Ziegeler (2014), Basile (2023) argues that some over-generalized uses of modal verbs can be explained by a recapitulation of earlier diachronic stages in the lexifier. The frequent non-epistemic use of *must* in present-day SgE can be seen as a replica of past dynamic/generic meanings of *must*, observed since Early Middle English times (Gregersen 2020).

These two approaches show that the influence of lexifier and substrate languages is crucial in assessing different contact phenomena. Nevertheless, past studies on the modal system of necessity and obligation in SgE (e.g. Bao 2010; Hansen 2018; Loureiro-Porto 2019; Basile 2023) have all been conducted on corpora, such as ICE, lacking metadata capable of identifying the substrate languages of their contributors. Importantly, the common production of a marked feature by different ethnic groups with different substrates would indeed pose a threat to the substratist approach.

Building on Basile (2023), the present paper re-explores the use of (semi-)modal constructions of necessity and obligation (i.e. *must*, *have to*, *(have) got to*, *gotta*, *need to*) in SgE by incorporating, for the first time in the literature, a sociolinguistic survey intended to elicit their production among participants belonging to different ethnicities (Chinese, Malay, Indian) and age groups (18-29, 30-49, 50+). Participants were asked what they would say in a series of imaginary situations, such as (1) and to select the most natural modal constructions from given alternatives.

(1) Your daughter had a party yesterday and left the house in a mess! You need to receive guests in an hour. What would you say to your daughter to make her clean the house immediately?)

Survey results reveal that ethnicity does not affect epistemic/non-epistemic *must* production, thus challenging Bao (2010) and supporting Basile (2023). Age exhibits, on the contrary, significant variance [ $p < 0.001$ ]. Following an apparent-time approach (Labov 1963), the paper shows that the survey data confirm a declining trend for *must* and a sharp increase of *need to* in SgE (cf. also Hansen 2018, Loureiro-Porto 2019, and Basile 2023). The rise of *need to*, in particular, is presented as an effect of substrate reinforcement from a parallel construction in Mandarin.

Bao, Zhiming. 2010. “Must in Singapore English”. *Lingua* 120: 1727–1737.

Basile, Carmelo Alessandro. 2023. “Necessity modal development in Singapore English: An investigation of substratist and contact-grammaticalisation approaches”. *English World-Wide* 44: 276–302.

Gregersen, Sune. 2020. *Early English Modals: Form, Function and Analogy*. LOT, Netherlands Graduate School.

Hansen, Beke. 2018. *Corpus Linguistics and Sociolinguistics: A Study of Variation and Change in the Modal System of World Englishes*. Leiden: Brill.

Heine, Bernd, and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Labov, William. 1963. “The social motivation of a sound change”. *Word* 19: 273–309.

Loureiro-Porto, Lucía 2019. "Grammaticalization of Semi-Modals of Necessity in Asian Englishes". *English World-Wide* 40: 115–142.

Ziegeler, Debra. 2014. "Replica Grammaticalisation as Recapitulation: The Other Side of Contact". *Diachronica* 31: 106–141.

---

## A response to Chambers and Trudgill (1980): The stylistic implications of (ing) across the lifespan of the individual

Lea Bauernfeind, James Grama, Johanna Mechler  
Sociolinguistics Lab, University of Duisburg-Essen, Germany

Based on Trudgill's foundational Norwich study (1974), Chambers and Trudgill (1980) explored style shifting across the lifespan, finding that co-variation between (ing) and age is impacted by stylistic parameters: retrenchment towards the standard form in post-adolescence was most noticeable in formal extemporaneous speech, whereas casual conversation was the last hold-out for vernacular forms.

Their analysis, and many others since (e.g., Labov 2001 on negation, (dh) and (ing)) have relied on apparent-time data, recordings from speakers of different age cohorts at one point in time. Recently, the sociolinguistic enterprise has made great strides in the construction of panel datasets, which collect repeated recordings from the same person at several points across their lives (e.g., Sankoff and Blondeau 2007). While much has been learnt about how speakers vary and change in their linguistic patterns as they age, very little attention has been paid to how speakers' stylistic repertoires change across their life history (but see Rickford and Price 2013).

This paper provides a response to Chambers and Trudgill's findings by adding a panel element to style shifting in the use of (ing) across the adult lifespan. Drawing on heritage data from a long-term project in the North-East of England, we built a novel panel corpus, which follows 28 people in six successive age-based cohorts, recording them two or three times at roughly seven-year intervals. Stylistic variation is conceptualised along two different dimensions: topic (careful vs. casual; see Rickford and McNair-Knox 1994; Labov 2001), and addressee (Bell's 1984 audience design model).

We focus on (ing), a stable variable, whose alveolar realisation in the broader North-East community is almost at ceiling (Mechler et al. 2022). Socio-psychological research has shown that (ing), while being associated with "monitored" registers (Schleef et al. 2011:209), lacks social stigma in the North of England and is less salient than in the South (Levon and Fox 2014; Levon and Buchstaller 2015). Contrary Chambers and Trudgill's findings, we observe little evidence that (ing) is sensitive to style in the North-East. While the professional trajectories of some speakers provide insights into stylistically-niched patterns (Author et al. to appear), the majority of our panel speakers do not style-shift towards [ŋ] in formal topics. Even less variability arises across addressee. Moreover, there is no evidence of a U-shaped curve with increased retrenchment in careful styles as speakers move into the middle age brackets.

Our data provide a panel window into a stable variable that is realised almost categorically in the community. The lack of stylistic differentiation suggests that we need more data from stable variables to test the assumptions underlying our models of age-grading. More specifically, our data indicates that the extent to which age-related style shifting occurs is as much a function of the socio-indexical landscape in which a variable operates as the linguistic marketplace pressures that speakers find themselves under.

Bell, A. 1984. "Language Style as Audience Design". In *Sociolinguistics: A Reader and Coursebook*, N. Coupland and A. Jaworski (eds). New York: St Martin's Press Inc. 240–250.

Chambers, J., and P. Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.

Labov, W. 2001. *Principles of linguistic change. II: Social factors*. Oxford: Wiley-Blackwell.

Levon, E., and I. Buchstaller. 2015. "Perception, cognition, and linguistic structure: The effect of linguistic modularity and cognitive style on sociolinguistic processing." *Language Variation and Change* 27(3):319–348.

Levon, E., and S. Fox. 2014. "Social salience and the sociolinguistic monitor: A case study of ING and TH-fronting in Britain." *Journal of English Linguistics* 42(3):185–217.

Mechler, J., Grama, J., Bauernfeind, L., Eiswirth, M.E., and Buchstaller, I. 2022. "Towards an empirically-based model of age-graded behaviour: Trac(ing) linguistic malleability across the entire adult life-span." *University of Pennsylvania Working Papers in Linguistics* 28(2):101–110.

Rickford, J., and F. McNair-Knox. 1994. "Addressee- and Topic-Influenced Style Shift: A Quantitative Socio-Linguistic Study". In *Perspectives on Register: Situating Register Variation within Sociolinguistics*, D. Biber and E. Finegan (eds). Oxford: Oxford University Press. 235–276.

Rickford, J., and M. Price. 2013. "Girlyz II women: Age-grading, language change, and stylistic variation." *Journal of Sociolinguistics* 17(2):143–179.

Sankoff, G., and H. Blondeau. 2007. "Language Change across the Lifespan: /r/ in Montreal French." *Language* 83(3):560–588.

Schleef, E., Meyerhoff, M., and L. Clark. 2011. "Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London." *English World-Wide* 32:206–236.

Trudgill, P. 1974. *The social differentiation of English in Norwich*. Cambridge: CUP.



---

## Intensifier Variation and Use: The Effects of Formality on the Internet

Joshua Baumgarten<sup>1</sup>, Carl Börstell<sup>2</sup>, Corrine Occhino<sup>3</sup>

<sup>1</sup>Syracuse University, <sup>2</sup>University of Bergen, <sup>3</sup>University of Texas at Austin

Intensifiers are a class of adverbs that change relatively rapidly compared to other parts of English (Peters 1994). Frequently, their functionality results from the process of grammaticalization in which they take on an additional emphatic role. Over time, certain intensifiers also take on connotations of informality, though *very* stands out from the other frequent English intensifiers because of its favored usage in formal speech and written text (see Jespersen 1922, Tagliamonte 2016). Despite much attention on the development and changing use of intensifiers (see Ito & Tagliamonte 2003, Tagliamonte & Roberts 2005), little has been done on the relationship between formality in discourse context and intensifier usage.

We examine the distribution of the four most frequent English intensifiers, *very*, *really*, *pretty*, and *so* in a corpus of Reddit utterances to address usage in a context of varied formality. Using Cornell University's ConvoKit Reddit corpus (Chang et al. 2020), we examined 21,216,663 utterances. In addition to investigating simple lexical distribution, we extracted 2,107,446 bigrams and 146,581 trigrams. Reddit was chosen because subreddits exhibit varying levels of formality across a large number of utterances. We analyzed the distribution of intensifiers across different registers (formal, neutral, informal) to identify whether certain intensifiers were common in certain formality contexts. Ngrams were defined by intensifier-adjective structure where the first word was one of the four target intensifiers.

Our study reveals that the distribution of the target intensifiers varies by formality of discourse genre. In-line with previous discussions, *very* is more frequent in formal contexts with decreasing use as formality decreases. It also has a wider distribution of adjectival-pairings in the formal register. In the informal contexts, *so* was the most frequent intensifier and had a wider distribution of adjectival pairings, with increasing use as formality decreases (Figures 1, 2). Across the formality cline, bigram and trigram collocations reveal that certain adjectives more frequently occur with specific intensifiers. For instance, *pretty* was heavily favored to precede *sure*, and *so* most frequently preceded *much*.

The trigram data also reveal a preference for repeated constructions (i.e., intensifier is reduplicated) in informal trigrams. Across the registers, there is noticeable variation of the repeated intensifiers in these trigrams, with *very very* trigrams being the most frequent formal repeated trigram, *so so* being the most frequent informal repeated trigram, and *pretty pretty* being very infrequent (Figure 3). Overall, our preliminary findings point to correlations between intensifier choice, adjective choice, and formality within an Internet discourse genre. Additionally, while *very* and *really* appear with a variety of adjectival-pairings, the frequent co-occurrence of *pretty* and *so* with specific adjectives suggests the emergence of fixed multi-word expressions and potentially unique grammaticalization pathways. With their propensity to change, understanding how intensifiers appear on Reddit may help to understand how English speakers use and perceive them in different formality contexts on the Internet.

Chang, J. P., Chiam, C., Fu, L., Wang, A., Zhang, J., and Danescu-Niculescu-Mizil, C. 2020. "ConvoKit: A Toolkit for the Analysis of Conversations." In Proceedings of the 21th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 57–60. 1st virtual meeting: Association for Computational Linguistics. <https://aclanthology.org/2020.sigdialog-1.8>.

Ito, R., & Tagliamonte, S. (2003). Well weird, right dodgy, very strange, really cool: Layering and recycling in English intensifiers. *Language in society*, 32(2), 257-279.

Jespersen, O. (1922). *Language: Its nature, development and origin*. G. Allen & Unwin Ltd.

Peters, H. (1994). Degree Adverbs in Early Modern English. In *Studies in Early Modern English*.

Tagliamonte, S. (2016). So sick or so cool? The language of youth on the internet. *Language in Society*, 45(1), 1-32. doi:10.1017/S0047404515000780

Tagliamonte, S., & Roberts, C. (2005). So weird; so cool; so innovative: The use of intensifiers in the television series *Friends*. *American speech*, 80(3), 280-300.

---

## A who yuh tink seh meet Tina: A reanalysis of "say" complementizers in Jamaican Patois

Kimberley Baxter

New York University, New York City, USA

In Jamaican Patois, there are two forms of *seh*: a verbal form meaning *say* and a complementizer-like form meaning that (1). This study focuses on complementizer-like *seh*, for which there have been conflicting analyses regarding whether it is truly a complementizer or strictly a verb. This creates an interesting problem because if *seh* is truly a complementizer, why is it allowed to flout complementizer-trace effects in a language with no post-verbal subjects

(2)? (Rizzi, 1982; Pesetsky, 2017) Conversely, if *seh* is strictly a verb and not a complementizer, why do Jamaican Patois speakers interpret *seh* in example (2) as that?

(1)  
Mi        seh            to    im    seh            im        nuh            fi            forward        ya.  
I-1SG    say-VERB    to    him    That-COMP    he        NEG-is not    to-PREP    come-VERB    here.

I said to him that he is not to come here./I told him not to come here.

(2)  
A            who    yuh    tink    seh            meet            Tina?  
FOC-is    who    you    think    that-COMP    meet-PST    Tina?

Who do you think (\*that) met Tina?

There are two analyses of complementizer-like *seh* in Jamaican Patois: one, presented by Besten and Veenstra (1994) claims that *seh* is an overt complementizer, over which subject extraction over an embedded clause is not allowed. The other, presented by Durreleman (2008), claims that *seh* is a verb which loses its individual meaning when it forms serial verb constructions in verb-*seh* phrases. Eg: *tink seh* would mean think, not think say or think that. However, I present a new set of data which challenges these arguments.

I conduct a judgment survey of Jamaican Patois speakers across the Jamaican Diaspora to test the acceptability of subject extraction over *seh* and other complementizers in Jamaican Patois, including stacked complementizers *seh dat* and *seh if*. In this survey, 43 example sentences are rated on a scale of 1-5, with one being the least acceptable and 5 being the most acceptable. The survey follows Sobin (1987), which tested the acceptability of subject-extraction over a series of complementizers in American English and found variable acceptability of this movement over the complementizer *that*.

Results show that while subject extraction over *seh* is rated very highly, ratings are even higher for Patois speakers in Jamaica than for Patois speakers outside of Jamaica. This trend continues in other tested categories, which suggests language variation in the Jamaican Diaspora. Additionally, subject extraction over *dat (that)* is widely accepted, with ratings of 3 or above for 71% of participants. This is important because where debate exists about the status of *seh* as a complementizer, there is no debate about the status of *dat* (Durreleman, 2008). This seems to suggest that subject extraction over complementizers is more widely accepted in Jamaican Patois than originally thought. Importantly, it also underscores the vital importance of the viewpoints of language speakers to the formation and analysis of syntactic theory.

Besten, H. D., & Veenstra, T. (1994). Fronting. Pidgins and Creoles: An introduction. Amsterdam: John Benjamins, 303-315.  
Durreleman, S. (Ed.). (2008). The syntax of Jamaican Creole: A cartographic perspective (Vol. 127). John Benjamins Publishing.  
Pesetsky, D. (2017). Complementizer-trace effects. Blackwell companion to syntax, 2nd edition. Oxford: Wiley-Blackwell. DOI: <https://doi.org/10.1002/9781118358733.wbsyncom108>.  
Rizzi, L. (1982). Negation, wh-movement, and the null subject parameter. An Annotated Syntax Reader, 169.  
Sobin, N. (1987). The variable status of Comp-trace phenomena. Natural Language & Linguistic Theory, 5(1), 33-60.

---

## Innovation and standardization in cross-sociolinguistic corpus development: A Case Study with Subject Pronoun Expression

Karen Beaman<sup>1</sup>, Aria Adli<sup>2</sup>, Robert Bayley<sup>3</sup>, Daniel Erker<sup>4</sup>, Gregory Guy<sup>5</sup>, Rafael Orozco<sup>6</sup>

<sup>1</sup>University of Tübingen, Tübingen, Germany. <sup>2</sup>University of Cologne, Cologne, Germany. <sup>3</sup>UC Davis, Davis, USA. <sup>4</sup>Boston University, Boston, USA. <sup>5</sup>New York University, New York, USA. <sup>6</sup>Louisiana State University, Baton Rouge, USA

The field of sociolinguistics is ripe for methodological standardization. Since the birth of variationist sociolinguistics, bolstered by an overall movement in research toward more rigorous quantitative analyses, the quantity of sociolinguistic data has proliferated. This trend is partly the result of expanding data availability but is also the outcome of more powerful computers, sophisticated tools, and advanced statistical methods to aid data analysis, compounded by sociolinguists' ongoing curiosity to answer new and different questions. Most sociolinguistic corpora have been collected by individual researchers, examining a few linguistic variables, investigating a single language variety, and using unique formats that are difficult to integrate and/or replicate. The one-researcher, one-community, one-variable approach introduces obstacles to synthesis, especially cross-linguistically. Moreover, the particularities of sociolinguistic data require specific treatment of variables and contexts beyond what is typically required in corpus linguistics. Ultimately, meta-analyses are not generally robust enough to confirm the presence or absence of probabilistic constraint effects or the relevance of differing typological structures. Thus, variationist sociolinguistics needs to evolve toward standardized, replicable, open-source, cross-linguistic data structures.

Arguably the most important challenge in cross-sociolinguistic analysis is to ensure apples-to-apples comparisons across corpora by clearly circumscribing the variable context and specifically defining the envelope of variation (Labov 1972). This effort has been facilitated for subject pronoun expression (SPE) – the variable presence or absence of subject pronouns – *yo canto ~ canto* / ‘I sing’ – a unique sociolinguistic variable that is found, to varying degrees, in many languages. SPE has been investigated in numerous studies across a number of languages (Beaman 2022; Guy, Orozco, and Erker 2021; Li and Bayley 2018; Torres Cacoullós and Travis 2019), notably in the seminal work of Otheguy and Zentella (2012). Using their work as a springboard, we have expanded the methodology to cover five different languages and more than 20 linguistic communities to build the first integrated, cross-linguistic corpus of subject pronoun expression, comprising over 850 speakers and 250,000 tokens collected through sociolinguistic interviews. The data are coded for four internal/structural constraints (person-number, clause type, TMA, verb semantics), three cognitive/discursive constraints (priming, referent continuity, lexical frequency), and four social constraints (community, age, gender, education).

The aim of this paper is to demonstrate this innovative approach to cross-language research, to share the challenges and successes we have experienced, and to showcase the lessons we have learned in cross-sociolinguistic corpus development, data analysis, and data visualization. We will demonstrate our bespoke Quant-Companion to illustrate the project’s analytics-oriented data architecture, which maximizes transparency and ensures replicability in terms of statistical analyses and visualizations (e.g., see Figure 1). The methods we have developed for our SPE corpus can be readily applied to other sociolinguistic variables that occur in multiple languages, such as consonant cluster reduction and various discourse-pragmatic variables. This paper demonstrates the importance of establishing global norms for data collection, specification, and visualization in order to facilitate a deeper and broader understanding of the inherent, universal processes of linguistic variation and change.

Beaman, Karen. 2022. “Social Meaning and Community Belonging in Subject Pronoun Expression.” In *New Ways of Analyzing Variation* 50, Stanford University.

Guy, Gregory R., Rafael Orozco, and Danny Erker. 2021. “Homogeneity and Heterogeneity in Null Subjects: A Cross-Linguistic and Cross-Community Comparison.” In *New Ways of Analyzing Variation* 49, Austin, Texas.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.

Li, Xiaoshi, and Robert Bayley. 2018. “Lexical Frequency and Syntactic Variation: Subject Pronoun Use in Mandarin Chinese.” *Asia-Pacific Language Variation* 4(2): 135–60.

Otheguy, Ricardo, and Ana Celia Zentella. 2012. *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. Oxford: Oxford University Press.

Torres Cacoullós, Rena, and Catherine E. Travis. 2019. “Variationist Typology: Shared Probabilistic Constraints across (Non-)Null Subject Languages.” *Linguistics* 57(3): 653–92.

## Lexical Variation in Terms for Romantic Partners in Québec French: A Diachronic Study

Guillaume Beauchamp, Nadège Fournier  
 Université de Montréal, Montréal, Canada

In Québec French (QF), many terms can designate a love partner. Some are restricted in meaning, e.g., *femme* ‘wife’/mari ‘husband’ and *époux/épouse* ‘spouse’ are used for married couples, while others are polysemic and specific to QF such as *blonde/chum* ‘love partner’ (Mercier, 2002). This study investigates this case of lexical variation. We examine how these terms have evolved over time in Québec and how they differ based on societal context and social factors such as age, gender, social class and marital status.

The numerous words for romantic partners in QF have already been the subject of some studies. Vincent and Lambert (2010) analyzed dictionaries and journalistic texts and conducted a perception survey of 350 Quebecers. The authors conclude that the Quiet Revolution (1960–1966) led to profound societal changes that affected the use and representations of these variants. For example, they report that people under 54 years of age prefer not to specify whether they are married or not, thus moving away from *femme* ‘wife’/mari ‘husband’ in favor of *blonde/chum* ‘love partner.’ Beaudoin and Fournier’s (2020) survey of 140 Quebecers reveals the same trends.

To complement these studies, we adopt a distinct approach to explore the same variable. Precisely, we analyzed spontaneous speech in sociolinguistic interviews using the variationist method (Labov, 1972). Drawing on corpus data collected in Montréal at two different time periods, 1971 (Sankoff et al., 1976) and 2012 (Blondeau et al., 2012), we examined the distribution of 1076 tokens of 21 variants. We propose an integrated approach that incorporates both statistical analysis and qualitative description, drawing upon our comprehensive knowledge of the 158 interviews.

Our findings indicate a significant evolution between 1971 and 2012 (see tables 1 and 2). In line with previous studies, we observed a noticeable decline in the use of conventional terms such as *femme* ‘wife’/mari ‘husband’ in favor of more informal terms such as *blonde/chum* ‘love partner.’ Surprisingly, the latter variants are often used to refer to married couples in 2012. This shift is more pronounced among younger speakers, suggesting a generational

renewal in couple designations. Furthermore, exploration of the most recent corpus (2012) shows differences in usage based on social factors. Married individuals, compared to unmarried ones, tend to use the first name of their partner more frequently to refer to them, which might be a strategy to avoid the obsolete forms *femme* ‘wife’/ *mari* ‘husband.’ Moreover, *copine* ‘girlfriend’/ *copain* ‘boyfriend,’ which were almost absent from the 1971 corpus, are commonly used by young speakers in 2012 and also serve to refer to same-sex partners. The legalization of gay marriage in 2004 and the growing acceptance of same-sex relationships might have played a role in the emergence of these forms.

In conclusion, our results underscore the intricate lexical variation within a societal context where marriage no longer serves as a marker of social achievement and where non-heteronormative realities are expressed openly. Our study provides insight into how the evolution of norms and representations influences variation.

- Beaudoin, F., & Fournier, N. (2020). L'évolution des dénominations des membres du couple au Québec entre 1957 et 2020, UdeM, VocUM 2020 - Langage : norme et pouvoir, Montréal, Canada, 19-20 novembre 2020.
- Blondeau, Hélène, Yves Frenette, France Martineau (Dir.), and Mireille Tremblay. (2012). Sous-corpus variationniste de Hochelaga-Maisonneuve du Corpus FRAN.
- Labov, W. (1972). Sociolinguistic patterns (vol. 4). University of Pennsylvania Press.
- Mercier, L. (2002). Le français, une langue qui varie selon les contextes. Dans *Le français, une langue à apprivoiser. Textes des conférences prononcées au Musée de la civilisation (Québec, 2000-2001) dans le cadre de l'exposition « Une grande langue: le français dans tous ses états »* (pp. 41-60).
- Sankoff, David, Gillian Sankoff, Suzanne Laberge, and Marjorie Topham. (1976). Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale. *Cahiers de Linguistique de l'Université du Québec* 6: 85–125.
- Vincent, D., & Lambert, S. (2010). Variation linguistique et changement social: les mots désignant les membres du couple en français québécois. Remysen, W/Vincent, D.(éds.), *Hétérogénéité et homogénéité dans les pratiques langagières, Mélanges offerts à Denise Deshaies*. Sainte-Foy, Les Presses de l'Université Laval.

---

## Automatic Transcription of Language Documentation Recordings for Sociolinguistic Analysis: Speech Recognition and Forced-Alignment for Northern Prinmi

Connor Bechler

University of Kentucky, Lexington, USA

Sociolinguistics and language documentation are highly complementary: applying sociolinguistic methods during documentation produces a more complete description of the language being documented while also broadening the scope of sociolinguistic inquiry to include a more diverse array of language varieties, speech genres, and social settings (Meyerhoff, 2019). However, (semi-)automated sociophonetic analysis requires a significant quantity of speech data. While most language documentation projects produce an immense number of recordings, they face the transcription bottleneck: transcription and annotation of documentation recordings is estimated to take 50–100 hours per hour of audio (Čavar et al., 2016).

Researchers have sought to overcome this challenge with partial automation via forced alignment, natural language processing, and automatic speech recognition (ASR) (Coto-Solano et al., 2021; Neubig et al., 2020). Transformer-based cross-lingual ASR models like XLSR-53 promise improved performance on under-resourced languages by leveraging massive data sets from many different languages (Conneau et al., 2020). However, as Coto-Solano (2022) notes, the process of adapting these tools for sociolinguistic analysis of minority and indigenous languages is still challenging.

This project extends these efforts to a novel context, adapting the XLSR-53 ASR model to produce automatic transcriptions of Northern Prinmi, a Tibeto-Burman Qiangic language spoken in Southwest China, with the aim of also implementing a forced-alignment system for sociophonetic analysis.

The training corpus is drawn from an oral art documentation project (Daudey & Gerong, 2018) and consists of three hours of transcribed recordings of 13 speakers from seven townships performing a range of songs and chanted rituals. Northern Prinmi is not standardized and does not have an established orthography, so transcriptions are in IPA and the recordings may include differing dialectal varieties. The small size, internal variation, and genre variation of the recordings make the corpus a rich site for sociolinguistic inquiry, though these same qualities also make ASR more difficult.

To assess the efficacy of XLSR-53 ASR models on language documentation recordings, I fine-tuned multiple XLSR-53 models on the corpus, testing a range of preprocessing methods and training hyperparameters. I tested two primary preprocessing methods: combining multi-character representations of lexical tones into single characters and combining characters representing consonants or vowels with their modifying diacritics. The best of the resulting models achieved a character error rate of 0.32. This error rate is comparable to previous ASR work for sociophonetics (Coto-Solano et al., 2021), suggesting that sociophonetic analysis of Northern Prinmi using XLSR-53 generated transcriptions is feasible.

I hope to explore several further questions, related both to improving ASR and implementing tools for sociophonetic analysis. Which recording variables—location, speaker, genre, gender—most impact model performance? What sort of errors does the model make, and how do these errors affect analysis? What approach to forced-alignment will work best for Northern Prinmi? How can the utility of the speech-processing systems for sociophonetic analysis be evaluated?

Through this work with Northern Prinmi, I aim to evaluate how recent computational tools can be made maximally useful for sociolinguists and documentation researchers working in similar and related contexts.

- Ćavar, M., Ćavar, D., & Cruz, H. (2016). Endangered Language Documentation: Bootstrapping a Chatino Speech Corpus, Forced Aligner, ASR. Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), 4004–4011. <https://aclanthology.org/L16-1632>
- Conneau, A., Baevski, A., Collobert, R., Mohamed, A., & Auli, M. (2020). Unsupervised Cross-lingual Representation Learning for Speech Recognition (arXiv:2006.13979). arXiv. <https://doi.org/10.48550/arXiv.2006.13979>
- Coto-Solano, R. (2022). Computational sociophonetics using automatic speech recognition. *Language and Linguistics Compass*, 16(9), e12474. <https://doi.org/10.1111/lnc3.12474>
- Coto-Solano, R., Stanford, J. N., & Reddy, S. K. (2021). Advances in Completely Automated Vowel Analysis for Sociophonetics: Using End-to-End Speech Recognition Systems With DARLA. *Frontiers in Artificial Intelligence*, 4. <https://www.frontiersin.org/articles/10.3389/frai.2021.662097>
- Daudey, H., & Gerong, P. (2018). Documentation of Northern Prinmi oral art, with a special focus on ritual speech. *Endangered Languages Archive*. Handle: <http://hdl.handle.net/2196/00-0000-0000-0010-8820-B>
- Meyerhoff, M. (2019). Unnatural bedfellows? The sociolinguistic analysis of variation and language documentation. *Journal of the Royal Society of New Zealand*, 49(2), 229–241. <https://doi.org/10.1080/03036758.2019.1619599>
- Neubig, G., Rijhwani, S., Palmer, A., MacKenzie, J., Cruz, H., Li, X., Lee, M., Chaudhary, A., Gessler, L., Abney, S., Hayati, S. A., Anastasopoulos, A., Zamaraeva, O., Prud'hommeaux, E., Child, J., Child, S., Knowles, R., Moeller, S., Micher, J., ... Littell, P. (2020). A Summary of the First Workshop on Language Technology for Language Documentation and Revitalization (arXiv:2004.13203). arXiv. <https://doi.org/10.48550/arXiv.2004.13203>

---

## Operationalizing gender: Methods for statistical modeling of open-response demographic data

Cooper Bedin, Lal Zimman, Marina Zhukova

University of California, Santa Barbara, Santa Barbara, USA

In the context of rapidly changing ideas about sex and gender, there is an increasing need for theories and methods that better reflect the complexity of these aspects of the self (e.g., Becker et al. 2023). This talk explores the implications of researchers' choices surrounding the way information about sex/gender is collected, binned, and modeled, and how these choices ultimately impact the results and interpretation of quantitative sociolinguistic analyses.

We present an analysis of sociophonetic data collected for a study on listener perception and queer speech. 14 speakers were recorded reading 25 sentences from the MOCHA-TIMIT corpus and reported their gender via open-response text after completing the task. They self-identified as "male" (n=8), "cis male" (n=3), "agender" (n=1), "gender non-conforming, amab, male-coded" (n=1), and "non-binary/questioning, but generally male-presenting" (n=1).

We compared three statistical models to examine the relationship between self-reported gender and center of gravity of the phoneme /s/, a sociophonetic variable associated with gender identity and expression (Stuart-Smith 2004; Zimman 2017). All models employed monofactorial linear regression with gender as the predictor and center of gravity as the response variable. The models differed in their gender categorization schemas (Figure 1): Model A represents a normative approach where speakers are categorized as "male," "female" (not present), or "other." Model B is an intermediate step where the largest sub-categories of the Model A schema are separated as "male" and "cis male." Model C treats each reported gender as distinct.

Based on the R2 and AIC scores we used to compare models, Model A (R2 = 0.07, R2Adjusted = 0.06, ΔAICc = 105.73) performed significantly worse than Model B (R2 = 0.39, R2Adjusted = 0.38, ΔAICc = 3.24) or Model C (R2 = 0.41, R2Adjusted = 0.40, ΔAICc = 0), as confirmed by the model effects plots in Figure 2. The effects plot for Model A indicated a significant difference in /s/ center of gravity between "male" and "non-male" categories. However, plots for Models B and C revealed that the difference was primarily driven by the three "cis male" speakers. Putting "cis male" into a separate category revealed a significantly higher average center of gravity compared to the other speakers. Interestingly, the center of gravity for "non-male" speakers did not significantly differ from that of "male" speakers, except for the "agender" speaker.

Our findings highlight the limitations of traditional approaches to operationalizing gender and their inability to capture the variation observed in our data. It may seem paradoxical that cis male-identified speakers are producing a higher center of gravity than other male speakers, but this fact supports the interpretation that the observed variation between male and cis male speakers is more strongly associated with their method of reporting gender than their

actual gender itself. We underscore the importance of considering the complexities of gender identification and self-reporting in quantitative research and offer our method as a way to move beyond binary gender categorizations and toward including the individualized ways research participants report their gender.

Stuart-Smith, Jane (2007). Empirical evidence for gendered speech production: /s/ in Glaswegian. In Jennifer Cole & José Ignacio Hualde (eds.), *Laboratory Phonology 9*, 65–86. New York: Mouton de Gruyter.

Zimman, Lal (2017). Variability in /s/ among transgender speakers: Evidence for a socially grounded account of gender and sibilants. *Linguistics* 55(5):993–1019.

---

## Intensifiers never go out of style: quantifying style and its effects on lexical variation

Barend Beekhuizen, Kaleigh Woolford  
University of Toronto, Toronto, Canada

Social identities are known to predict lexical variation. Across studies of intensifiers (e.g., Bolinger 1972; Partington 1993; Ito & Tagliamonte 2003) correlations such as ‘men use pretty more than women’, and ‘younger speakers use really more than older speakers’ are frequently observed. As work in the ‘third wave’ of variationist sociolinguistics (e.g., Eckert 2008, 2012) stresses, however, correlations between linguistic variables and macrosocial categories are not static associations marking social categories, but stylistic practices through which individuals index social identities and construct social meaning. To integrate such insights into a traditional variationist methodological framework, we propose a quantitative, corpus-based method for (1) measuring identity-associated styles and (2) assessing their effects on lexical choice. We do so for intensifier choice using an 8-million word corpus of social media data from the ‘Am I The Asshole’ subreddit, where ‘redditors’ discuss moral dilemmas, often self-identifying with age and gender in their posts.

We propose a measure to assess the explanatory power of stylistic age and stylistic gender, alongside the bare age and gender of the redditor (see Figure 1). In Step I we measure whether each word in our corpus is used significantly more or less by male than female redditors, or young [18-24yo] than old [30-36yo] redditors, and if so, by how much, resulting in a log Odds Ratio score representing the word’s age or gender bias. In Step II, we sum the age/gender biases for the 10 nearest context words of each intensifier token, resulting in an aggregated association measure reflecting how strongly the intensifier token’s context is associated with certain ages or genders. Step I measures mere associations to groups of people (approximating first order indexicalities, cf. Silverstein 2003) while Step II captures higher-order indexical purposes of such first-order indexicalities.

Using these stylistic factors as potential predictors alongside bare age and gender, in Step III, we model the choice between the 5 most frequent adjective-modifying intensifiers in our corpus (really, very, so, pretty, super), replicating previous findings. For all intensifiers, however, our two stylistic factors predict intensifier choice independently of the ‘bare’ demographic factors, and (per model comparison techniques) explain more of the variance, meaning that identity-associated stylistic factors are stronger predictors than the corresponding ‘bare’ demographic factors. Table 1 illustrates this with descriptive statistics and examples: while older Redditors use more very than really, using an ‘old style’ informs the choice between these intensifiers more strongly. Our results suggest that intensifier choice is made in lockstep with the wider decision to speak in a particular style, thus contributing a novel perspective on intensifiers, as well as a new method for understanding the role of style in lexical variation.

Bolinger, D. (1972). *Degree Words*. The Hague: Mouton.

Eckert (2008). Variation and the indexical field. *Journal of Sociolinguistics* 12(4): 453-476

Eckert (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41: 87-100.

Ito, R. & S. A. Tagliamonte. (2003). Well weird, right dodgy, really cool: Layering and recycling in English intensifiers. *Language in Society* 32, 257-279

Partington, A. (1993). Corpus evidence of language change: The case of intensifiers. In Baker et al. (eds.), *Text and Technology: In Honour of John Sinclair*. Amsterdam: Benjamins, 177-92.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23(3-4): 193-229.

---

## Underspecification and gender/sexuality-based meanings of noun-phrase number agreement

Ronald Beline Mendes, Isabel Pie  
University of Sao Paulo, Sao Paulo, Brazil

One of the most enregistered stereotypes in Brazilian Portuguese is that non-standard noun-phrase number agreement (NP $\emptyset$ , as in *as casa- $\emptyset$* , instead of *as casa-S* ‘the houses’) indexes lower class and level of education (Mendes & Oushiro 2015). By analyzing perception data collected experimentally, Mendes (2016) shows that NP $\emptyset$  can also index stereotypical masculinity: four male voices were perceived as more effeminate- and gay-sounding in their standard guises (NPs, as in *as casa-S*). According to him, the link between NP $\emptyset$  and masculine-soundingness is

indirect: NP $\emptyset$  is first and foremost associated with lower class and education, which is then potentially interpreted as masculine-soundingness. This paper discusses perception data that suggest a different interpretation: depending on attitudes toward homosexuality, NP $\emptyset$  can also be associated with effeminacy and gay-soundingness in male voices.

Using the same four male voices as in Mendes (2016), we designed a modified matched-guise experiment with stimuli that combine variable NP agreement (NP $\emptyset$  or NPs) with variable F0 (original or +30Hz) and variable (-s) duration (original or +80ms average). Each of the 204 listeners that participated in the experiment listened to each male voice once (in one of the possible combinations of the linguistic variants) and then evaluated them in various scales (including effeminacy and gay-soundingness). Additionally, the participants were asked to answer an attitude questionnaire (based on Gato, Fontaine & Carneiro 2012) intended to assess how conservative they tend to be towards male homosexuality. A Principal Components Analysis reveals that the first component in their responses comprises the scales for effeminacy and gay-soundingness (in an expected positive correlation). Taking the scores generated for this component as a dependent variable, we built regression models to test the effects of the linguistic variables and listener's social characteristics (gender, age, education and attitudes toward homosexuality, among others) in their responses. Results show that the male voices were perceived as more effeminate/gay-sounding in their Fo+30Hz and longer (-s) guises, regardless of participants' social characteristics. As for NP agreement, listeners with positive attitudes toward male homosexuality perceived the male voices as more effeminate/gay-sounding in their NP $\emptyset$  guise (differently from our expectation, based on Mendes 2016). Plus, a representation of the indexical field of NP agreement using Minimum Spanning Trees (Oushiro 2019) shows that, in the collected responses, the meanings "effeminate" and "gay" are closer to (more frequently associated with) "lower class" and NP $\emptyset$  than to NPs (see attached figure).

As our data yields interpretations that differ from those proposed by earlier studies for the link between NP agreement and gender/sexuality-based meanings, they also further illustrate the property of underspecification of the relationship between a linguistic form and social meanings (Eckert 2016). In addition, our study provides further evidence for the salience of variable F0 and (-s) lengthening in the perception of male voices, adding Brazilian Portuguese to the list of languages in which those variables have been analyzed in reference to gender and sexuality.

Gato, Jorge; Fontaine, Anne Marie; Carneiro, Nuno Santos (2012). Multidimensional scale of attitudes toward lesbians and gay men: Construction and preliminary validation. *Paidéia* 22/51: 11–20.

Eckert, P., 2016. Variation, meaning and social change. In: N. Coupland, ed. *Sociolinguistics. Theoretical Debates*. Cambridge: Cambridge University Press.

Mendes (2016a) Nonstandard plural noun phrase agreement as an index of masculinity. In: Levon, Erez; Mendes, R. B. *Language, sexuality and power. Studies in intersectional sociolinguistics*. Oxford University Press.

Mendes, R. B.; Oushiro, L. (2015) Variable number agreement in Brazilian Portuguese: An overview. *Language and Linguistics Compass* 9/9: 358-368.

Oushiro, Livia (2019). A computational approach for modeling the indexical field. *Revista de Estudos Linguísticos* 27(4): 1737-1786.

---

## Literal or detail-oriented? Exploring the source of Nerdy persona effects on pragmatic reasoning

Joyce He, Andrea Beltrama  
University of Pennsylvania, USA

Persona-based information has been shown to guide language comprehension, not only in speech processing (see e.g., [5]) but also in semantic interpretation. For example, comprehenders interpret numerals more precisely when uttered by Nerdy speakers vs. Chill speakers: they are less likely to select a price of \$207 as the referent of "The price is \$200" with the former persona than the latter ([1]). The exact sources of this effect, yet, are still under-explored; two alternative hypotheses are that Nerdy speakers, compared to Chill speakers, are perceived as especially:

- Hyp.A: attentive to literal meaning, hence committed to carefully observing extant form-to-meaning mappings and avoiding loose talk
- Hyp.B: detail-oriented towards the world, i.e. attuned to pick up on minute details of objects/events, and represent them in their descriptions.

Both possibilities are compatible with prior sociolinguistic analyses of Nerdy ([3-4]) and lead to associate Nerdy speakers with precise interpretations of numerals. But they crucially generate divergent predictions with other linguistic phenomena for which attention to literal meaning vs. detail-orientedness don't converge. One case in point are vague adjectives, which we investigate here.

Consider (1) used to describe two minimally different routes: Route A is 595 miles, Route B is 593.

(1) Route A is long, but Route B is not.

Being a vague predicate, long is literally true of an object if it exceeds a contextually relevant threshold by a significant amount ([8-9]). But in (1), whatever the threshold is, it can't be significantly exceeded if Route A counts as long whereas Route B, only minimally shorter, doesn't. Following Hyp.A, a Nerdy speaker shouldn't be expected to utter (1), since it entails using long loosely. But following Hyp.B, one could posit that a Nerdy speaker might fixate on a minor mileage difference between the routes, and thus utter (1) – even at the price of stretching literal truth.

We'll test Hypothesis A-B adapting [1]'s paradigm: participants will see dialogues where one speaker asks for information and another checks their phone and responds with construction like (1) containing a vague adjective. Participants will then choose between two pictures of a phone – VISIBLE vs. COVERED. They'll choose the VISIBLE one if they believe it might be the one the speaker was actually looking at; the COVERED one otherwise.

We'll manipulate:

- Speaker persona: Nerdy vs. Chill (between-subjects, Fig.1-2);
- Diff(ERENCE) between the numbers on the VISIBLE screen, with 3 (within-subjects) levels (Fig. 3):
  - Large-Diff: sentence clearly true;
  - No-Diff: sentence necessarily false
  - Small-Diff: sentence true only on loose reading

We expect ceiling and floor rates of VISIBLE choices in Large-Diff and No-Diff. In critical Small-Diff, we predict modulations by speaker Persona. VISIBLE choice rates will be:

- Hyp.A: Nerdy < Chill, since the literal meaning of long doesn't hold.
- Hyp.B: Nerdy > Chill, since uttering (1) requires extreme detail-orientedness.

This study provides a more fine-grained picture of the interaction between social identity and pragmatic reasoning, enriching extant work on how socio-indexical information affects meaning interpretation ([1-2-7-10]) and linguistic cognition more broadly (e.g., [11-6-12]).

- [1] A. Beltrama and F. Schwarz. Imprecision, personae, and pragmatic reasoning. *Semantics and Linguistic Theory (SALT)*, 2021.
- [2] A. Beltrama and F. Schwarz. Social identity, precision and charity: when less precise speakers are held to stricter standards. *Semantics and Linguistic Theory* 32, 2022.
- [3] M. Bucholtz. "Why be normal?": Language and identity practices in a community of nerd girls. *Language in Society*, 28(2):203–223, 1999.
- [4] M. Bucholtz. The whiteness of nerds: Superstandard English and racial markedness. *Journal of Linguistic Anthropology*, 11:84–100, 2001.
- [5] A. D'Onofrio. Personae in sociolinguistic variation. *WIREs Cognitive Science*, 11(6), 2020.
- [6] K. K. Drager. *Linguistic variation, identity construction and cognition*. Language Science Press, 2015.
- [7] S. Fairchild and A. Papafragou. Sins of omission are more likely to be forgiven in non-native speakers. *Cognition*, 181:80–92, 2018.
- [8] D. G. Fara. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, 28:45–81, 2000. Originally published under the name "Delia Graff".
- [9] C. Kennedy. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30(1)(1):1–45, 2007.
- [10] C. Kennedy and L. McNally. Scale structure, degree modification and the semantics of gradable predicates. *Language*, 81(2):345–381, 2005.
- [11] T. Mahler. *Social identity information in projection inferences: a case study in social and semantic-pragmatic meaning*. PhD thesis, The Ohio State University, 2022.
- [12] N. Niedzielski. The effect of social information on the perception of sociolinguistic variables. *Journal of Social Psychology (Special Edition)*, 18:1:62–85, 1999.
- [13] L. Wade. *The Linguistic and the Social Intertwined: Linguistic Convergence toward Southern Speech*. PhD thesis, University of Pennsylvania, 2020.

---

## Operationalizing gender: Implications of participant-driven measurements

Montreal Benesch<sup>1</sup>, Cooper Bedin<sup>2</sup>

<sup>1</sup>N/A, USA. <sup>2</sup>University of California, Santa Barbara, Santa Barbara, USA

This talk fits into a larger discussion about how gender is measured in linguistics (e.g., Becker et al., 2023). The third wave viewpoint of social meaning considers speakers to be social agents, constructing identities through linguistic variation (Bucholtz & Hall, 2005). In line with this perspective, when working with participants, we should empower them throughout the process. Involving participants in the design of measurement tools emphasizes their agency, while simultaneously ensuring that the design accurately measures participants' own experiences. This allows for a more nuanced understanding of gender, where current methods are limited in their ability to capture many non-normative gender experiences.



The data we present were collected during a study of how genderfluid individuals (N=6) use language to construct their genders. Over the course of a month, participants made journal-style recordings of themselves and, following each recording, completed a survey indicating their gender at the time of recording. Measuring gender for these participants required balancing (1) accurately and affirmingly representing a fluid range of genders with (2) supporting quantitative analysis. Recognizing that genderfluid people often have unique, personal genders that may not align with normative notions of masculinity and femininity, the first author took an emic approach to creating measures. They met with each participant before data collection and collaboratively devised two independent Likert scales per participant to measure gender on post-recording surveys.

While these scales succeed at point (1), the question becomes how they affect a quantitative analysis of the data. To this end, we present an analysis of data from one speaker, Tarragon, and examine one sociolinguistic variable, /s/ center of gravity (CoG)—which is known to have a complex indexical field of gendered meaning in American English (e.g., Calder, 2019; Podesva & van Hofwegen, 2016). Tarragon provided five recordings, containing 241 /s/ tokens of duration >50ms. For each recording, Tarragon rated their gender from 0–6 on two scales collaboratively devised with the first author: defiant and blending. We then examine two quantitative methods for evaluating whether and how Tarragon’s self-rated genders correspond to variation in /s/ CoG.

The first method is the Kendall rank correlation coefficient,  $\tau$  (tau) (Kendall, 1938), given in Table 1. The second method is regression modeling, with gender as an ordinal predictor and /s/ CoG as a numeric response. One model was trained for each gender scale (Table 2). Model effects are plotted in Figure 1. Both methods support the hypothesis that Tarragon’s defiant~blending schema is indicative of their production of the gendered variable /s/, where defiant is inversely associated with /s/ CoG, and blending is positively associated with /s/ CoG.

In conclusion, this talk demonstrates that the results produced by devising such measures can be used to reach quantitatively-supported conclusions, while creating an affirming research environment for gender-diverse participants. This inherently enhances data quality, because researchers can better align their measurements with individuals’ senses of self. The data collection and analytic methods demonstrated in this paper can extend to a myriad of research questions, and warrant further development.

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4–5), 585–614. <https://doi.org/10.1177/1461445605054407> [<https://doi.org/10.1177/1461445605054407>]

Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika*, 30(1-2), 81–93. <https://doi.org/10.1093/biomet/30.1-2.81>

---

## /t/-release and burst duration in the construction of Jewish masculinity

Jaime Benheim

Brown University, Providence, USA

/t/-release carries a wide indexical field of potential social meanings [1]. Among Jewish speakers, /t/-release is associated with Orthodox men in particular [2], due to its ability to index social meanings related to education and authority, highly valued traits in Orthodox communities, and to Orthodox men’s greater exposure to speech styles involving this variable at yeshivas (Jewish seminaries). Despite this link to Orthodox masculinity, Jewish speakers of other genders and denominational backgrounds also recruit /t/-release indexically ([3,4]). Beyond Jewish speakers, phonetic detail such as burst duration can impact the social meaning associations of /t/-release [5]). In this study, I bring together these two lines of inquiry to examine how a group of Jewish men stylistically use /t/-release frequency and burst duration in tandem. I find that different combinations of these acoustic dimensions can index different locally-meaningful Jewish personae. Data come from sociolinguistic interviews collected in 2019 with 7 Jewish men (born 1984-1997; 2 Orthodox, 3 Conservative, 2 unaffiliated) from the Chicago metro area. This demographically restricted sample was selected to control for other macrosocial factors (gender, age, place). /t/-release was coded binarily (released v. unreleased, following [2]) and submitted to logistic regression models. Tukey contrasts were used to assess between-speaker pairwise comparisons. Following [5], differences in burst duration (normalized for speech rate) were assessed qualitatively. Finally, interview transcripts were coded for speakers’ stances towards and ideologies about Judaism and their Jewish identities.

Speakers produced between 4%-24% of word-final /t/ as releases. Whereas the two Orthodox speakers in the sample produced the most /t/-release in absolute terms, only one differed significantly in his rate of /t/-release from the other speakers in the sample. This occurred despite both Orthodox speakers’ strong connections to Orthodox Judaism and experience in yeshivas. Based on interview content, I argue that this difference may be due to these speakers’ different motivations to index a specifically yeshivish ([2]) speech style. Burst duration, meanwhile, varies according to speakers’ ideologies about the values associated with Jewish identity. A contrast emerged between speakers’ stances towards Jewish identity, with some associating it with education and learnedness (e.g., [2]), and others with an emphasis on social justice. Speakers who primarily associate Jewishness with social justice produced

longer burst durations in comparison to those who associated it with education (perhaps drawing on its associations with expressiveness; [5]), and this distinction crossed denominational backgrounds.

Through an analysis of interview content, I argue that speakers' differential uses of burst duration and /t/-release frequency enable them to index different Jewish personae. A non-Orthodox speaker, for example, is able to index that he values Jewish education (through short burst durations) while simultaneously affiliating in opposition to Orthodoxy (by producing the lowest /t/-release rate in the sample). Meanwhile, speakers who see their Jewish identities as inherently linked to social justice are able to index this through long burst durations and intermediate rates of /t/-release. Finally, I situate these results in relation to /t/-release's broader indexical field.

- [1] Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4): 453-476.  
[2] Benor, S. (2001). The learned /t/: Phonological variation in Orthodox Jewish English. *University of Pennsylvania Working Papers in Linguistics*, 37(2): 1-16.  
[3] Levon, E. (2006). Mosaic identity and style: Phonological variation among Reform American Jews. *Journal of Sociolinguistics*, 10(2): 181-204.  
[4] Leppala, J. (2020). Word-final /t/-release and linguistic style: An investigation of the speech of two Jewish women from Detroit. M.A. Thesis, Eastern Michigan University.  
[5] Podesva, R. (2006). Phonetic detail in sociolinguistic variation: Its linguistic significance and role in the construction of social meaning. Doctoral Dissertation, Stanford University.

---

## Rhythm as a contested marker of ethnicity in Modern Hebrew

Si Berrebi<sup>1</sup>, Sharon Peperkamp<sup>2</sup>

<sup>1</sup>Tel Aviv University, Tel Aviv, Israel. <sup>2</sup>Ecole Normale Supérieure - Université PSL, Paris, France

Social segregation based on socioeconomic or ethnic grounds typically boosts linguistic variation. In this context, Modern Hebrew is often presented as an exceptional case, where a language with distinct dialects corresponding to ethnic variation is undergoing a process of leveling, despite a persistent pattern of correlations between place of residence, wealth, and ethnicity (e.g., Ducker 2006, Tzfadia & Yiftachel 2008). The claim for leveling is based primarily on the decline in usage since at least the 1960s of three segmental features: the pharyngeals [h] and [ʕ] and the apical [r], the production of which puts speakers on the Mizrahi (Middle Eastern) side of the ethnic divide (Davis 1984; Bentolila 2002; Gafter 2019), contrary to Ashkenazi (European). Nevertheless, in this paper we show that Modern Hebrew speakers are sensitive to ethnically-conditioned phonetic variation – and in particular, rhythm – in both perception and production, even when the aforementioned markers are absent.

In study 1, listeners (N = 105) rated recordings of a single sentence on a 7-point scale (from “definitely Ashkenazi” to “definitely Mizrahi”). Stimuli were recorded by 14 young speakers who do not produce the known segmental markers, as well as by a control group of 14 older speakers, of whom some produce them. Listeners consistently rated half of the young voices as sounding Mizrahi or Ashkenazi. An exploratory acoustic analysis was conducted to detect possible acoustic correlates of these ratings. We found a correlation with rhythm (measured as in Ramus et al. 1999): the total duration of vocalic intervals in a phrase divided by the total duration of all segments in that phrase (see Figure 1). Specifically, a shorter proportion of vocalic intervals was correlated with Ashkenazi ratings.

Study 2 was designed to go beyond this correlation and examine whether rhythm participates in style shifting. Since style shifting is performative, modulation of a variant by style is taken to indicate that the variant is a marker (Labov 1972). We relied on the fact that there are Hebrew speaking actors who perform both Mizrahi and Ashkenazi characters. Four actors (2 men) of different personal ethnic backgrounds were chosen. Characters were selected according to their portrayed ethnicity, 2-3 short segments of freely available online content were used for each. Every sketch was divided into intonation phrases and fully segmented on PRAAT (Boersma 2006), as in Figure 1. The proportion of vocalic intervals was again significantly correlated with ethnicity in the expected direction: a shorter proportion of vocalic intervals was produced for Ashkenazi characters (see Figure 2).

Having established that rhythm is a sociolinguistic marker, we next show – based on discourse analysis of social media – that the social meaning of phonetic variation outside the segmental markers is contested, with some speakers associating it with ethnic identity and others with socio-economic status and place of residence. We discuss how the inter-relatedness of ethnicity, place of residence and wealth in Israel shaped the social meaning of variation in Hebrew, and why the ethnicity-related meaning of the segmental markers is so salient, compared to that of rhythm.

Bentolila, Y. (2002) Linguistic variation across generations in Israel. In S. Izre'el (ed.) *Speaking Hebrew: Studies in the Spoken Language and in Linguistic Variation in Israel*. (Te'uda 18) 131–140. Tel Aviv: Tel Aviv University Press.

Boersma, P. (2006). Praat: Doing phonetics by computer (version 4.4. 24). <http://www.praat.org/>

Davis, L. M. (1984). The pharyngeals in Hebrew: Linguistic change in apparent time. *Folia Linguistica Historica*, 18(Historica vol. 5, 1), 25–32.

- Ducker, C. (2006). Jews, Arabs and Arab Jews: the politics of identity and reproduction in Israel. ISS Working Paper Series/General Series, 421, 1-58.
- Gafter, R. J. 2019. "Modern Hebrew sociophonetics". Brill's Journal of Afroasiatic Languages and Linguistics. 11: 226-242.
- Labov, W. (1972). Sociolinguistic Patterns. University of Pennsylvania Press.
- Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265–292.
- Tzfadia, E., & Yiftachel, O. 2008. Sfariferia: the development towns and the Mizrahi space [Hebrew: Sfariferia: Arej ha-pituax ve-hamakom ha-mizrahi]. *Block: Architecture/ Media/ Theory/ City Journal* (6), 22-30.

## The triggering role of age-grading: The rise of simple pronouns over 40 years in Laurentian French

Hélène Blondeau<sup>1</sup>, Raymond Mougeon<sup>2</sup>, Mireille Tremblay<sup>3</sup>

<sup>1</sup>University of Florida, Gainesville, USA. <sup>2</sup>York University, Toronto, Canada. <sup>3</sup>Université de Montréal, Montréal, Canada

### Introduction

In varieties of Laurentian French simple forms of the non-clitic plural pronouns, i.e., nous, vous, elles, eux—in line with Standard French—alternate with complex vernacular counterparts nous autres, vous autres, eux autres (1-2).

Eux autres ils étaient tout le temps chez L'Amère à Boire mais eux ils avaient...eux c'était pas pire parce qu'ils étaient un peu loin de l'action. FRAN-HOMA 024F36

'Them, they were always at L'Amère à Boire, but them, they had... them, it wasn't bad because they were a bit far from the action.'

On a une célébration qui va être un peu bizarre pour nous mais nous autres on fête la journée de la Parade de Welland. FRAN-Welland 08F26 'We have a celebration which will be a bit weird for us but us, we celebrate the day of the Welland Parade.'

Previous research has found that in Québec French, the complex forms are gradually replaced by simple variants (Auger 1994; Blondeau 2011; Morin 1982). We re-examine this change with data collected in the 1970s and in the 2010s in Montréal, where French is in the majority and in Welland, where it is in the minority.

### Methods

We document change with real- and apparent-time data. We assess the impact of the following factors: pronoun type (simple vs. complex), speaker gender, age, and socioeconomic status.

### Results

Our results (Table 1) confirm the rise of the simple forms over time in both Montréal (14% à 48%) and Welland (10% à 22%). Multivariate analyses reveal that in both communities and at both periods the simple forms are favored by high SES speakers and women. However, in the 1970s, in Montreal, we observe a pattern of age-grading across the social classes and gender groups. Further, in the 2010s, there is a linear correlation with age, with the younger generation clearly at the vanguard of change. In contrast, in Welland, at both points in time, the younger generation is spearheading change.

### Discussion

According to Labov (1994:97), "it is possible that age-grading is involved in the mechanism of real-time change for certain types of linguistic change." The Montréal findings lend support to this hypothesis. We argue that, after the 1960s' Quiet Revolution, levels of education and literacy rose markedly in the 1970s, in Montreal and thus that the mid-age speakers (peak users of the simple forms) provided an increased input of simple forms to the younger speakers. This kickstarted the process of change, which then gained momentum in the 2010s due, in part, to: i) the growth of the French media, ii) the inflow of foreign-born Francophones—whose French lack complex pronouns—and iii) the continued rise of levels of education. In contrast, we point out that in Welland, the very limited opportunities to work in French account for the absence of age-grading. Further, the slower pace of change reflects notably the paucity of French-medium post-secondary education and of French media, as well as the marginal presence of Francophone immigrants.

### Corpora

Beniak, Edouard, Raymond Mougeon, et Daniel Valois. 1985. Contact des langues et changement linguistique: étude sociolinguistique du français parlé à Welland (Ontario). Publication / Centre international de recherche sur le bilinguisme, B-146. Québec: Centre international de recherche sur le bilinguisme.

Blondeau, Hélène, Mireille Tremblay, Anne Bertrand, et Elizabeth Michel. 2021. « A new milestone for the study of variation in Montréal French: The Hochelaga-Maisonneuve sociolinguistic survey ». *Corpus*, no 22 (janvier). <https://doi.org/10.4000/corpus.6221>.

Martineau, France, et Marie-Claude Séguin. 2016. « Le Corpus FRAN : réseaux et maillages en Amérique française ». *Corpus*, no 15 (octobre). <https://doi.org/10.4000/corpus.2925>.

Sankoff, David, Gillian Sankoff, Suzanne Laberge, et Marjorie Topham. 1976. « Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation grammaticale ». *Cahier de linguistique*, no 6: 85. <https://doi.org/10.7202/800043ar>.

Auger, Julie. 1994. « Pronominal Clitics in Québec Colloquial French: A Morphological Analysis ». University of Pennsylvania.  
Blondeau, Hélène. 2011. *Cet « autres » qui nous distingue: tendances communautaires et parcours individuels dans le système des pronoms en français québécois. Les voies du français*. Québec: Presses de l'Univ. Laval.

Labov, William. 1994/2010. *Principles of Linguistic Change. 1: Internal Factors*. Repr. *Language in Society* 20. Chichester: Wiley-Blackwell.

Morin, Yves-Charles. 1982. « De quelques [l] non étymologiques dans le français du Québec: Notes sur les clitiques et la liaison ». *Revue québécoise de linguistique* 11 (2): 9. <https://doi.org/10.7202/602486ar>.

---

## Individual and Group-variation: The Pragmatic Functions of Rising Tonal Accents on South Funen.

Kirstine Boas

Aarhus University, Aarhus, Denmark

Danish has the characteristic prosodic feature, *stød*, which is typically associated with a creaky articulation due to either a glottal stop or laryngelization in the vocal folds. The *stød* has a South Funen variant: a non-*stød*, which can be accompanied with a rising tonal accent (RTA). This presentation focuses on the RTA, which affects the intonation of the South Funen dialect to be more melodic than the Standard Danish intonation (Pedersen 2002: 46). Whereas the variant was a natural part of the traditional dialect (which is almost completely vanished (Pedersen 2002), it is abundant in the young generation today. It has thus never been investigated if the use of RTAs on South Funen varies indexically.

RQ's: Does the use of RTA vary indexically, and what meanings and values are connected to the use of South Funen RTAs that will arise in specific contexts? (Eckert 2008, 2012)

The analysis is based on 35 9th graders (21 identifying as girls and 14 as boys, 15-16 years) from a South Funen school, whom I have followed closely for a month to gather knowledge about their individual language use, social groups, values etc. Data from all the students includes self-recordings, school-group-work-recordings, 1:1-interviews and readings of a text with minimal pairs (*stød* – no *stød*) and monosyllables with high vowels, which traditionally often lead to a RTA (Andersen 1958: 36). Based on the readings, I have acoustically analyzed the RTAs in Praat (Boersma & Weenik 2023). I identify the RTA as a LH\* or LH\*H% tone on a significantly lengthened vowel (Andersen 1958; Boas 2023). I have auditorily coded this corpus for the presence of South Funen RTA and a range of social, pragmatic and interactional factors (inspired by Levon 2016). For this presentation I will focus on 6 selected students and their use of RTA.

Through a qualitative analysis, my results show that the RTA has become a resource the young generation in South Funen can choose to draw on in specific contexts, rather than it being a moribund feature of the language associated with older generations. In this corpus the variant varies indexically among the boys, not the girls. Several groups of boys used RTAs systematically when being together, but not necessarily individually. The RTA served a group membership function for group A and B, however there were indexical group specific differences. Group A had a more positive stance, and with the use of RTA they expressed chill, humor and having a good time. In group B it had a more negative stance; here the use of RTA expressed lazy and cocky attitudes, and fun in a more teasing sense. Both groups have members who switched into it during the breaks in school, while in class they would speak the standard variety. It seems that this traditional dialectic feature has a local, social benefit, and has been reallocated to be put at use differently by individuals seeking certain groups, even in young generations of Danish speakers.

Andersen, P. (1958). *Fonemsystemet i Østfynsk*. København: Schultz

Boas, K. (2023). Tone-*stød* og tonale accent i nyere fynsk dialekt. In: Boas, K., Hansen, I. S., Hougaard, T. T. & Overgaard, E.L. (Ed), 19. *Møde om Udforskningen af Dansk Sprog*.

Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.10, retrieved 3 May 2023 from <http://www.praat.org/>.

Eckert, P. (2008). Variation and the indexical field. In: *Journal of Sociolinguistics* 12/4, 453-476. DOI: 10.1111/j.1467-9841.2008.00374.x

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41, 87–100. DOI: 10.1146/annurev-anthro-092611-145828.

Levon, E. (2016). Gender, interaction and intonational variation: The discourse functions of High Rising Terminals in London. In: *Journal of Sociolinguistics* 20/2, 133-163. DOI: 10.1111/josl.12182

Pedersen, I. L. (2002). Hovedtrækkene i de danske dialekter. In: *Dialekter – sidste udkald?* 39-81. Modersmål-Selskabet & C.A. Reitzels Forlag, Odense.

---

## The use of the unstressed past tense prefix among Greek-Canadians

Laurens Bosman, Panayiotis Pappas

Simon Fraser University, Vancouver, Canada

We examine the usage pattern of a key dialectal feature in the speech of Greek immigrants to Canada during the period 1945 to 1975 in order to determine whether this pattern provides evidence of dialect contact among the population of these new communities. Specifically, we investigate the variation between the presence of the unstressed past tense prefix (/e/) (non-standard) and its absence (Standard Modern Greek, SMG, cf. Ralli, 2005). This prefix is added to the stem of verbs in order to form the past tense (perfective or imperfective). For example, the imperfective past of /'fevɣo/ 'I leave' is /'efevɣa/ 'I was leaving'. In the standard variety, this prefix is deleted if it is not stressed and this interacts with a stress assignment rule that requires that the stress is placed on one of the three final syllables (Kaisse, 1982). For example, the imperfective past of /psi'fizo/ 'I vote' is /'psifiza/ 'I was voting' and not \*/e'psifiza/. Thus, verb stems (i.e., root + past tense suffix) which have three or more syllables do not have such a prefix. However, there are dialects which maintain the prefix in tetrasyllabic or larger forms, such as those of Cyprus, the Ionian islands, the Peloponnese and Crete (Kontosopoulos, 2001).

Our dataset is based on a corpus of 443 recorded sociolinguistic interviews (Anastassiadis et al., 2017). We have examined 145 of these interviews (over 700,000 words), which have been fully transcribed, are balanced in terms of the sex of the speaker (70 females and 75 males) and represent 8 of the major Modern Greek varieties. The speakers were born mostly between 1935 and 1955 and moved to Canada when they were young adults (median age of 22 years old).

Up to this point, we have extracted (23,705) tokens of verb forms which can vary between a form without the unstressed prefix /e/ and one in which the prefix is maintained (e.g. /'psifiza/ vs. /e'psifiza/). Our preliminary analysis shows that speakers of varieties that maintain the unstressed prefix do so at a low rate (10 - 20%) suggesting that they have adopted the SMG pattern. At the same time we also see examples of usage of the non-standard form by speakers of varieties that delete the unstressed /e/. We hypothesize that the elevated use of the SMG pattern is evidence that, at the time of immigration, the synthesis of the common vernacular had already begun (Horrocks, 2010). We also suggest that the use of the unstressed /e/ by speakers of other non-standard varieties is based on it being a feature of Peloponnesian Greek whose speakers form the majority population in Greek-Canadian communities (Chimbos, 1980) and is considered the most prestigious dialect due to its resemblance to SMG (Browning, 1983; Horrocks, 2010). The convergence of these two factors (preponderance and prestige) may have led some speakers from other dialects to interpret its use as an acceptable variant of the Standard.

Anastassiadis, A., Ralli, A., Gekas, A., Papanagiotou, C., Siotou, A., Tsimpouris, C., & Tsolakidis, S. (2017). Immigration and Language in Canada: Greeks and Greek Canadians. [Electronic Database]. <https://immigrec.com/en>.

Browning, R. (1983). *Medieval and Modern Greek* (2nd ed.). Cambridge University Press.

Chimbos, P. (1980). *The Canadian Odyssey: The Greek Experience in Canada*. McClelland Stewart Ltd.

Horrocks, G. C. (2010). *Greek: A History of the language and its speakers* (2nd ed.). Wiley-Blackwell.

Kaisse, E. M. (1982). On the preservation of stress in Modern Greek. *Linguistics*, 20, 59–82.

Kontosopoulos, N. (2001). *Διάλεκτοι και ιδιώματα της Νέας Ελληνικής [Dialects and varieties of Modern Greek]* (3rd ed.). Grigori Publications.

Ralli, A. (2005). *Morphology*. Athens: Patakis.

---

## The Universality of Polarity Sensitivity

Josh Phillips, Claire Bower

Yale University, USA

“NPIs are typologically very common ... and seem to exist in virtually every language we consider” (Giannakidou 2011: 1661; cf. 2019: 70). We report the results of a comparative study of languages of Indigenous Australia, showing that Negative Polarity Items (NPIs) are absent from almost all of these languages. We discuss results and their implications for cross-linguistic semantic variation, particularly regarding scalarity and veridicality.

An influential line of research on polarity sensitivity concerns their participation in scales (e.g. Horn 1972, Krifka 1995, Israel 2011 a.m.o). We relate the paucity of NPIs reported for Australian languages to these languages' apparent lack of explicit comparatives and (other) degree constructions, paralleling Geraci et al.'s 2022 modality-specific suggestions for signed languages, which also tend to lack NPIs (Quer 2020). Alternative analyses (e.g. Giannakidou 2019) relate polarity sensitivity to (non)veridicality.

46 Australian languages were surveyed for empirical material on NPIs. While no grammar contained overt discussion of polarity sensitivity, all grammars included sentences where phrases were translated with English NPIs (e.g. any, yet, ever, at all). From this material, we compiled a comparative dataset illustrating concepts in negative polar contexts. The languages are drawn from all primary divisions of Pama-Nyungan and a range of non-Pama-

Nyungan families from across north Australia (see Bower 2023), to maximize genetic and areal coverage. Though a broad range of NPI contexts were surveyed, we focus here on items translated with ‘indefinite’ pronominals (anyone, anything) and aspectual operators (e.g. yet, anymore).

In only one of the languages was there clear evidence of an NPI. The Diyari intensifier *marla* occurs in negative polar contexts with a “cessative” aspectual function, corresponding to ‘anymore’ (1). From comparative and sociohistorical evidence, this usage (in addition to a related usage in comparative constructions) is likely a contact-induced innovation (Wegner et al. 2022). Grammars of 23 languages report the use of ‘ignorative’ pronouns (cf. Wierzbicka 1980, Haspelmath 1997, cf. Mushin’s 1995 “epistememes”). These items see use both as indefinites — in free choice and negative quantificational contexts — and as *wh*-words (interrogatives), as illustrated for Gurindji in (2). There is no evidence that these items show polarity sensitivity; they appear equally felicitous in both negative and positive environments. Emphatic negation also frequently seems to be realized by negative interjections/quantifiers (3).

Given the wealth of linguistic literature on polarity sensitivity and the prediction—in semantic and pragmatic theory—that polarity sensitive items should be ubiquitous cross-linguistically, data from Australia provides a challenge. It also raises a conundrum. Given that Australian languages tend to lack degree constructions (Bowler & Kapitonov 2023), it is tempting to correlate the lack of NPIs with the lack of degree semantics. However, Australian languages do show polarity sensitivity in the modal domain (McGregor & Wagner 2006; Phillips 2023), providing empirical evidence for a link between polarity and veridicality. This contradiction points to a need for further work on degreeless languages.

Bower (ed.) 2023. *The Oxford Guide to Australian Languages*. OUP

Bowler & Kapitonov. 2023. Quantification. In Bower (ed.) 2023.

Ennever, T. 2021. *A grammar of Ngardi*. Berlin: De Gruyter

Haspelmath, M. 1997. *Indefinite pronouns*. Oxford: OUP

Geraci C, M Oomen & M Santoro 2022. As strong as an NPI in LSF, NGT and LIS. *Semantics and Linguistic Theory* 32. doi: 10.3765/salt.v1i0.5353

Giannakidou, A 2011. Negative and Positive Polarity Items. In C Maienborn, K von Stechow & P Portner (eds.) *Semantics: An international handbook of natural language meaning*. Berlin: De Gruyter

Horn, L. 1972. *On the semantic properties of logical operators in English*. UCLA, PhD Dissertation

Israel 2011. *The grammar of Polarity: Pragmatics, Sensitivity, and the Logic of Scales*. Cambridge: CUP

Krifka, M. 1995. Semantics & pragmatics of polarity items. *Linguistic Analysis* 25

Meakins, F & P McConwell. 2021. *A grammar of Gurindji*. Berlin: De Gruyter

McGregor, W & T Wagner. 2006. *Ocean. Linguist.* 45(2). 339–79.

Mushin, I. 1995. Epistememes in Australian languages. *Aust. J Linguist.* 15. 1–31.

Phillips, J. 2023. Negation. In Bower (ed.) 2023.

Quer, J. 2020. Expression of Negation in Sign Languages. In V Déprez + M-T Espinal (eds.) *The Oxford Handbook of Negation*. Oxford: OUP

Wierzbicka, A 1980. *Lingua Mentalis*. Sydney: Academic Press.

---

## Salience and Southernness: A perception study on [hw]

Keiko Bridwell

University of Georgia, Athens, GA, USA

In United States English, morpheme-initial exhibits sociophonetic variation between competing segments [hw] and [w]. Once the standard and prestige variant, [hw] has over time lost ground to [w], now largely remaining only in the speech of Southern speakers over the age of 50 (Labov, Ash & Boberg 2006; Bridwell 2018; 2019). Yet while a limited number of studies have been conducted to determine the locations in the US where [hw] is spoken and to track the history of its decline (McDavid & McDavid 1952; Kurath & McDavid 1961; Labov, Ash & Boberg 2006), no full study has to date investigated how this feature is perceived. One exploratory study has shown that, among online commenters, [hw] is associated with Southernness and with older speakers (Bridwell 2019), but current online usage suggests a more complex social meaning. In social media discourse, the spelling or pronunciation *hw*, particularly in the word “White”, is used to mockingly reference or mimic the stereotypical persona of a casually racist White person. It is unclear, however, whether this usage derives from the perception that conservative Southerners actually use [hw], or whether it is a form of mocking hypercorrect, orthographically-based speech. Drawing on a tradition of perceptual dialectology, the present study seeks to address this question by testing whether [hw] is a salient, noticeable feature, and if so, if it is perceived as Southern by US listeners.

Data for this experiment was collected via an online survey in which 170 participants listened to six speakers, some natively producing [hw] and others [w], read the same passage aloud, and were asked to identify the area of the United States the speaker was from and to click on the specific words that served as cues to this region. Data from the three Southern speakers was considered for this study: an older man using [hw], a younger man using [hw],

and a younger woman using [w]. The words clicked by participants and the perceived regions assigned to speakers were then evaluated to investigate how often [hw] was noticed and whether clicking on it influenced participants' regional categorization of a speaker.

Results showed that [hw] was noticed at a high rate: [hw] words received a relatively high number of clicks relative to words with other Southern features, and [hw] was explicitly commented on more frequently than any other feature. However, this finding was not consistent across speakers: the younger male speaker's [hw] words were clicked more frequently than the female speaker's [w] words, but this was not true of the older male speaker, suggesting that salience may be largely driven by the unexpectedness of a token given who is using it or how it is produced. Additionally, participants who clicked on [hw] were not more likely to judge the speaker as Southern. These results suggest that while listeners are highly sensitive to the presence of [hw], it has little connection with Southernness for the typical speaker of US English, and is instead perceived as a strange or unusual, but aregional, feature.

Bridwell, Keiko. 2018. The distribution of [ɹ]: Geographic and age-related factors governing the wine-whine merger. Presented at the LCUGA 5, University of Georgia, Athens GA.

Bridwell, Keiko. 2019. The distribution of [ɹ]: An acoustic analysis of sociophonetic factors governing the wine-whine merger in Southern American English. University of South Carolina MA Thesis.

Kurath, Hans & Raven I. Jr. McDavid. 1961. The Pronunciation of English in the Atlantic States: Based upon the Collections of the Linguistic Atlas of the Eastern United States. Ann Arbor: University of Michigan Press.

Labov, William, Sharon Ash & Charles Boberg. 2006. The Atlas of North American English: Phonetics, phonology, and sound change: A multimedia reference tool. Berlin; New York: Mouton de Gruyter.

McDavid, Raven I. & Virginia Glenn McDavid. 1952. h before semivowels in the Eastern United States. *Language*. JSTOR 28(1). 41–62.

---

## “We was one big happy family there”: tracking the evolution of past BE in Falkland Island English

David Britain, Hannah Hedegard  
University of Bern, Bern, Switzerland

Non-standard constellations of past tense BE are common across varieties of English. These include levelling to was (e.g. it was, we was) (e.g. Feagin 1979, Smith & Tagliamonte 1998), levelling to were (e.g. it were, we were) (e.g. Moore 2010), and a split system whereby levelling to was is found in positive sentences and levelling to weren't in negative (e.g. it was, weren't it?, we was, weren't we?) (Schilling-Estes & Wolfram 1994). Further linguistic constraints have been identified, including a favouring of was with plural existentials (Meechan & Foley 1997), a preference for weren't levelling in tags (e.g. Tagliamonte 1998) and a preference for was after plural NPs rather than after pronoun they (the NP>Pro constraint, e.g. Smith & Tagliamonte 1998).

Investigations of Southern Hemisphere Englishes have suggested that a) there is a preference for levelling to was (e.g. Schreier 2002 for Tristan da Cunha), b) much more non-standardness in existentials than with other subject types (e.g. Hay & Schreier (2004) for New Zealand), and c) mixed evidence of the NP>Pro constraint – present in 19th century New Zealand, not in Australia. Many such studies, however, are based on a narrow demographic (adolescents only in Eisikovits (1991)) or focus largely on non-standardness in existentials since variability in other contexts has almost disappeared (e.g. Hay & Schreier (2004) for 20th century New Zealand).

Here we investigate past BE in the English of the Falkland Islands (FIE), a British Overseas Territory in the South Atlantic Ocean, settled in the 19th century by migrants largely from Scotland and Southwest England. We base our analysis on a corpus of almost 200 hours of informal conversation with 120 Islanders. Recordings were transcribed into ELAN, with the corpus totalling 920,000 words. All unambiguous tokens of past BE were extracted, 17087 in total. Each was coded for a range of linguistic and social constraints. Results showed, first, that there was virtually no non-standardness in past BE with singular subjects. Secondly, and as in many other studies (e.g. Hay & Schreier 2004), past BE in plural existentials behaved very differently than after other plural subjects. We fitted a mixed-effects binomial logistic regression model on the 3305 tokens of non-existential plural forms, with a range of social and linguistic factors as fixed effects and speaker as random intercept. Age(\*\*), gender(\*\*\*), urban/rural socialisation(\*\*\*), subject(\*\*\*), animacy(\*\*\*), polarity(\*\*\*) and word order(\*\*\*) were all shown to be significant by the model.

As in New Zealand, non-standardness is declining over time but not as rapidly and the NP>Pro constraint remains robust – a feature common also to Scottish and Southwestern English varieties of British English. Non-standardness is more common among male, rurally-socialised speakers as well as with inverted word order, with inanimate subjects, and, unusually, in negated verbs. Overall, then, FIE appears broadly typical of southern Hemisphere Englishes, but has retained non-standardness longer than elsewhere, and maintains the NP>Pro constraint characteristic of its settler varieties.

- Eisikovits, E. (1991). Variation in subject-verb agreement in inner Sydney English. In J. Cheshire (Ed.), *English around the world: Sociolinguistic perspectives* (pp. 235-256). Cambridge: Cambridge University Press.
- Feagin, C. (1979). *Variation and change in Alabama English: A sociolinguistic study of the white community*. Washington: Georgetown University Press.
- Hay, J., & Schreier, D. (2004). Reversing the trajectory of language change: Subject-verb agreement with *be* in New Zealand. *Language Variation and Change*, 16, 209-235.
- Meechan, M., & Foley, M. (1994). On resolving disagreement: Linguistic theory and variation - *there's* bridges. *Language Variation and Change*, 6, 63-85.
- Moore, E. (2010). The interaction between social category and social practice: Explaining *was/were* variation. *Language Variation and Change*, 22, 347-371.
- Schilling-Estes, N., & Wolfram, W. (1994). Convergent explanation and alternative regularization patterns: *Were/weren't* levelling in a vernacular English variety. *Language Variation and Change*, 6(3), 273-302.
- Schreier, D. (2002). Past *be* in Tristan da Cunha: The rise and fall of categoricity in language change. *American Speech*, 77, 70-99.
- Smith, J., & Tagliamonte, S. A. (1998). 'We were all thegither, I think we was all thegither': *Was* regularization in Buckie English. *World Englishes*, 17, 105-126.
- Tagliamonte, S. A. (1998). *Was/were* variation across the generations: View from the city of York. *Language Variation and Change*, 10, 153-192.

---

## A sociophonetic study of creaky voice in Canadian English and French

Jeanne Brown

McGill University, Montreal, Canada

Creaky voice, otherwise known as vocal fry, is a voice quality often characterized by low pitch, irregular vocal pulses, and decreased glottal airflow. Older studies of creaky voice in English find that it is commonly associated with men's speech and therefore can be interpreted as a sign of masculinity or authority (Henton & Bladon, 1988; Klatt & Klatt, 1990; Laver, 1968). However, more recent studies find that women use creak more often than men (Podesva, 2013; Yuasa, 2010). Yuasa (2010) suggests that increased use by young upwardly-mobile American women is causing a perceptual shift towards a more nonaggressive, urban-oriented, and informal interpretation. Despite previous work, claims of gender effects and sound change in progress are contentious, requiring more empirical investigation. Additionally, non-modal voice quality remains largely understudied in Canada, especially with respect to bilingualism.

The present study is a sociophonetic analysis of creaky voice in Canadian English and French which examines the acoustic correlates of creak across language, gender, age, and time. Spontaneous speech from monolingual and bilingual speakers was collected from videos online to create a current (2020s) corpus which was then compared to an older (1990s) corpus of broadcast speech, ICE-Can (Greenbaum & Nelson, 1996). One spectral slope measure ( $H1^*-H2^*$ ) and two Harmonics-to-Noise Ratios (CPP and HNR from 0-500Hz) were extracted at vowel mid-points. Spectral slope measures are acoustic correlates of glottal constriction and HNRs measure waveform periodicity. Vowel quality, stress, phonological context, and utterance position were also included as factors. Comparisons between the acoustic measures of creak in relation to the social and linguistic factors were analyzed using linear mixed-models.

The main results (see Figure 1) show that men's vowels have less reliable  $f_0$  tracks, lower spectral slope measures ( $H1^*-H2^*$ ) and lower HNR measures (CPP and HNR05). Altogether, these acoustic cues indicate more glottal closure and higher levels of noise/aperiodicity, ultimately providing evidence of stronger creak for men than for women. Notably, this finding is in direct opposition with recent and frequent claims of increased creaky voice among young women. There is also a slight effect of age in the data, interacting with gender. For men, younger speakers produce more creak (lower  $H1^*-H2^*$  and CPP values), whereas for women, no age effect is found. When comparing corpora roughly 20 years apart, few consistent differences in the acoustic measures arise, suggesting that there is limited if any evidence of diachronic change with regards to the prevalence of creaky voice. As for language, the subset of bilingual speakers shows mixed results and much individual variation (see Figure 2). Patterns that emerge are lower  $H1^*-H2^*$  values (more creak) in English than in French for young women, but the inverse for older women and men. HNR05 values are also lower (creakier) in English than in French for two thirds of the bilingual speakers. This study provides new insight into the relation between creak and gender while also highlighting the importance of acoustic measures in quantifying creak and cautioning against assertions of widespread sound change without larger-scale empirical inquiry.

- Greenbaum, S., & Nelson, G. (1996). The International Corpus of English (ICE project). *World Englishes*, 15:3-15.
- Henton, C., & Bladon, A. (1988). Creak as a sociophonetic marker. In L. Hyman & C. Li (Eds.), *Language, speech and mind: Studies in honor of Victoria A. Fromkin* (pp. 3-29). London, England: Routledge.



- Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87(2), 820–857. <https://doi.org/10.1121/1.398894>
- Laver, J.D.M. (1968). Voice quality and indexical information. *International Journal of Language & Communication Disorders*, 3, 43-54. <https://doi.org/10.3109/13682826809011440>
- Podesva, R. J. (2013). Gender and the social meaning of non-modal phonation types. In C. Cathcart, I. Chen, G. Finley, S. Kang, C. S. Sandy, & E. Stickles (Eds.), *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society*, pp. 427-448
- Yuasa, I. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85, 315-337. [10.1215/00031283-2010-018](https://doi.org/10.1215/00031283-2010-018).

---

## Spanish grammatical gender assignment of Basque lexical items: Exploring the role of the double definite article construction

Alicia Brown

University of Arizona, Tucson, USA

Previous work on grammatical gender assignment to lexical items inserted into phrases of another language have pinpointed three general strategies: the unmarked, ‘default,’ assignment, generally the masculine; the grammatical gender of the translation equivalent; and the phonological form, wherein grammatical gender is assigned based on frequency correlations of grammatical gender to the phonological shape of the word. In the Spanish/Basque context, a double definite article construction has emerged, wherein the Spanish-language clitic and the Basque-language suffix are both applied to an individual lexical item (e.g., *liburutegi*, ‘library’, *la liburutegia*, ‘DEF-FEM library+DEF’). As the (standard) Basque singular definite article suffix is *-a*, this presents the possibility of a higher application of the phonological shape strategy, as the suffix is homophonous with the canonically feminine morphological ending. The present study uses a story-elicitation task with Basque/Spanish speakers in Vitoria-Gasteiz (n=42) to analyze the application of these strategies and includes the double definite article construction as a factor.

While each strategy (default, analogy, phonological form) is widely accepted, what is less agreed on is what each strategy’s relative influence is and whether they act independently or in tandem. Previous studies have found phonological shape to override analogical assignment (Dubord, 2004, for English/Spanish in Arizona); and analogical gender to be a stronger factor than phonological shape (Trawick & Bero, 2022, for English/Spanish in New Mexico). Specific to Basque/Spanish, both analogy and phonological shape were found to be relevant factors (Parafita Couto et al., 2016), with some arguing that the double definite article construction triggering high rates of feminine gender assignment due to the resultant phonological form is causing a shift from the masculine to the feminine for the default assignment (Deuchar et al., 2010, Parafita Couto et al., 2016). Following Badiola and Sande (2018) and Nicolás and Badiola (2019), this paper views the double definite article construction as a possible source for higher rates of feminine assignment due to the phonological form assignment strategy, without positing that the default strategy has shifted from masculine to feminine.

The present study reports on the findings from a variationist approach to a story elicitation task, part of a larger project on grammatical gender assignment in Vitoria-Gasteiz, Spain, a region known for speaking standard Basque (Batua). Participants were given a children’s book and asked to narrate a story in Spanish and including the Basque words provided. Results indicate that the double definite article construction is more likely to be employed when the translation equivalent is grammatically feminine, thereby changing the word’s form to that of a canonically grammatically feminine Spanish word. In addition to not supporting the hypothesis of the default gender shift in this region, this also suggests that the translation equivalent is the predominant strategy: since the majority of the participants are sequential bilinguals, this is in line with findings from Cantone and Müller (2008), Licerias et al. (2008), and Trawick and Bero (2022), particularly for Latin-based words (e.g., cognate similarity, cf. Gradoville, Waltermire, & Long, 2021)

- Badiola, L., and Sande, A. (2018). Gender assignment in Basque/Spanish mixed determiner phrases. *Issues in Hispanic and Lusophone Linguistics*, 19, 15–38.
- Cantone, K., & Müller, N. (2008). Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language facility. *Lingua*, 118, 810–826.
- Deuchar, M., Epelde, I., Oyharçabal, B., & Parafita Couto, M. C. (2010). Gender agreement in Spanish–Basque and Spanish–English nominal constructions. XXXIX Simposio Internacional de la SEL [Conference Presentation]. Santiago de Compostela, Spain.
- DuBord, E. M. (2004). Gender assignment to English words in the Spanish of Southern Arizona. *Divergencias: Revista de estudios lingüísticos y literarios*, 2, 27–39.
- Gradoville, M., Waltermire, M. and Long, A. (2021). Cognate similarity and intervocalic /d/ production in Riverense Spanish. *International Journal of Bilingualism*, 25, 727-746.

- Liceras, J. M., Fernández Fuentes, R., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and number agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua*, 118, 761–852.
- de Nicolás, I., and Badiola, L. (2019). Asignación de género en sintagmas determinantes con alternancia de código de castellano/euskera. *Deusto Workshop on Basque Linguistics*.
- Parafita Couto, M.C., Munarriz, A., Epelde, L., Deuchar, M., & Oyharcabal, B. (2016). Gender conflict resolution in Spanish-Basque mixed DPs. *Bilingualism: Language and Cognition*, 18(2), 304-323.
- Schwartz, A., and Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, 55(2), 197-212.
- Trawick, S. & Bero, T. (2022). Gender assignment: Monolingual constraints contribute to a bilingual outcome. *International Journal of Bilingualism*, 26(2), 181-197.

---

## Language attitudes and stylistic variation over time: Results of a panel study from Ulrichsberg (Austria)

Lars Bülow, Philip Vergeiner  
University of Munich, Munich, Germany

While there are several studies that investigate the relationship between language attitudes and linguistic variation and change from a synchronic perspective using cross-sectional data, panel studies that examine this relationship for the same speakers at different points in time are rare for research pragmatic reasons. This is particularly true for Austria, where there is only one larger panel study – the Ulrichsberg project (Vergeiner et al. 2022; Wallner in prep.). In this project, data on language attitudes were collected in addition to data on language use (Scheutz 1985; Wallner in prep.). Data collection took place in Ulrichsberg, a small rural village with about 2,800 inhabitants in the Central Bavarian dialect region of Austria (see Figure 1) where the same twelve speakers (six males and six females) were recorded in two settings (formal interview and informal conversation among friends), once in 1975/76 and once in 2018/19. The Ulrichsberg project thus covers variation and change over 43 years. Initial results on phonetic-phonological change indicate retrograde change for some speakers and variables, i.e. the increasing use of dialect features, especially in the 2018/19 formal interviews (Vergeiner et al. 2020; Bülow & Vergeiner 2021; Vergeiner et al. 2022). However, these findings have not yet been correlated with data on language attitudes.

This paper examines the relationship between language attitudes and phonetic-phonological variation and change in the Bavarian dialect of Ulrichsberg over a 43-year period. In order to operationalize language attitudes, specifically language identity, a Bavarian Orientation Index (BOI) was created for each point in time (for a similar approach, see Beaman 2021). The effect of language attitudes on 17 phonetic-phonological variables belonging to two different lects – the local base dialect (Basisdialekt) and the regional contact dialect (Verkehrsdialekt) (see Table 1; for the operationalization of these two lects cf. Vergeiner et al. 2022) – was measured using linear mixed models.

The results indicate that attitudes towards the Bavarian dialect have remained stable over 40 years. Moreover, it is shown that positive attitudes towards the dialect have a positive effect on dialect use. However, the effect of language attitudes varies between the two different types of dialect features (local base dialect features vs. contact dialect features). In formal settings, a higher BOI leads to increased use of contact dialect features, whereas in informal settings a higher BOI has a stronger effect on the use of local base dialect features (see Table 2). Thus, the study highlights the complex interplay between linguistic, situational, and attitudinal factors over time.

- Beaman, K. V. (2021). Identity and Mobility in Linguistic Change across the Lifespan: The Case of Swabian German. In A. Ziegler, St. Edler, N. Kleczkowski & G. Oberdorfer (eds). *Urban Matters. Current Approaches of International Sociolinguistic Research*. Amsterdam: John Benjamins, 27–60.
- Bülow, L. & Vergeiner, P. C. (2021). Intra-Individual Variation Across the Lifespan: Results from an Austrian Panels Study. *Linguistics Vanguard* 7(s2):1–11.
- Scheutz, H. 1985. *Strukturen der Lautveränderung. Variationslinguistische Studien zur Theorie und Empirie sprachlicher Wandlungsprozesse am Beispiel des Mittelbairischen von Ulrichsberg/Oberösterreich*. Wien: Braumüller.
- Vergeiner, P. C., Wallner, D. & Bülow, L. (2022). Language Change in real-time. 40 years of Lectal Coherence in the Central Bavarian Dialect-Standard Constellation of Austria. In K. V. Beaman & G. R. Guy (eds.), *The Coherence of Linguistic Communities. Orderly Heterogeneity and Social Meaning*. London: Routledge, 281-300.
- Vergeiner, P. C., Wallner, D., Bülow, L. & Scheutz, H. (2020). Redialektalisierung und Alter. Ergebnisse einer real-time-Studie zum age-grading in Ulrichsberg. In H. Christen, B. Ganswindt, J. Herrgen & J. E. Schmidt (eds.), *Regiolekt – Der neue Dialekt? Akten des 6. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD)*. Stuttgart: Franz Steiner, 101–122.
- Wallner, D. (in prep.). *Language Change in Real Time: A Panel Survey on Dialectal Sound Change in Ulrichsberg/Upper Austria*. PhD Thesis. Universität Salzburg, Salzburg.

---

## Replenishing an impoverished inventory: Evidence of attenuative typographical variation from ten European languages.

Sophia Burnett

CY Cergy Paris, Cergy, France

Non-standard minuscules are not the reserve of the English language but are making syncretic appearances in CMC in other languages of Latin script. The aim of this paper is to provide evidence that languages such as English, Dutch, Italian, are augmenting the relative paucity of their inventories with attenuative allomorphs in lieu of ex novo glyphs. The evolution of scripts is driven by the pressure of simplification (Milton & Morin 2021), so there naturally comes a time in which the script becomes oversimplified (by this I do not infer that complex scripts are ‘unevolved’). Standardised non-lexical signs for emphasis abound (Heath, 2018); ALLCAPs, italics, bold, exclamation marks(!), however there are no standardised non-lexical attenuators. Non-standard minuscules impact word perception in eye-tracking (Pauly & Nottbusch, 2020), and I propose that it disturbs perception of indexical order, which according to Silverstein (2003) is central to analysing how semiotic agents access macro-sociological plane categories and concepts as values in the indexable realm of the micro-contextual. In English, we have recently witnessed the appearance of the lowercase ‘l’ not as a replacement, but as a complement (Burnett, 2022), proposing a novel grapheme-phoneme pairing—a non-dialectical monophthongisation—which doubles authorial choice in both iconic self-representation, but also prosodically; the shift from /ai/ to /ə/ , /ɒ/ , or /ʌ/ affording a ‘prosodic slouch’ which can be useful to offset impassioned lexicon, for example. These variant forms are allomorphs, with the same functional layer in the syntax as their majuscule ‘parent’, but with a novel descriptive layer. Non-standard orthography can be observed as pure indexicals (Androutsopoulos, 2011), or as sociophonetically enregistered signs (Ilbury, 2020) (Agha, 2015), depending on usage and contingency. The diacritic/non-diacritic data is approached as glyphs with semantic potential; as visual objects (see Kinoshita et al., 2021) and as graphic signs (typography) (Spitzmüller, 2022) hence, we shall not restrict our definition of diacritic languages to those with diacritics indexing pitch patterns, but also those indexing vocalic pronunciation, such as the German umlaut. The central hypothesis is that social actors relying on small inventories have seized the opportunity to implement non-standard lower case glyphs as supplementary signs and grapheme-phoneme pairings.

I will present a brief case-study to demonstrate how minuscule variants are pragmatically useful in CMC, allowing social actors to attenuate their Self or their discourse. We shall then see how  $\pm 100k$  tokens from ten other languages of Latin script were ‘Wordsworth’ ranked for diacritic use, then analysed for evidence of non-standard minuscules—to wit, lower-case letters at the start of sentences. Five of the languages have commonly used diacritics (Lithuanian, Latvian, Romanian, French, Croatian), the other five languages less so, or—as is the case with English—no diacritics at all (Slovenian, German, Italian, Catalan, Dutch). Finally we shall discuss the results ( $\alpha$  0.01,  $p < 0.001$ ) that show greater use of non-standard minuscules in the five ‘diacritic poor’ languages, and fewer in the five ‘diacritic rich’ languages, supporting the hypothesis of a user led replenishing of inventories.

- Agha, A. (2015). Enregisterment and communication in social history. *Registers of Communication*, 18(1), 27-53.
- Androutsopoulos, J. (2011). Language change and digital media: a review of conceptions and evidence. *Standard languages and language standards in a changing Europe*, 1, 145-159.
- Burnett, S. (2022). Offsetting love and hate: The prosodic effects of the non-standard lsg in tweets to Boris Johnson and Jeremy Corbyn over four days of the UK general election. *European Journal of Applied Linguistics*, 10(2), 388-414.
- Chomsky, N., et Halle, M. (1968) *The Sound Patterns of English*. Harper & Row.
- Dubois, Jean. (2002) *La linguistique textuelle [Text linguistics]*. Paris : Armand Colin
- Eisenstein, J. (2015) Systematic patterning in phonologically motivated orthographic variation.
- Heath, M. (2018). Orthography in social media: Pragmatic and prosodic interpretations of caps lock. *Proceedings of the Linguistic Society of America*, 3(1), 55-1.
- Ilbury, C. (2020). “Sassy Queens”: Stylistic orthographic variation in Twitter and the enregisterment of AAVE. *Journal of sociolinguistics*, 24(2), 245-264.
- Kinoshita, S., Yu, L., Verdonshot, R. G., & Norris, D. (2021). Letter identity and visual similarity in the processing of diacritic letters. *Memory & Cognition*, 49, 815-825.
- Spitzmüller, J. (2015) Graphic variation and graphic ideologies: a metapragmatic approach, *Social Semiotics*, 25 :2, 126-141
- Spitzmüller, J. (2022). Approaching typographic variation. *Vestnik of Saint Petersburg University. Language and Literature*, 18(4), 910-928.
- Tholly, A. (2022). *Analyse sémantique de la prosodie imitative* (Doctoral dissertation, Sorbonne université).
- Miton, H., & Morin, O. (2021). Graphic complexity in writing systems. *Cognition*, 214, 104771.
- Nguyen, D., Liakata, M., DeDeo, S., Eisenstein, J., Mimno, D., Tromble, R., & Winters, J. (2020). How we do things with words: Analyzing text as social and cultural data. *Frontiers in Artificial Intelligence*, 3, 62.
- Pauly, D. N., & Nottbusch, G. (2020). The influence of the German capitalization rules on reading. *Frontiers in Communication*, 5, 15.
- Pincemin, Bénédicte. (2018) *Sept logiciels de textométrie [Seven textometry software tools]*. halshs-01843695f.

Ryland Williams, J., Lessard, P., Desu, S. et al. (2015) Zipf's law holds for phrases, not words. *Sci Rep* 5, 12209.  
Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & communication*, 23(3-4), 193-229.  
Zaefferer, D. (1990). On the coding of sentential modality. *Toward a typology of European languages*, 215-238.

---

## The systematicity behind US Spanish variable grammar: The case of mood variation in Southern Arizona

Isabella Calafate

Baylor University, Waco, USA

Mood selection in Spanish has been traditionally approached from a prescriptive perspective and a normative effort has been made for dictating when subjunctive should and should not be used. Based on obligatory and variable contexts of subjunctive, several studies claim that US Spanish speakers overextend the use of indicative in contexts where subjunctive is expected. These studies (e.g., Silva-Corvalán, 1994; Lynch, 1999; Montrul, 2007, 2009) often conclude that this is due to contact with English, claiming that US Spanish speakers' use of subjunctive results from simplification or even incomplete acquisition. Nevertheless, mood selection in Spanish is much more variable than prescribed, illustrating yet another case of the little overlap that exists between prescription and praxis (Poplack et al., 2013).

Following the quantitative methods of variationist sociolinguistics, this study investigates the linguistic factors conditioning mood variation in the Spanish variety spoken by bilinguals in Southern Arizona. Based on data extracted from 48 sociolinguistic interviews from the Corpus of Spanish in Southern Arizona (Carvalho, 2012-), this research offers innovative analytical methods by combining two different types of analysis of linguistic factors. First, it explores the impact of different productivity measures on mood variation, following LaCasse (2018), Poplack et al. (2018), Schwenter and Hoff (2020), and Torres Cacoullós et al. (2017). Then, it applies a multivariate mixed-effect logistic regression run in R (Picoral, 2020) that includes governor frequency, polarity, coreferentiality, and morphological form of the embedded verb as fixed effects and lexical identity of the governor and speaker as random effects.

Findings uncover the systematicity and complexity of mood variation in US Spanish, typically hidden by comparisons of bilingual Spanish to idealized monolingual standards resulting from prescriptive accounts (Bessett & Carvalho, 2022; Otheguy, 2016). Mood variation is found to be mainly determined by the lexical identity and frequency of the governor, supported by other linguistic factors such as polarity, and coreferentiality. Subjunctive is triggered by a set of non-frequent governors and favored by negative polarity and non-coreferential subjects. In addition, results call attention to the interaction between governor and polarity, establishing that research on mood selection in Spanish needs to consider the effect of polarity on specific governors.

In addition to advancing our understanding of the linguistic factors underlying mood variation in Spanish in general, these findings shed light on variable patterns in US Spanish specifically. Findings point to clear continuities across monolingual and bilingual varieties, countering previous claims of grammatical simplification in US Spanish and revealing patterns that show both systematicity and complexity of mood variation. Ultimately, by integrating two types of analysis, this variationist study offers strong evidence that counters previous generalizations about the effect of contact in US Spanish and highlights the integrity of its variable grammar.

- Bessett, R. & Carvalho, A. M. (2022). The structure of US Spanish. In S. Loza and S. M. Beaudrie (Eds.) *Heritage Language Teaching: Critical Language Awareness Perspectives for Research and Pedagogy*. (pp. 44–61). New York: Routledge.
- Carvalho, A. M. (2012-). *Corpus del Español en el Sur de Arizona (CESA)*. University of Arizona. [cesa.arizona.edu](http://cesa.arizona.edu).
- LaCasse, D. L. (2018). *The subjunctive in New Mexican Spanish: Maintenance in the face of language contact*. [Doctoral dissertation, The Pennsylvania State University].
- Lynch, A. E. (1999). The Subjunctive in Miami Cuban Spanish: Bilingualism, Contact, and Language Variability. 60(8), 2895-A.
- Montrul, S. (2007). Interpreting mood distinctions in Spanish as a heritage language. In K. Potowski & R. Cameron (Eds.), *Spanish in Contact: Policy, Social and Linguistic Inquiries* (pp. 23–40). John Benjamins Publishing Company.
- Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior*, 13(2), 239–269.
- Otheguy, R. (2016). The linguistic competence of second-generation bilinguals: A critique of “incomplete acquisition”. In C. Tortora, M. den Dikken, I. L. Montoya & T. O’Neill (Eds.), *Romance Linguistics 2013: Selected papers from the 43rd Linguistic Symposium on Romance Languages (LSRL)* (pp. 301-319). Amsterdam/Philadelphia: John Benjamins.
- Picoral, A. (2020, December 20). *Quantitative Language Data Analysis in R: Regression and Contrasts* [Blog post]. Adriana Picoral. Retrieved from <https://picoral.github.io/>
- Poplack, S., Lealess, A., & Dion, N. (2013). The evolving grammar of the French subjunctive. *Probus*, 25(1), 139–195. <https://doi.org/10.1515/probus-2013-0005>
- Poplack, S., Torres Cacoullós, R., Dion, N., Berlinck, R. A., Digesto, S., Lacasse, D., & Steuck, J. (2018). Variation and grammaticalization in Romance: A cross-linguistic study of the subjunctive. In W. Ayres-Bennett and J. Carruthers (Eds.), *Manuals in Linguistics: Romance Sociolinguistics* (pp. 213- 246). de Gruyter.

- Schwenter, S. A., & Hoff, M. (2020). Cross-dialectal productivity of the Spanish subjunctive in nominal clause complements. In S. Sessarego, J. J. Colomina-Almiñana, and A. Rodríguez-Riccelli (Eds.), *Variation and Evolution: Aspects of Language Contact and Contrast Across the Spanish-speaking World* (pp. 11–31). John Benjamins Publishing Company.
- Silva-Corvalán, C. (1994). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change*, 6, 255–272.
- Torres Cacoullous, R., LaCasse, D., Johns, M., & De la Rosa Yacomelo, J. (2017). El subjuntivo: hacia la rutinización. *Moenia* 23, 73–94.

---

## Explaining the variance in US Spanish variable patterns: A comparison of different operationalizations of bilingual experiences

Isabella Calafate<sup>1</sup>, Adriana Picoral<sup>2</sup>

<sup>1</sup>Baylor University, Waco, USA. <sup>2</sup>University of Arizona, Tucson, USA

The relation between contact and grammatical change has more often than not been seen as unequivocal, leading to premature assumptions regarding the effect of bilingualism on contact varieties (Poplack & Levey, 2010). US Spanish, for instance, is frequently seen as unstable and not as complete as monolingual Spanish (Bessett & Carvalho, 2022). In research on mood variation in US Spanish, the operationalization of contact with English through immigrant generations has gained popularity since the pioneering work of Silva-Corvalán (1994) in Los Angeles. Based on differences in subjunctive use across generations, research claims for an accelerated simplification of subjunctive due to contact with English in Los Angeles (Silva-Corvalán, 1994), Miami (Lynch, 1999), among other areas. However, contact with English can be operationalized in different ways and not without challenges. A unidirectional operationalization based on birthplace such as immigrant generation may in fact be an oversimplification of bilingual experiences given the linguistic realities of the Southwest US, which result from the constant influx of Spanish-speaking immigrants, the proximity to Mexico, and the unique position this region has historically and culturally (Villa & Rivera-Mills, 2009).

This variationist study examines how contact with English affects variable patterns of US Spanish regarding mood variation. Data were collected from sociolinguistic interviews from the Corpus of Spanish in Southern Arizona (Carvalho, 2012-). In a previous stage of this research, three extralinguistic factors—immigrant generation, language dominance, and language use in the family—were analyzed. Since factors examined in a multivariate logistic regression should not be mutually dependent, separate analyses were run in R. All three factors were found to be significant to mood variation in Southern Arizona. The variance explained (first calculated for each separate model and then through bootstrap simulations) indicated that the model with language use in the family was the one that best explained mood variation in the region ( $R^2 = 69.3\%$ ), drawing attention to the importance of language socialization for one's linguistic behavior.

Despite previous results in favor of the model with language use in the family, the relevance of immigrant generation and language dominance is undeniable given that the models with these factors were found to be significant and to also have a high variance explained ( $R^2 = 67.9\%$  and  $67.8\%$ ). To account for these factors as well, we calculated a bilingual index that combines them. Preliminary results point to the significance of this bilingual index to mood variation ( $p < .001$ ) and a variance explained of  $68.3\%$ , which is in fact not higher than the value previously found for the model with language use in the family. We present these results with a discussion of how model fit, variance explained, and resampling simulations can help determine which factor groups best explain the variable being studied. As we continue this research, more alternatives for operationalizing contact with English and bilingual experiences are explored to find a better way to assess their effect on mood variation and potentially on other variables as well.

- Bessett, R. & Carvalho, A. M. (2022). The structure of US Spanish. In S. Loza and S. M. Beaudrie (Eds.) *Heritage Language Teaching: Critical Language Awareness Perspectives for Research and Pedagogy*. (pp. 44–61). New York: Routledge.
- Carvalho, A. M. (2012-). *Corpus del Español en el Sur de Arizona (CESA)*. University of Arizona. [cesa.arizona.edu](http://cesa.arizona.edu).
- Lynch, A. E. (1999). The Subjunctive in Miami Cuban Spanish: Bilingualism, Contact, and Language Variability. 60(8), 2895–A.
- Poplack, S. & Levey, S. (2010). Contact-induced Grammatical Change: A Cautionary Tale. In Auer, P. & Schmidt, J. (Eds.) *Language and Space: An International Handbook of Linguistic Variation*. Volume 1: Theories and Methods (391–419). Berlin: Mouton de Gruyter.
- Silva-Corvalán, C. (1994). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change*, 6, 255–272.
- Villa, D. & Rivera-Mills, S. (2009). An integrated multi-generational model for language maintenance and shift: The case of Spanish in the Southwest. *Spanish in Context*, 6(1), pp. 26–42.

---

## Spanish Lexical Borrowings in Portuguese among the Venezuelan Diaspora in Brazil: A Variationist Perspective

Marina Cárcamo

Seattle University, Seattle, USA

This presentation delves into a systematic variationist analysis of Spanish lexical borrowings in L2 Portuguese by a balanced sample of 32 Venezuelan migrants residing in Brazil. Following Poplack's (2018) perspective on borrowing analysis, it investigates the linguistic and extralinguistic factors that influence the use of Spanish lexical items compared to monolingual Portuguese nouns. The linguistic factors analyzed include the semantic classification (unique or synonymic) and cognate status of the lexical items, while the extralinguistic factors comprise the participants' gender, length of residence in Brazil, age of arrival, and social integration with native Portuguese speakers.

The analysis is based on a corpus of sociolinguistic interviews that I collected in the Campinas region of Brazil in 2019, from which 2,348 noun phrase (NP) heads in subject position (around 75 tokens per speaker) were extracted for examination. The data underwent two multivariate mixed-effects logistic regression analyses using Rbrul (Johnson, 2009).

The findings indicate that lexical borrowings account for a small proportion (13.8%) of the analyzed tokens, with Portuguese NP heads being more dominant (86.2%). The statistical analyses reveal that borrowings are more frequently employed when they are unique terms compared to Portuguese NP heads. This observation implies that borrowings are associated with innovative cultural elements tied to the contact culture and are utilized to address lexical gaps in the recipient language, aligning with the expectations outlined by Aaron (2015). Additionally, the study finds that borrowings from Spanish are favored by semi-cognate words, and disfavored when the Portuguese counterparts are unrelated, suggesting that the migrants' L1 is particularly activated in contexts where the languages resemble, which leads to a higher permeability. Regarding extralinguistic factors, gender, length of residence, and social integration emerge as significant predictors of borrowing usage. Women show a higher usage of borrowings compared to men, consistent with their historical association with language innovation and change. The results also reveal that individuals with a shorter period of residence in Brazil and limited interaction with native Portuguese speakers tend to employ borrowings more frequently, indicating a relationship between lower L2 exposure and borrowing usage.

By providing insights into the linguistic and social circumstances that shape borrowing patterns during L2 acquisition in immersion contexts, this study's contributions are multiple. First, it highlights the benefits of a community-based quantitative approach afforded by variationist sociolinguistics for exploring the social factors behind the development of L2 varieties among migrants. Second, it adds to the field of SLA by considering the impact of typological similarities in the analysis of the acquisition of cognate languages. Finally, it explores lexical borrowings in a community beyond the WEIRD pool[1], addressing Bayley and Escalante's (2022) call to include L2 learners usually excluded from social science research.

[1] Western, Educated speakers from Industrialized, Rich, Democratic nations.

Aaron, J. E. (2015). Lone English-origin nouns in Spanish: The precedence of community norms. *International Journal of Bilingualism*, 18(2), 191-211.

Bayley, R. & Escalante, C. (2022). Variationist Approaches to Second Language Acquisition. In K. Geeslin (Ed.), *The Routledge Handbook of Second Language Acquisition and Sociolinguistics*. New York: Routledge.

Johnson, D. (2009). Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3 (1), pp. 359–383. Doi: 10.1111/j.1749-818x.2008.00108.x

Poplack, S. (2018). *Borrowing: Loanwords in the speech community and in the grammar*. Oxford University Press.

---

## Split ergativity and loss of rich verbal agreement

Julien Carrier

Université du Québec à Montréal, Montréal, Canada

The standard morphosyntactic alignment across all Inuit varieties is generally held to be ergative-absolutive (see Johns 1992; Compton 2017), as intransitive subjects and transitive objects both bear the (null) absolutive case whereas transitive subjects are marked with the ergative case, as shown with examples from Inuktitut in (1a-b).

arna-up angut taku-jaa

TRANSITIVE (or ERGATIVE)

woman-ERG man(ABS) see-PART.3SG/3SG

'The woman sees the man'

angut sinik-tuq  
man(ABS) sleep-PART.3SG

#### INTRANSITIVE

‘The man is sleeping’ (from Johns, 2001: 131)

However, there is another syntactic construction that can also express a dyadic event, namely the antipassive construction, as shown with another example from Inuktitut in (2).

arnaq anguti-mik taku-juq  
woman(ABS) man-MOD.SG see-PART.3SG

#### ANTIPASSIVE

‘The woman sees the/a man’ (from Johns, 2001: 131)

In early descriptions and many grammars of the language (cf. Bergsland 1955; Fortescue 1984), it is often said that speakers use the transitive construction (often called “the ergative construction” in the literature) when the object is definite but the antipassive construction when it is indefinite. However, antipassive objects in Inuktitut can also be interpreted as definite, as exemplified in (2). There is a growing consensus that the possibility of antipassive objects being interpreted as definite in fact signals an ongoing change in Inuktitut, whereby the antipassive construction is on its way to replace the ergative construction as the basic transitive construction (see Johns 1999, 2001, 2006; Yuan 2018; Carrier 2012, 2021). An important shortcoming of this proposal, though, is the lack of statistical evidence (but see Carrier 2021). The conditioning factors are thus not well understood. Crucially, if this change is truly taking place, it would indicate a morphosyntactic alignment shift from ergative-absolutive to nominative-accusative, as antipassive and intransitive subjects are both marked the absolutive case while antipassive objects are marked with an oblique case commonly called the modalis case in the Inuit literature (compare (1b) with (2)).

In the first part of this talk, using a variationist approach (e.g., Labov 1972; Tagliamonte 2012), I demonstrate that the antipassive construction is indeed being used progressively more in disfavor of the ergative construction in North Baffin Inuktitut. The corpus is based on naturalistic data from 40 native speakers (21 men and 19 women) who were all born and raised in Mittimatalik, Nunavut, between 1902 and 1998. The statistical results show that the antipassive construction started to be favored over the ergative construction with third-person dual/plural subjects and subsequently also with third-person singular subjects, while the ergative construction is still predominantly employed when subjects are first or second person. From a typological perspective, this emerging split system based on subject person represents an anomaly, since it has been widely assumed that there exists a universally fixed directionality of person splits, where the ergative-absolutive alignment surfaces with third-person subjects and the nominative-accusative one with first/second-person subjects (see, e.g., Dixon 1994; Deal 2016; Coon & Preminger 2017). Yet, we observe the opposite here.

In the second part of this talk, I argue that this split system stems from the neutralization of subject number agreement in ergative inflection, which has made many forms encoding dual/plural subjects disappear and mainly affected those encoding third-person subjects. It is often argued that loss of rich verbal agreement may lead to loss of null subjects (e.g., Roberts 1993; Duarte 2000).

Using again a variationist approach, I show that third-person null subjects appear progressively more in the antipassive construction, suggesting a failure in licensing this type of null subjects in the ergative construction. From a formal perspective, Holmberg (2010) argues that the presence of a D-feature in T/Infl for languages with rich verbal agreement is what makes them able to license null subjects. Languages undergoing a loss of rich verbal agreement lose this D-feature in T/Infl, which makes null subjects either more constrained or no longer allowed. However, this proposal cannot account for the fact that it is not “all” null subjects that can no longer be licensed in the ergative construction but only third-person ones (due to specific impoverished agreement forms). Building on Torrego and Pesetsky (2004), I propose that verbal agreement forms not only have uninterpretable  $\phi$ -features but also valued ones if they are not impoverished, allowing them to license null subjects, or unvalued ones if they are, making them unable to license null subjects.

- Bergsland, Knut. (1955). *A Grammatical Outline of the Eskimo Language of West Greenland*. University of Oslo, Norway.  
Carrier, Julien. (2012). *L’expression de la transitivité en Itivimiut*. Master’s thesis, Université du Québec à Montréal.  
Carrier, Julien. (2021). *Ergativity on the move*. Doctoral dissertation, University of Toronto.  
Coon, Jessica, and Omer Preminger. (2017). Split ergativity is not about ergativity. In J. Coon, D. Massam and L. Travis (eds.), *The Oxford handbook of ergativity*, (pp. 226–252). Oxford: Oxford University Press.  
Compton, Richard. (2017). Ergativity in Inuktitut. In J. Coon, D. Massam and L. Travis (eds.), *The Oxford handbook of ergativity*, (pp. 832–850). Oxford: Oxford University Press.  
Deal, Amy Rose. (2016). Person-based split ergativity in Nez Perce is syntactic. *Journal of Linguistics*, 52, 533–564.  
Dixon, Robert M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press

Duarte, Eugênia. (2000). Main and embedded null subjects in Brazilian Portuguese. In M. Kato and E. Negrão (eds.), *The Null Subject Parameter in Brazilian Portuguese*, pp. 127–146. Frankfurt: Vervuert-IberoAmericana.

Fortescue, Michael. (1984). *West Greenlandic*. London: Croom Helm.

- Holmberg, Anders. (2010). Null subject parameters. *Parametric variation: Null subjects in minimalist theory*, 88-124.
- Johns, Alana. (1999). The Decline of Ergativity in Labrador Inuttut. In R. M. Déchaine, L. Barel, and C. Reinholtz (eds.), *Workshop on Structure and Constituency in Native American Languages*, (pp. 73-90). MIT: MIT Occasional Papers in Linguistics.
- Johns, Alana. (2001). An inclination towards accusative. *Linguistica Atlantica*, 23, 127-144.
- Johns, Alana. (2006). Ergativity and Change in Inuktitut. In A. Johns, D. Massam and J. Ndayiragije (eds.), *Ergativity: Emerging Issues*, (pp. 293-311). Dordrecht: Springer.
- Labov, William. (1972). *Sociolinguistic Patterns*. Malden, Massachusetts: Blackwell.
- Roberts, Ian. (1993). *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht: Kluwer.
- Tagliamonte, Sali A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Wiley-Blackwell, Malden, Mass.
- Pesetsky, David, and Esther Torrego. (2007). The syntax of valuation and the interpretability of features. *Phrasal and clausal architecture*, 262-294.
- Yuan, Michelle. (2018). *Dimensions of Ergativity in Inuit: Theory and Microvariation*. Doctoral dissertation, MIT.

**Perceiving prosody: Dialect contributions to perceptions of pitch and speech rate in Miami**  
Phillip Carter<sup>1</sup>, Salvatore Callesano<sup>2</sup>

<sup>1</sup>Florida International University, Miami, USA. <sup>2</sup>University of Illinois, Urban-Champaign, Urbana, USA

In remarking on perceived differences among varieties of Spanish, speakers often invoke prosodic features as a key site of differentiation, linking perceived speech rate to judgements of intelligibility. But while prosody clearly figures in explicit judgements about Spanish dialects, the ways in which ideas about prosodic difference mediate nonconscious perceptions of Spanish language varieties is a topic that has received scarce empirical attention. This study presents the findings of a verbal-guise experiment designed to consider the relationship among Spanish dialect differences in pitch and speaking rate alongside national-origin stereotypes. As stimuli, three male residents of Miami read a text in their home variety of Spanish: Peninsular, Highland Colombian, and “post-Castro” Cuban (Alfaraz, 2002). All three were college educated in their respective countries of origin, are professionally employed in Miami, and are between the ages of 25-40. College students (n=302) - Latinx (n=210) - and non-Latinx (n=90) - were asked to listen to these recordings and asked a) How would you rate the pitch of this person’s voice? and b) How would you rate the speed of this person’s speech? Response possibilities were presented as 5-point Likert scales, with ranges of “very deep” to “very high” (pitch) and “very slow” to “very fast” (rate). To account for actual acoustic differences in the speech stream, we measured the mean pitch in PRAAT for each voice used in the instrument (Peninsular: 129.41 Hz, SD: 27.67 Hz; Cuban: 138.12 Hz, SD: 51.45 Hz; Colombian: 133.77 Hz, SD: 19.85 Hz). All voices align within 9 Hz of one another. To account for differences in speaking rate, we calculated words per minute and found minimal variation (Peninsular: 197.40 words/min; Cuban: 213.60 words/min; Colombian: 212.40 words/min). Linear mixed effects were built in R to test effects of dialect on perceptions of pitch and speech rate. ‘Participant’ was considered a random variable and fixed effects were included for ethnicity (Latinx vs. non-Latinx), national origin (e.g. Cuba, Venezuela), birthplace (U.S. vs. non-U.S.), and income. For pitch, results show a significant effect in the aggregate in which the Cuban voice is perceived as “lower pitched” than the Peninsular and Colombian guises. Latinx participants made more fine-grained discriminations, perceiving the Cuban voice with the lowest pitch, and the Peninsular voice as significantly lower than the Colombian. When looking at just those participants born in the U.S., we find significant differences among dialect but also a length of residency effect, in which longer length of residency in Miami correlates positively with higher perceptions of pitch across the board. Regarding speech rate, we find significant perceptual differences across all dialect comparisons, such that the Colombian voice is perceived to have the lowest rate, the Spanish voice in the middle, and the Cuban voice is perceived to have the fastest. This pattern is replicated for the Latinx-only analysis. U.S.-born participants perceive the Colombian voice to be the slowest. We discuss these and the full suite of significant interactions in light of stereotypes about these varieties of Spanish in Miami and Latin America.

- Alcorn, S., K. Meemann, C. Clopper, and R. Smiljanic. 2020. Acoustic cues and linguistic experience as factors in regional dialect classification. *The Journal of the Acoustical Society of America* 147: 657.
- Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change* 21.1: 135-156.
- Campbell-Kibler, K. (2010). Sociolinguistics and perception. *Language and Linguistics Compass* 4.6: 377-389.
- Clopper, C. and R. Smiljanic. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics* 39: 237-245.
- Jacewicz E, Fox RA, O'Neill C, Salmons J. Articulation rate across dialect, age, and gender. *Lang Var Change*. 2009 Jul 1;21(2):233-256.



- Levon, E. (2007). Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society* 36.4: 533-554.
- Lipski, J. M. (1994). *Latin American Spanish*. Addison-Wesley.
- Pratt, T. (2023). Affect in sociolinguistic style. *Language in Society* 52.1: 1-26.
- Smiljanic, R. and C. Clopper. 2017. Prosodic and segmental cues to regional dialect variation in American English. *Journal of the Acoustical Society of America* 142.4: 2679-2679.
- Tyler, J. C. (2015). Expanding and Mapping the Indexical Field: Rising Pitch, the Uptalk Stereotype, and Perceptual Variation. *Journal of English Linguistics*, 43(4), 284-310.

---

## Slur reclamation and metalinguistic discourse among LGBTQ speakers

Allison Casar

University of Illinois, Urbana-Champaign, Urbana, USA

The word *queer* has undergone significant semantic and pragmatic shift since the rights movements of the 1980s and 90s. Today, many consider it a reclaimed slur, as attested by the existence of Queer Studies as a discipline in major universities. This paper analyzes data from Twitter and interviews with young, self-identified members of the LGBTQ community in a college town in Central Illinois in order to explore variation in metalinguistic and metadiscursive attitudes towards reclaimed use of *queer*. Using Corpus-Assisted Discourse Analysis (Baker 2004, 2006; Ancarno 2020), I analyze sentiment and stance-taking tactics towards *queer* as a slur and/or an identity label, and the connections between discourse on *queer* and broader discourses of transphobia, homonormativity, and heteronormativity (Jaffe 2009, Bucholtz and Hall 2004, O'Brien 2015, Duggan 2002, Koyama 2000).

For the interview portion of this study, twelve self-identified LGBTQ young adults (18-35) living in the American Midwest were recruited using word-of-mouth and flyering. During semi-structured interviews they were asked about their preferred terms for their sexual orientation and gender, their preferred label for the community ("LGBT community", etc), and their experiences as an LGBTQ person. They were also asked to discuss how they felt about the reclamation of anti-LGBTQ slurs, especially the word *queer*.

For the second stage of this study, Twitter data was gathered and organized into a small corpus using the rtweet package for R (Kearney 2019). Tweets were pulled from the Twitter API using the keywords "*queer*" and "*q-slur*," the latter of which is an avoidance term sometimes employed by people who wish to discuss the word *queer* but refuse to use it, because they consider it a slur.

Both datasets were analyzed using discourse analysis informed by Queer Theory and Queer-of-Color Critique (Koyama 2000, Ferrada 2018). Interview results show overwhelmingly positive metalinguistic attitudes towards *queer*; interviewees' discourse positions *queer* as a useful umbrella term and/or as a marker of non-normativity in the face of mainstream trends towards homo-, hetero-, or cisnormativity. The Twitter corpus shows largely positive metalinguistic attitudes towards *queer*, but also shows some variation. In particular, preliminary analysis suggests that the use, censorship/avoidance, or condemnation of *queer* in LGBTQ Twitter discourse coincide with homonormative and cisnormative stances.

These results reinforce the hypothesis that those who use the avoidance term *q-slur* are also more likely to espouse normative stances on sexuality and gender. This includes the attitude that there are correct versus incorrect, or normal versus abnormal, ways of performing sexual or gender identity. The co-occurrence of *q-slur* with transgender-excluding hashtags such as #LGBwithouttheT serves as further evidence that, among many LGB Twitter users, condemnation of the word *queer* and of the concept of "*queer community*" coincides with homonormative and anti-trans attitudes.

This research is crucial because it focuses on discursive variation and ideological divides within online LGBTQ communities and updates pragmatic theories of felicitous slur reclamation to reflect heterogeneity among minoritized in-group speakers.

- Du Bois, John W., and Elise Kärkkäinen. "Taking a Stance on Emotion: Affect, Sequence, and Intersubjectivity in Dialogic Interaction." *Text & Talk*, vol. 32, no. 4, Jan. 2012. DOI.org (Crossref), <https://doi.org/10.1515/text-2012-0021>.
- Ferrada, Juan Sebastian. "Resignifications: Linguistic Resistance and Queer Expressions of Latinidad." *The Oxford Handbook of Language and Sexuality*, edited by Kira Hall and Rusty Barrett, 1st ed., Oxford University Press, 2021. DOI.org (Crossref), <https://doi.org/10.1093/oxfordhb/9780190212926.013.73>.
- Jaffe, Alexandra. *Stance: Sociolinguistic Perspectives*. Oxford University Press, 2012.
- Koyama, Emi. "Whose Feminism Is It Anyway? The Unspoken Racism of the Trans Inclusion Debate." *Sociological Review*, vol. 68, no. 4, July 2020, pp. 735-44. EBSCOhost, <https://doi.org/10.1177/0038026120934685>.
- O'Brien, Jodi. "Heterosexism and Homophobia." *International Encyclopedia of the Social & Behavioral Sciences* (Second Edition), edited by James D. Wright, Elsevier, 2015, pp. 790-95. ScienceDirect, <https://doi.org/10.1016/B978-0-08-097086-8.10204-1>.

---

## Effects of phonological length and morphological variation in code-switching processing and representation

Daniella Castillo

Graduate Center & Queens College/CUNY

This project investigates the acceptability and processing of Spanish-English code-switches in progressive versus perfect constructions by manipulating phonological and morphological aspects of either the auxiliary or the participle. Previous studies have found an asymmetry between the two structures in naturalistic corpora [1], acceptability judgment tasks [2], and processing [3], such that the code-switch in (1) is not only more frequent, but it is also rated higher and read faster than the code-switch in (2) when both are compared to a baseline.

(1) Los estudiantes están *writing their assignments*. the students are

(2) Los estudiantes han *written their assignments*. the students have

Whereas the asymmetry costs in comprehension have been attributed to distributional patterns in production, factors that are strictly linguistic (e.g., verb morphology) have not been investigated. At the phonological level, while present participles in English and Spanish show congruency in minimal length requirements (i.e., both are at least disyllabic, e.g., *giving*, *dando*), past participles present variation: English past participles can be monosyllabic (e.g., *died*), but in Spanish, they are at least two-syllables long (e.g., *ido*). Therefore, the predictive expectations after encountering a Spanish auxiliary may call for a disyllabic form. The phonological length of the auxiliary may also play a role in code-switching comprehension costs, considering that in monolingual sentences *ha*, but not *habían*, is blocked in certain syntactic structures (e.g., subject inversion) [4]. At the morphological level, a code-switch after *estar* predicts a regular verb complement, a condition that Spanish and English participles satisfy by default. In contrast, a code-switch after *haber* can be followed by either regular or irregular participles, some of which may not overlap cross-linguistically in their morphology (e.g., *won* vs. *ganado*). Taking into account that the processing of regular forms requires decomposition (i.e., separating the stem from the inflection) while irregular forms are non-compositional [5], it is possible that these two routes may modulate switch costs. In light of these observations, we ask whether the phonological length of the auxiliary, as well as the phonological length and morphology of the past participles modulate the acceptability and the magnitude of the cost in code-switching.

To test the various linguistic variables and their effects on representation and processing I will employ the eye-tracking while reading methodology and collect acceptability judgments from highly proficient Spanish-English bilingual code-switchers living in New York. Four experiments have been designed: Experiment 1 compares regular and irregular verbs (e.g., *shot* vs. *reached*); experiment 2 compares the phonological length of the auxiliary (i.e., *han* vs. *habían*); experiment 3 compares the phonological length of the past participle (e.g., *grouped* vs. *organized*); experiment 4 compares progressive and perfect structures (e.g., *están reaching* vs. *han reached*) to replicate previous findings.

- [1] Salig, L.K., Valdés Kroff, J.R., Slevc, L. R., & Novick, J.M. (2023). Linking frequency to bilingual switch costs during real-time sentence comprehension. *Bilingualism: Language and Cognition*, 1–16. <https://doi.org/10.1017/S1366728923000366>
- [2] Dussias, P. E. (2003). Spanish-English code mixing at the auxiliary phrase: Evidence from eye-movement data. *Revista Internacional de Lingüística Iberoamericana*, 2, 7-34.
- [3] Guzzardo Tamargo, R. E., Valdés Kroff, J. R., & Dussias, P. E. (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, 89, 138-161. <https://doi.org/10.1016/j.jml.2015.12.002>
- [4] Lema, J. (1995). Distinguishing copular and aspectual auxiliaries: Spanish *ser* and *estar*. *Contemporary research in Romance linguistics*, 257-274.
- [5] Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes*, 18(4), 405-430. <https://doi.org/10.1080/01690960244000054>

---

## Orientation towards the vernacular and style-shifting as language behaviours in speech of first-generation Polish migrant communities speaking Norwegian in Norway

Chloe Castle<sup>1</sup>, Kamil Malarski<sup>2</sup>

<sup>1</sup>The Arctic University of Norway (UiT), Tromsø, Norway. <sup>2</sup>Adam Mickiewicz University, Poznań, Poland

Dialect acquisition in a foreign language is too often not discussed in research on multilingual acquisition. The choice and the use of the dialect is, however, vitally important for the social meanings which language

communicates, i.e. it matters whether one uses the standard or the vernacular forms. The question is to what extent non-native speakers of a language are able to pragmatically engage in processes like style-shifting.

Indeed, it has been shown that sociolinguistic variation is acquired relatively early by children (Labov 1971, Labov 2013), together with understanding the social norms this variation communicates (Okumura et al. 2014). Moreover, some studies have shown that L2 speakers are more likely to shift or adapt their dialects than L1 speakers (Gnevshva et al. 2022).

Our study focuses on style shifts and the frequency of using the vernacular in Ln Norwegian by speakers with Polish as their first language, residing in Norway. The sociolinguistic interviews were recorded in Tromsø and Oslo, two separate dialect regions, among Polish fluent speakers of Norwegian. The analysis focuses on dialectal and accentual variation in their speech, and the question whether the high use of the dialect is influenced by socio-cultural factors like their sense of belonging within Norwegian society, or the percentage of native Norwegian speakers in their friends' circles.

The results are twofold. First, the researched speakers display dialectal features from the region (as opposed to the written standard forms of Bokmål) mostly in unscripted speech, when talking about their interests and family, as opposed to in more formal speech styles elicited through the reading tasks. This shows that an L3 speaker is capable of developing sensitivity towards the norm and the vernacular form, and understands that the two may send two different messages in different social surroundings. Moreover, it shows that first-generation migrant communities in fact may develop their L2/L3/Ln language competencies in a similar process that native speakers do, also at the level of sociolinguistic variation. It has sometimes been shown previously that their language use would often be governed by the use of idiolect structures, and only their children would orient themselves towards community standards and the regional variation in a language (Labov 2014). Second, the results returned in the linear regression statistical model show that the two factors strongly conducive to the general use of the dialect are the very high fluency of Norwegian, as well as a high percentage of Norwegian native speakers present in one's social circles. The former signals that the full pragmatic understanding of the concept of the vernacular and dialectal variation comes only when one is fully fluent in a language. The latter, in turn, may be showing that one is ready to fully assimilate with language communities in a foreign country only after understanding how sociolinguistic variation works in a foreign language; yet trying to assimilate with such communities perhaps enforces and accelerates the process of acquiring the understanding of sociolinguistic variation, too.

Gnevshva, K., Szakay, A., & Jansen, S. 2022. Lexical preference in second dialect acquisition in a second language.

*International Journal of Bilingualism*, 26(2), 163–180.

Labov, W. 1971[1964]. Stages in the acquisition of standard English. In Roger Shuy (ed.), *Social dialects and language learning*, New York: Appleton-Century-Crofts, 77–103.

Labov, W. 2013. Preface: The acquisition of sociolinguistic variation. *Linguistics* 51(2): 247–250.

Labov, W. 2014. The sociophonetic orientation of the language learner. In Silvia Calamai and Chiara Celat. (eds.) "Advances in Sociophonetics", Amsterdam/Philadelphia: John Benjamins, 17–28.

Okumura Y, Kanakogi Y, Takeuchi S, Itakura S. 2014. Twelve-month-old infants show social preferences for native-dialect speakers. *Shinrigaku kenkyu: The Japanese Journal of Psychology* 85(3): 248–56.

---

## Investigating Communication Accommodation and Endonormativity: The case of TH-stopping and postvocalic /r/ in Malaysian English

Le Xuan Chan

National University of Singapore, Singapore

Why do you change your accent when speaking to foreigners? This question straddles two important concepts in sociolinguistics: 1) accommodation towards the interlocutor, and 2) endo- and exonormative standards. Unlike its neighboring Singapore English (SgE), existing literature on Malaysian English (MalE) has yet to provide evidence for progression past nativization into the endonormative stabilization stage of Schneider's dynamic model of new Englishes [5,7]. Within the context of MalE, this study investigates the interaction between endonormative and exonormative norms in conversational speech, and to what extent convergence, divergence, or maintenance [2] play a role in indicating endonormative stabilization.

The current study looks at two MalE features, TH-stopping and postvocalic (non)rhoticity, utilizing spontaneous speech from 6 (3 male, 3 female) proficient MalE speakers elicited using spot-the-difference tasks with a MalE and non-MalE (Canadian English) interlocutor. 1097 tokens of initial /θ, ð, t, d/ and 919 tokens of postvocalic /r/ were analyzed. The presence of TH-stopping and postvocalic rhoticity were coded, alongside Voice Onset Time (VOT) measurements for stopped /θ, ð/ and /t, d/ tokens. Statistical analyses were conducted using binary logistic regression and linear-mixed effects models in R.

For TH-stopping, /θ/ and /ð/ were regularly stopped, with no significant differences between the two interlocutors ( $z=-.11$ ,  $p=.92$ ). A three-way laryngeal contrast ( $f=50.3$ ,  $p<.001$ ), however, was found: /t/ is aspirated,

stopped /θ/ is unaspirated, while /d/ and stopped /ð/ showed bimodal distributions of short-lag and negative VOT (Table 1).

For postvocalic /r/, however, we see an overall interlocutor effect ( $z=5.7$ ,  $p<.001$ ), with rhoticity increasing from the MaE task (42%) to the non-MaE task (59%). We also see an effect of environment driven mainly by /r/ before vowels ( $z=3.3$ ,  $p<.01$ ). Though significant differences were not found between other environments, we still see a trend where /r/ preceding vowels are most likely to be rhotacized (67%), followed by phrase-final /r/ (54%), /r/ preceding fricatives (46%), and /r/ preceding stops (39%), which were least likely to be rhotacized (Fig. 1).

These results show convergence towards the exonormative norm in MaE for postvocalic rhoticity, but not for TH-stopping. The findings pattern with [1] on Hong Kong English, which proposes phonological salience in explaining the difference in accommodation between /r/ and /θ/. Alternatively, it is possible that /r/-lessness is stigmatized in MaE as in SgE [8], or that the adoption of postvocalic /r/ serves a practical function to increase distinction [6]. The presence of convergence towards the exonormative norm, however, MaE might have yet to progress into the endonormative stabilization, though evidence from more phonotactic features and comparisons with SgE are needed.

This study also points to ongoing sound changes in MaE phonology – evidence for rhoticity found here concurs with previous observations in MaE [3] and SgE [8], though unlike SgE, postvocalic /r/ before vowels are not stigmatized and are regularly produced. Evidence of aspiration in /t/ also departs from reports of deaspirated voiceless stops in [4], indicating a trend of increasing aspiration in MaE.

- [1] Cao, W. (2019). Phonetic convergence of Hong Kong English: sound salience and the exemplar-based account. *Proceedings of the 19th international congress of phonetic sciences*, 2334-2338.
- [2] Dragojevic, M., Gasiorek, J., & Giles, H. (2015). Communication accommodation theory. *The international encyclopedia of interpersonal communication*, 1-21.
- [3] Pillai, S. (2015). Rhoticity in Malaysian English: The emergence of a new norm. In U. Gut, R. Fuchs, & E.-M. Wunder (Eds.), *Universal or Diverse Paths to English linguistics*, 23-40.
- [4] Phoon, H. S., Abdullah, A. C., & Maclagan, M. (2013). The consonant realizations of malay-, chinese-and indian-influenced Malaysian english. *Australian Journal of Linguistics*, 33(1), 3-30.
- [5] Schneider, E. W. (2003). Evolutionary patterns of New Englishes and the special case of Malaysian English. *Asian Englishes*, 6(2), 44-63.
- [6] Starr, R. L. (2019). Cross-dialectal awareness and use of the BATH-TRAP distinction in Singapore: Investigating the effects of overseas travel and media consumption. *Journal of English Linguistics*, 47(1), 55-88.
- [7] Tan, R. S. K., & Low, E. L. (2010). How different are the monophthongs of Malay speakers of Malaysian and Singapore English?. *English World-Wide*, 31(2), 162-189.

---

## Assessing Literacy in the Home Languages of Immigrant Students

Jennifer Chard, Anthony Vicario, Elizabeth Garza  
Multilingual Literacy SIFE Screener

This presentation describes a multilingual assessment tool used in New York State during public school intake for newcomer immigrant students. The battery of assessments includes reading comprehension and math tests in 18 languages (the most common languages of instruction for newcomers prior to arriving in New York State). Students who are new to English and who arrive in New York State public schools from other countries are given these assessments if educators believe that the student may have had an interruption or inconsistency in education prior to arriving in the United States. These tests can help school personnel understand students' abilities in reading comprehension and math, regardless of their proficiency in English. The results of these multilingual tests are provided to school personnel in English, which allows them to make better informed placement, programming, and instructional decisions to best support students as they transition to school life in the United States.

Home language testing is important for several reasons. First, the knowledge and abilities students have in the home language (for example, in mathematical skills or critical thinking while reading a passage text) can help them as they learn English. The students taking the tests are usually new to English and can't describe their academic history to teachers easily; however, by understanding some of the skills and abilities students have upon arrival, educators can build instructional plans for students which build on these skills, while teaching the English needed to express these skills in the United States. In the presentation we will discuss specific ways that home language skill transfer can assist students as they learn English.

As part of developing these tests, we analyze the text complexity of passages by grade level across a wide number of countries. Text complexity increases as grade level increases, and we will discuss generalizations from analyses of texts in Arabic, English, and Spanish.

---

## Talking sushi: cross-linguistic variations and the construction of a cosmopolitan identity in Changsha, China

Ziqi Chen

Linguistics, University of Chicago, Chicago, USA

Everyday materials have been shown to be enregistered indexicals of social meanings, e.g., wine (Silverstein, 2006) and porcelain (Gal, 2017), yielding a material register concurrently with sociolinguistic variants. How do cross-linguistic variations operate in forming a material register in the contemporary context of urban China? In this paper, I explore this question by observing the semiotics of sushi in the making of elite ways of talking Japanese food in Chinese, based on the similarity and differences between Japanese and Chinese script systems. Ethnographically, I focus on a sushi restaurant in the style of omakase (entrusting the chef with deciding a multiple-course meal, usually expensive), located in Changsha, a rapidly growing second-tier city in south-central China. By analyzing the menus created by the chef and owner, the restaurant's food reviews posted on social media, and the opinions shared by the restaurant's long-term customers/reviewers, I show 1) how Japanese lexical items are borrowed into Chinese through orthographic variations, producing the material register of sushi; 2) how the review language and the menu language manifest the style of Japaneseness differently; 3) what qualities are picked out and enregistered in appreciating a 'good' piece of sushi as connoisseur, rendering a linguistically and culturally foreign image-text (Mitchell, 2013); and 4) the language ideologies shared by people in talking sushi. Findings are discussed pertaining to the projection of cosmopolitan identity, social differentiation embedded in the city-specific context, and the third-wave sociolinguistics.

Agha, A. (1999). Register. *Journal of Linguistic Anthropology*, 9(1/2), 216-219.

Eckert, P. (2008). Variation and the indexical field 1. *Journal of sociolinguistics*, 12(4), 453-476.

Gal, S. (2017). Qualia as value and knowledge: Histories of European porcelain. *Signs and Society*, 5(S1), S128-S153.

Mitchell, W. T. (1986). *Iconology: image, text, ideology*. University of Chicago Press.

Silverstein, M. (2006). Old wine, new ethnographic lexicography. *Annual Review of Anthropology*, 35, 481-496.

Zhang, Q. (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in society*, 34(3), 431-466.

---

## Within-group talker variation and memory asymmetries

William Clapp<sup>1</sup>, Charlotte Vaughn<sup>2</sup>, Meghan Sumner<sup>1</sup>

<sup>1</sup>Stanford University, Stanford, CA, USA. <sup>2</sup>University of Maryland, College Park, MD, USA

Decades of research in sociolinguistics has documented the fundamental nature of between-group and within-group variation in language. Meanwhile, work on memory for spoken language (e.g., Palmeri, Goldinger, & Pisoni, 1993) has often focused on macrosocially homogenous speaker groups (i.e., middle-class, educated, white speakers), and has not investigated effects of individual speakers, perhaps expecting few differences in how various speakers' language is encoded and stored in memory. These omissions are especially curious given that this memory research has spurred the development of exemplar theory in speech, which has been a major influence within sociolinguistics.

Recently, Clapp, Vaughn, & Sumner (2023) found an asymmetry between memory for words produced by white talkers and identifiably Black talkers, an effect which held across Black and white listeners, suggesting a role of macrosocial variation on memory representations. In this study we used Clapp et al.'s (2023) publicly available data to pursue the idea that memory for spoken words depends not only on macrosocial patterns, but also on individual talkers within a group. Specifically, we tested the hypothesis that listeners vary in their memory for words produced by talkers within a macrosocial category, suggesting that fine-grained social cues that influence how spoken words are encoded and subsequently remembered, resulting in asymmetries at both the macrosocial and individual levels.

We analyzed data from 527 participants from Clapp et al.'s Exp. 1. Participants completed a continuous recognition memory experiment in which 140 words were presented and then repeated at varying lags (e.g. with 1–67 intervening trials). Listeners identified whether words were Old (repeated) or New (first presentation) on each trial. Words were repeated in the Same voice or a Different voice.

We analyzed hits (correct Old responses on repeated trials) and  $d'$  (sensitivity:  $z$ -transformed hit rate minus  $z$ -transformed false alarm rate, i.e., incorrect Old responses on New trials). Hits were analyzed with a mixed-effects logistic regression model including fixed effects of Talker, RepVoice (Same vs. Different), Lag (number of intervening trials), and Trial Number.  $d'$  values were modeled with a mixed-effects linear regression including the fixed effect of Talker. Estimated marginal means tests were conducted to obtain pairwise comparisons for members of each demographic category (B/W: Black/white; F/M: female/male; A/C: Alabama/California). There were two talkers in each of the eight possible groups.

Results revealed considerable variability within demographic groups. Hit rates (Fig. 1, left) varied significantly between talkers within the groups WFA, WMC ( $p < 0.001$ ), BFA ( $p < 0.01$ ), and BFC ( $p < 0.05$ ) and marginally within BMC ( $p < 0.1$ ). Differences in sensitivity (Fig. 1, right) were found between members of the categories WFA, WMC ( $p < 0.001$ ), BFA, BFC ( $p < 0.01$ ), and BMC ( $p < 0.05$ ).

The results provide evidence that encoding strength differs across talkers within the same macrosocial categories. This finding has implications for theories of speech perception, highlights the importance of using diverse samples of talkers and listeners in spoken language research, and demonstrates the rich interplay between sociolinguistic variation and the cognitive processes that underlie language understanding.

Clapp, W., Vaughn, C., & Sumner, M. (2023). The episodic encoding of talker voice attributes across diverse voices,” *Journal of Memory and Language*, 128. 104376.

Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (1993). Episodic Encoding of Voice Attributes and Recognition Memory for Spoken Words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(3):309–328.

### Variation in Differential Object Marking in Copala Triqui

Lauren Clemens<sup>1</sup>, Jamilläh Rodriguez<sup>2</sup>, Rebecca Tollan<sup>3</sup>

<sup>1</sup>University at Albany, Albany, USA. <sup>2</sup>University of North Carolina, Chapel Hill, USA. <sup>3</sup>University of Delaware, Newark, USA

**Overview** Differential Object Marking is attested in Copala Triqui, an Otomanguean language of Oaxaca, Mexico, as well as in Spanish, a common L2 for Triqui speakers. This project focuses on within-speaker variation with respect to DOM in Copala Triqui, e.g. what syntactic variables are more likely to condition the accusative marker *man* in so-called optional contexts, as well as between-speaker variation, e.g. how an individual’s unique linguistic profile in this largely bilin- gual population impacts acceptability judgments. Our results suggest that both formal syntactic and usage-based constraints condition DOM in Copala Triqui.

**Background** Copala Triqui has a rich tone system, isolating morphology, and VSO word order, but SVO is also common, especially with those who use more Spanish. Unlike the other Triqui languages, Copala Triqui has differential object marking (DOM): direct objects are only morpho- logically marked in a subset of transitive clauses, and the presence of the accusative marker is triggered by factors pertaining to animacy and specificity (e.g. Silverstein 1976; Comrie 1989). In (1a) below, the object pronoun is obligatorily preceded by the marker *man*, in (1b) the non-human animate object is optionally preceded by *man*, and in (1c) the inanimate direct object may not be preceded by *man*. The example in (1c) illustrates another hallmark of DOM languages, which is that differential object marking is often syncretic with the dative case on indirect objects (Manzini and Franco 2016; Kalin 2018). For more on the emergence of DOM in Copala Triqui and the relationship between accusative and dative case marking in the language see Broadwell (2022).

1.

- a. Niyaj xtaj maree \*(man) ‘unj a’  
 [ni<sup>3</sup>jah<sup>2</sup> ʃtah<sup>32</sup> mare:<sup>13</sup> ma<sup>3</sup> ʔu<sup>h1</sup> a<sup>ʔ2</sup> ]  
 see bird red ACC 1SG DECL  
 ‘A/the red bird sees me.’
- b. Niyaj xtaj maree (man) xcuáá gáá a’  
 [ni<sup>3</sup>jah<sup>2</sup> ʃtah<sup>32</sup> mare:<sup>13</sup> ma<sup>3</sup> ʃkua:<sup>5</sup> ga:<sup>5</sup> a<sup>ʔ2</sup> ]  
 See bird red ACC snake ? DECL  
 ‘A/the red bird sees a/the rattlesnake.’
- c. Rque’ xtaj maree (\*man) tachrii \*(man) ‘unj a’.  
 [ʃke<sup>4</sup> ʃtah<sup>32</sup> mare:<sup>13</sup> taʃsi:<sup>31</sup> ma<sup>3</sup> ʔu<sup>h1</sup> a<sup>ʔ2</sup> ]

give bird red feather DAT 1SG DECL  
 ‘A/the red bird gave a/the feather to me.’

While non-human animate objects occur both with and without *man* (1b), Broadwell (2022) reports that at least some speakers strongly prefer it when the subject and object are equal in animacy. He also reports that some speakers are also more likely to require *man* in VSO clauses, where the subject and object are in adjacent positions, than in SVO order, where they are not. Observations like these led him to conclude that one factor in the presence of DOM is differentiation of the object from the subject (see also Aissen 2003; Richards 2010; Fedzechkina et al. 2012).

Experiment In order to probe the role of individual variation, word order, the role of DOM with non-pronominal, non-human animate objects, and animacy/specificity equivalence for subject and object, we designed a 2x2 acceptability judgment experiment that we are currently administering and plan to complete in July 2023. The experiment crosses two factors (within subjects, within items): The first manipulation is the presence/absence of *man* (“DOM”; *man*, null); and the second is the relative order of subject and verb (“Word Order”; VSO, SVO). This yielded the four conditions, as in (2), which all mean ‘a/the goat nudged a/the cow’:

<b>a. Condition 1: VSO <i>man</i></b>					<b>c. Condition 3: SVO <i>man</i></b>				
Tiguúj	tana	man	scúj	a’.	Tana	tiguúj	man	scu’j	a’.
[ti <sup>5</sup> gih <sup>5</sup>	ta <sup>3</sup> na <sup>32</sup>	ma <sup>3</sup>	skuh <sup>5</sup>	a’ <sup>?</sup> ]	ta <sup>3</sup> na <sup>32</sup>	[ti <sup>5</sup> gih <sup>5</sup>	m <sup>3</sup> a <sup>3</sup>	skuh <sup>5</sup>	a’ <sup>?</sup> ]
Nudge	goat	ACC	COW	DECL	Goat	nudge	ACC	COW	DECL
<b>b. Condition 2: VSO null</b>					<b>d. Condition 4: SVO null</b>				
Tiguúj	tana	scúj	a’.		Tana	tiguúj	scu’j	a’.	
[ti <sup>5</sup> gih <sup>5</sup>	ta <sup>3</sup> na <sup>32</sup>	skuh <sup>5</sup>	a’ <sup>?</sup> ]		ta <sup>3</sup> na <sup>32</sup>	[ti <sup>5</sup> gih <sup>5</sup>	skuh <sup>5</sup>	a’ <sup>?</sup> ]	
Nudge	goat	COW	DECL		Goat	nudge	COW	DECL	

Stimuli were presented auditorily. Participants rated each sentence on a 3-point scale represented by the intuitive and familiar (☹️ = 1; not a good sentence), (😐 = 2; ok not great), or (😊 = 3; fully acceptable). We created a total of 12 item sets like in (2); each participant saw a total of 3 items per condition. We coupled these with 6 filler items (all intransitive, half with SV order; half with VS order); each participant rated 18 items in total (thus, the experiment was of relatively short duration, so as to accommodate the time constraints of our target participant population). Items were created in consultation with two native speakers of Copala Triqui and recorded by a third. In addition to rating for each trial, we recorded participants’ response time. Finally, the experimenter, who used a mix of Spanish and Copala Triqui, conducted an informal survey about language experience, use, and proficiency for both languages.

Preliminary results Data collection is ongoing; here we report results from 7 participants (total projected  $n = 32$ ; full dataset to be quantitatively analyzed using mixed-effects linear regression). The mean rating across all trials was 2.9, which is consistent with the premise that non-pronominal, non-human animate objects are acceptable with and without *man*. We then examined response time to trials rated as “3”, adopting the standard view in psycholinguistic literature that constructions which are easier to process and/or are implicitly more natural trigger faster responses of acceptability (e.g. Bader and Haussler 2010). Data collected thus far (Figure 1) shows evidence of a crossover interaction of word order and case, with presence of DOM facilitating an “acceptable” response for SVO but hindering it for VSO. At this time, there is no evidence that these response times correlate with Spanish experience, use or proficiency.

Implications One way to understand the emerging pattern is that while participants were faster to respond to VSO orders overall, the presence of *man* in SVO facilitates the interpretation of the object in clauses where the object is in the canonical subject position, i.e. immediately after the verb. In cases like these, *man* does not serve to differentiate between the subject and object *per se*, but increases the saliency of the object in a canonical subject position. This study also highlights how variability in syntactic processing can be reflected in the timecourse of acceptability judgments in linguistic fieldwork.

Aissen 2003. Differential object marking: Iconicity vs. Economy *NLLT*  
 Bader & Haussler 2010. Toward a model of grammaticality judgments *J of L*  
 Broadwell 2022. The emergence of accusative case in Copala Triqui *Linguistic Discovery*  
 Comrie 1981. Language universals and linguistic typology: Syntax and morphology  
 Fedzechkina et al. 2012. Language learners restructure their input to facilitate efficient communication *PNAS*  
 Kalin 2018. Licensing and differential object marking: the view from Neo-Aramaic  
 Manzini & Franco 2016. Goal and DOM datives *NLLT*

---

## Who Speaks Taiwanese? Language Variation, Standardization, and Creation

Paul Cockrum

The Ohio State University, Columbus, USA

What is in a name? Who determines the terminology we use? These questions are often taken for granted yet present important implications when considering the position of marginalized languages. The present paper extends this argument to examine how the act of naming a language serves to either disrupt or reinforce sociopolitical power structures, particularly when combined with other language variation. We present a case study of Taiwanese, also known as Taiwanese Southern Min, Holo, and Hokkien, in an analysis of the English, Mandarin, and Taiwanese terms to refer to the language.

Taiwanese is a Sinitic variety spoken on the island of Taiwan in East Asia. Typically, linguists classify Taiwanese as belonging to the Min branch of the broader Chinese language family (Klötter 2005:1, Wu 2011:2, Lin 2008:3). Whether or not these branches are afforded the prestige of being called a language or relegated to a dialect is subject to an individual's political views (Heylen 2021:76; Tse 2000:154). In recent years, government policy has shifted from suppression to revitalization as Taiwan projects a multicultural democratic image (Klötter 2009:113; Dupré 2017:90). However, not all have supported the government's efforts, noting that the colonial power which once prohibited the public usage of Taiwanese should not play the central role in standardizing a variant—particularly when it differs from the “indigenous” speakers (Klötter 2005:248). From this, we can see that the Taiwanese language debate has become a proxy battleground for larger issues such as national identity.

The present paper conducts a qualitative analysis of metalinguistic views stated in pedagogical materials, both formal and informal, which often teach both language and a prescribed view of history. Thematic analyses of the content of videos placed speakers on a spectrum ranging from pro-Taiwanese to pro-Pan-Chinese sentiments through a deductive approach analyzing statements on history, linguistic membership, and sociopolitical usage. These sentiments correlated with terminology choice. To give a single example, language teacher Ayo rejects the government-approved terminology of “Taiwanese Southern Min” Taiwan minnanyu (Tai-Lam, 2020). In her rationale for doing so, Ayo states that the term Southern Min is not a term that has local history—it was a creation of a colonial government (referring to the Chinese Nationalist Party) to “brainwash” a population into perceiving a kinship with other Chinese languages, reflective of the political milieu at the time. For Ayo, she refers to refer to the language as “Taiwanese” due to its local history of usage.

It is in this context that modern linguists find themselves in a dilemma—how should one refer to the very language of study? In choosing an option, one unknowingly perpetuates or disrupts historical power structures—in essence, taking a side. This question is not unique to Taiwanese, as the question of balancing one's views as a linguist and respecting the wishes of the language community is something that has plagued Creole Linguistics as well (Winford 2013:406). What is in a name? For speakers of Taiwanese, it is identity construction.

Dupré, Jean-François. 2017. *Culture Politics and Linguistic Recognition in Taiwan*. London: Routledge.

Heylen, Ann. 2021. A sociolinguistic approach to the standardization of Taiyu in the quest for a Taiwanese identity. In Chris Shei (ed.), *Taiwan: Manipulation of ideology and struggle for identity*. New York: Routledge. 75-97.

Klötter, Henning. 2005. *Written Taiwanese*. Wiesbaden: Harrassowitz Verlag.

Lin Qing Xun 林慶勳. 2001. *Taiwan Minnanyu Gailun 臺灣閩南語概論 [Survey of Taiwanese Southern Min]*. Taipei: Xinli Chubanshe Gufen Youxian Gongsi 心理出版社股份有限公司.

Sebba, Mark. 2007. *Spelling and Society*. Cambridge: Cambridge University Press.

Tâi-lâm muē-á kàu lí kóng tâigi 台南妹仔教你講台語 [Tainan Little Sister Teaches You To Speak Taiwanese]. 2020. EP0 【台語的迷思】台語為什麼不叫閩南語？學台語的重要性是啥？ | 台南妹仔教你講台語 [EP0 Taiwanese Myths. Why is Taiwanese not called Southern Min? What are the important parts of studying Taiwanese? | Tainan Little Sister Teaches You to Speak Taiwanese]. YouTube. <https://www.youtube.com/watch?v=vFjSqu8HLpI>

Tse, John Kwock-Ping. 2000. Language and a rising new identity in Taiwan. *International Journal of the Sociology of Language* 143.151-164.

Tsiok ing tâi sann siann tō pōng bí phang 足英台三聲道磅米芳. 2021. 我按怎看「Tâi-gí是毋是閩南語」[How I see “is Taiwanese Southern Min?”]. YouTube. <https://www.youtube.com/watch?v=XgnUCI6k0SA>

Winford, Donald. 2013. Social factors in contact languages. In Peter Bakker and Yaron Matras (eds.), *Contact Languages: A comprehensive guide*. Boston: De Gruyter Mouton. 363-416.

Wu Zai-ye 吳在野. 2011. *Heluo Minnanyu Zonghengtan 河洛閩南語縱橫談 [Holo Southern Min: A comprehensive discussion]*. 2nd ed. Taipei: Dongda tushu gufen youxiangongsi 東大圖書股份有限公司.



---

## The actuation of unstressed /a/-raising in Modern Hebrew

Uriel Cohen Priva<sup>1</sup>, Roey Gafter<sup>2</sup>

<sup>1</sup>Brown University, USA. <sup>2</sup>Ben-Gurion University of the Negev, Israel

Low vowels tend to raise when their duration is shorter. Such phonetic pressures lead many languages to have phonological processes at unstressed positions that categorically raise low vowels (Flemming 2004). But how do phonetic tendencies become phonological properties? We analyze novel data from Modern Hebrew (MH), in which we observe several distinct patterns of unstressed /a/-raising that shed new light on this issue. Such data challenge Ohala's (2003) understanding of sound change as a reanalysis of "free" variation. We propose that the observed patterns can be accounted for by a combined model, which extends current variationist approaches to sound change (e.g. Guy 2003), by allowing a continuous phonetic variation parameter to feed into variation in a categorical phonological variation parameter.

MH has a five vowel system (i-e-a-o-u), which can all occur in stressed and unstressed positions. We elicited data from 23 speakers (12F; 11M; age: M=43 [14–77]) who produced 22 words using a carrier phrase, which targeted vowels in stressed and unstressed contexts. Previous accounts observed that duration is a key acoustic correlate of stress, but quality differences were not considered (Bat-El, Cohen & Silber-Varod 2019). Nevertheless, our data show considerable variation in the quality of /a/ (Figure 1).

All the speakers' unstressed tokens were substantially shorter than their stressed tokens (2.6 times shorter on average). 16 out of 23 speakers also produced an average lower F1 (that is higher /a/) in unstressed tokens. However, the relation between duration and vowel height plays out in different ways (Figure 2): for ten speakers, the variation in F1 is best explained as being driven exclusively by duration rather than stress (based on model comparisons). Conversely, for four speakers, stress explains the variation better than duration. For two speakers, stress and duration produce equally good models. In other words, while some speakers simply manifested the phonetic pressure to raise short low vowels, others appeared to have generalized it to a categorical stress-unstressed distinction. Notably, the seven remaining speakers did not significantly differ in F1 between stressed and unstressed positions, demonstrating that the general tendency to raise short /a/ is speaker-dependent. Parallel patterns were observed for F2 centralization.

The phonologization of /a/-raising in MH therefore requires two distinct parameters of variation. First, the phonetically motivated relationship between duration and F1 is moderated by individual variation: not all speakers yield to the pressure to the same extent. Second, only some speakers re-interpret vowel quality as being mediated by stress rather than duration. Both parameters may spread through the community (Guy 2003), but the phonological parameter is dependent on the phonetic one: sufficiently many speakers need to produce raised unstressed /a/. While sociophonetic research on vocalic variation typically focuses only on stressed vowels, our data reveal the importance of contrasting stressed and unstressed realizations, as well as studying variation in unstressed vowels more broadly (Barajas 2021). Focusing further on stress contrasts may shed important light on the variation that underlie the ubiquitous raising and neutralization patterns in the world's language.

Barajas, Jennifer. 2021. Vocalic variation: A sociolinguistic analysis of atonic vowel raising in rural michoacán, Mexico. In *The Routledge handbook of variationist approaches to Spanish*.

Bat-El, Outi, Evan-Gary Cohen & Vered Silber-Varod. 2019. Modern Hebrew stress: Phonology and phonetics. *Brill's Journal of Afroasiatic Languages and Linguistics* 11(1).

Cohen Priva, Uriel. 2022. Using mixed effects models for vowel normalization. Presented at NWAV 50.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In *Phonetically-Based Phonology*.

Guy, Gregory R. 2003. Variationist approaches to phonological change. In *The handbook of historical linguistics*.

Ohala, John J. 2003. Phonetics and historical phonology. In *The handbook of historical linguistics*.

---

## When ze is hir gender: morphosyntactic approaches to neopronouns in English

Kirby Conrod

Swarthmore College, Swarthmore, PA, USA

**Introduction:** I propose a formal analysis of the extant variation in English neologistic pronouns – neopronouns – such as *xe*, *fae*, and *thon*. While the majority of English-speakers<sup>1</sup> do not produce neopronouns frequently, and

---

<sup>1</sup> This paper is intentionally agnostic as to whether this varies between global Englishes or particular regional Englishes; much of the sociolinguistic research has focused on either internet communities or American English speakers, the majority of whom were White.

neopronouns are rated lower than canonical pronouns in acceptability judgment tasks (Rose et al. 2023), they are also not rated as entirely ungrammatical, and neopronouns *are* used by some English speakers. While sociolinguistic research has not yet identified a detailed profile of who produces neopronouns in daily conversation, it is clear that these forms are attested, albeit by a small minority of speakers. In this paper, I discuss possible approaches in a generative framework of morphosyntax, drawing on Minimalist microparametric approaches as well as Distributed Morphology.

**Background:** So-called ‘neopronouns’ are a subcategory of third-person singular pronouns (primarily discussed in English) coined with the intent to create a suitable gender-neutral alternative to *she* and *he*. Baron (2020) gives a history of these coinages, which date back to at least the early 20th century, although he notes that none of the neopronouns he investigates (*hiser*, *thon*, i.a.) gained the same widespread uptake that singular ‘*they*’ has undergone. Other analyses of neopronouns have focused on ‘nounself’ pronouns (in which a noun is used as a pronoun, yielding an inflectional paradigm such as *pup*, *pups*, *pupself*; Miltersen 2016), and have compared them to the acceptability of *they* with a singular specific referent (Hekanaho 2020). The consensus is that no neopronouns in English have achieved mainstream adoption. In an acceptability ratings survey, Rose et al. (2023) found that *xe*, *fae*, *ey*, and *ze* were rated higher than ungrammatical controls, but lower than canonical pronouns. This was true for definite NP and proper name antecedents. Hekanaho (2020) and Rose et al. (2023) found that sociolinguistic variables influenced acceptability; primarily age, gender, and LGBTQ+ identity.

**Data:** In a new analysis of the data shared by Rose et al. (2023), I performed a k-means clustering analysis on participant ratings (n=978, excluding participants with missing data). This unsupervised machine-learning categorization approach found 2 groups of respondents; in post-hoc analysis of these groups, the participants differed primarily in how they rated neopronouns (not singular *they* or un/grammatical controls). Participants in Group 1 (n=561) primarily rated neopronouns high with proper names and definite NPs (mean=5.3), while Group 2 (n=417) primarily rated neopronouns low (mean=2.5). This suggests that speakers fall into roughly two groups: either neopronouns are mostly acceptable, or mostly not acceptable. Sociolinguistic variables (age, gender, LGBTQ+ identity) were related to group membership, but were not completely deterministic except in one respect – that all participants whose own pronouns are neopronouns fell into Group 1.

**Analysis:** The differences between ratings by Group 1 and Group 2 above suggest that acceptability of neopronouns is a morphosyntactic parameter, which begs the question of what exactly that parameter consists of; there are several logical possibilities. The first, which I present and then immediately discard, is that a neopronoun (say, *xe*) and definite specific singular *they* expone the same set of features, and which lexical item is spelled out is largely stochastic. Since neopronouns are so low frequency in daily conversation, this seems immediately implausible. A second related possibility is that *xe* is a possible exponent of some set of features for one set of speakers, but not others; if this were the case, however, it would be expected that neopronouns would be readily and easily transmitted between speakers, and that they would spread rapidly. This, too, does not seem to be empirically valid.

Instead, I propose that speakers who accept and produce neopronouns have in their lexicon additional morphosyntactic features, meaning that the featural makeup of *xe* and (singular) *they* are not identical. That is, neopronouns are not morphosyntactically genderless, but have unique, idiosyncratic gender features that differentiate them from each other. This then begs the question of whether the gender features associated with any given neopronoun are a type of novel gender features or whether they are unrelated to the features that differentiate *she* and *he*. In this, I follow Conrod (2019, 2022), who argues that so-called ‘gender’ features in English pronouns are less similar to noun-class systems of grammatical gender (e.g. Bantu, Romance, Germanic) and more similar to social honorific features like pronouns in Thai or T/V 2nd person pronouns in European languages.

Thus, I propose that the morphosyntactic features differentiating *xe* from *she* are social features, whose denotations are along the lines of (1) below (based on denotations given by Conrod 2019:78, based on the R<sub>0</sub> *is-called* function proposed by Matushansky 2015:340).

(1)  $\iota x.R_0(x)(\text{“xe”})$

≈person who is called ‘xe’

The difference, then, between speakers in Group 1 (who accept neopronouns) and Group 2 (who do not) is not a Borerian microparameter per se, but rather the existence in their individual lexicons of each feature associated with each neopronoun; and, more broadly, apparently the ability of individuals to add features to their inventory of possible genders. This also predicts that referring to a person who uses *xe/xyr* with *fae/faer* will be experienced as misgendering, since neopronouns are therefore predicted *not* to be individually interchangeable across paradigms.

**Discussion:** The analysis I propose – that neopronouns are acceptable for speakers who allow or are willing to acquire new gender features into their lexical inventory – predicts that neopronoun-permitting speakers may also be more permissive of neologisms in other functional categories, such as determiners or prepositions; this is a robust area of potential future research. Additionally, this analysis lends itself to a very direct social-semiotic link between neopronoun acceptance and gender binarist ideology: people who believe that many diverse genders are possible are, it follows, more likely to organize their mental inventory of features to reflect that fact. Future studies may test this hypothesis by gathering information on social ideological beliefs about gender (as a binary *vs* expansive and diversiform), as well as prescriptivist linguistic ideology.

Hekanaho, L. (2022). A thematic analysis of attitudes towards English nonbinary pronouns. *Journal of Language and Sexuality*, 11(2), 190-216.

Rose, E., Winig, M., Nash, J., Roepke, K., & Conrod, K. (2023). Variation in acceptability of neologistic English pronouns. *Proceedings of the Linguistic Society of America*, 8(1), 5526.

Matushansky, O. (2015). The other Francis Bacon: On non-bare proper names. *Erkenntnis*, 80(Suppl 2), 335-362.

Conrod, K. (2019). *Pronouns raising and emerging* (Doctoral dissertation).

Conrod, K. (2022). *Abolishing Gender on D*. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 67(3), 216–241. Cambridge University Press.

Miltersen, E. H. (2016). Nounself pronouns: 3rd person personal pronouns as identity expression. *Journal of Language Works - Sprogvidenskabeligt Studentertidsskrift*, 1(1).

Baron, D. (2020). *What's Your Pronoun?* (1st ed.). Liveright Publishing Corporation.

---

## “The girls in the valley concur:” stancetaking and language ideologies regarding the speech of

Julia Fox

Kathryn Copeland

University of Kentucky, Lexington, KY, USA

In this project, I analyze the metalinguistic commentary from TikTok users regarding the speech of actress and model Julia Fox, using the frameworks of stancetaking, indexicality, affect theory, and language regard (Ahmed, 2014; Dubois, 2007; Eckert, 2008; Evans et al., 2018; and Kiesling, 2009). This commentary surrounds a TikTok audio that went viral because of the way Fox pronounced the name of the 2019 film *Uncut Gems* (which she starred in) in a podcast episode, which was stylized by commenters as “unka jams” to reflect Fox’s lowering of /e/ characteristic of the California Vowel Shift (Pratt and D’Onofrio, 2017)

Though there are over 47,000 videos on TikTok using this audio and many more comments, I focus on the comments on two videos: the original video uploaded by Spotify and a video uploaded by a “Valley Girl in solidarity,” defending Fox from the negative comments (nattyissues, Spotify). In both comment sections, people fixate on Fox’s perceived “stupidity,” “arrogance,” and “fakeness,” all of which are indexed by the linguistic features called out by commenters, specifically her “vocal fry” and California-like vowels and are implicated in the stylistic bundles of the enregistered characterological figures which commenters connected her to, especially “valley girl” and “Kardashian” (Chao and Bursten, 2021; Habasque, 2021).

To better understand the evaluation of Fox’s speech, I organized the commentary in a version of the stance triangle from Dubois (2007) with more specific labels and appended an indexical field in line with Eckert (2008) to the end of the triangle that highlights the indexical associations made by commenters in their evaluations, which dialogically draw upon previous evaluations of the linguistic features present in the stance object, “unka jams.” I then further dissect the commentary specifically using the frameworks of epistemic and affective stance, and their connection to culturally constructed ideas about who is allowed to feel and know what (Jaffe, 2009; Kiesling, 2009).

The epistemic stance taken by a variety of commenters in their evaluations not only indicates their confidence in the truth value of their own statements, but the truth value of Fox’s assertions, as they continually reference Fox’s supposed “cluelessness” — that she does not know what she is talking about. Closely connected to the idea of “cluelessness” is the evaluation of Fox’s speech as “annoying,” an affective stance taken frequently in the comments, and which illustrates the way that the circulation of this object and the surrounding commentary as textual series become “saturated with affect” (Ahmed, 2014, p. 11). Overall, an examination of the stances taken and the alignment and disalignment from Fox and the cloud of qualities and characters indexed by her speech provide insight into the way that such linguistic features as creaky voice are evaluated based upon the embodiment of a speaker and broader ideologies (language-related and otherwise) about gender and intelligence. These results demonstrate the value of social media commentary as a resource for understanding language ideologies and identity negotiation through speech (Preston, 2019).

Ahmed, S. (2014). *The Cultural Politics of Emotion*. Edinburgh University Press.

Evans, Betsy E, Erica Benson, and James Stanford (eds.) 2018. *Language Regard*. Cambridge University Press.

- Chao M., and Bursten, J. R. S. (2021). Girl Talk: Understanding Negative Reactions to FemaleVocal Fry. *Hypatia* 1–18. <https://doi.org/10.1017/hyp.2020.55>
- Du Bois, J. W. (2007). The stance triangle. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 164(3), 139-182.
- Eckert, P. (2008). Variation and the indexical field. *Journal of sociolinguistics*, 12(4), 453-476.
- Habasque, P. (2021). Is Creaky Voice a Valley Girl Feature? *Stancetaking & Evolution of a Linguistic Stereotype*. *Anglophonia. French Journal of English Linguistics*, (32).
- Jaffe, A. (2009) Introduction. In Jaffe, A. (Ed.), *Stance: sociolinguistic perspectives* (pp. 1-28). Oxford University Press.
- Kiesling, S. F. (2009). Style as stance. In Jaffe, A. (Ed.), *Stance: sociolinguistic perspectives* (pp.171-194). Oxford University Press.
- nattyissues. [@nattyissues]. (2022, February 15). Part 10 | but also girly isn't even from california so idk why she said it like that LMAOO #fy #justiceforvalleygirls #juliafox [Video]. TikTok. <https://www.tiktok.com/t/ZTRwUmJho/>.
- Pratt, T., & D'Onofrio, A. (2017). Jaw setting and the California Vowel Shift in parodic performance. *Language in Society*, 46(3), 283-312.
- Preston, D. 2019. How to get respondents to reveal implicit attitudes — Talk to them. *Linguistic Vanguard*, 5(1).
- Spotify. [@spotify]. (2022, February 9). Part 8 | Here for @juliafox #kanyewest being each others' muses ❤️ #juliafox #callherdaddy #daddygang #ye #popculturetea [Video]. TikTok. <https://www.tiktok.com/t/ZTRwUSUqg/>.

---

## Phonetic variation of duration measures for palatal plosives in Australian languages

Coralie Cram, Hannah Morrison, Claire Bowern  
Yale University, New Haven, USA

Australian languages are known for their phonological uniformity (Dixon 1980; Fletcher & Butcher 2014; Round 2023). However, considerable anecdotal evidence for variation in the realization of phonological categories exists (cf. Tabain 2023; Babinski 2020,2022). Tabain (2023) suggests that at least some Australian languages have postalveolar affricates (e.g. /tɛ/) rather than true palatals (/c/, /j/). Tabain and Beare (2011) note a difference in FFT spectral peaks between Arrernte and Pitjantjatjara, and tentatively suggest the difference could be partially due to phonemic inventory (presence or absence of a contrasting laminal stop series). Here we investigate the temporal realization of palatal plosives, arguing that crosslinguistic variation in realization is inconsistent with phonemic inventory differences. This contributes to our understanding of variation exhibited in non-contrastive systems, necessary for elucidating phonological change and its precursors.

We examine 11 Australian (Pama-Nyungan and non-Pama-Nyungan) languages (cf. Babinski 2022) using public transcribed archival materials. All data come from monologic narratives, representing monitored fluent speech. Materials were automatically force-aligned with manual correction. While this is a sample of convenience based on high-quality available materials, it covers languages from multiple areas and families. Four languages have two series of stops, variably described as signaled by differences in VOT; for simplicity we refer to them as “voiced” [j] and “voiceless” [c].

The timepoints of the closure onset and consonant offset were recorded in Praat for each palatal stop, with the timepoint of the stop burst included for tokens that showed a visible burst in the spectrogram. Approximately 150 tokens were measured for each language, with three speakers per language represented where possible. Durations of the closure, burst and full gesture were extracted from Praat and plotted in RStudio using ggplot2, with significance tested through ANOVA using the stats package.

While there was little interspeaker variation within each language, languages vary extensively in the percentage of tokens which exhibit an identifiable release burst (ranging from c. 30% of tokens in Gija and Kunbarlang to over 90% in Yan-nhangu and Yidiny; Fig. 1). Ngan’gi shows fully released voiceless plosives but voiced plosives are always debuccalized; a less marked pattern is found in Golpa. For plosives with release bursts, closure and burst duration vary extensively (Fig. 2), and significantly ( $P < 0.05$ ).

This work provides further evidence that the phonetics of categories are language-specific (rather than varying by speaker or in free variation), even in systems which are phonologically identical. Australian languages vary in both the degree of release and in the duration of closure and release burst of palatal segments. This variation is not accounted for by consonant inventory, either in number of place contrasts or in number of stop series. For example, Golpa, Yan-nhangu, Gija, and Kayardild all have a laminal contrast, and the first two (but not the last) have a voicing contrast. However, this clearly does not explain the variation in these results, since Gija patterns more similarly with Warlpiri (in both figures) and MalakMalak (in figure 2) than with Kayardild.

Babinski, Sarah. 2020. Lexical stress: Phonetic variation under phonological stability in 23 Australian languages. In *LabPhon17*. online.

Babinski, Sarah. 2022. *Archival Phonetics & Prosodic Typology in Sixteen Australian Languages*. Yale University.

Dixon, R. M. W. 1980. *The languages of Australia (Cambridge Language Surveys)*. Cambridge: Cambridge University Press.

Fletcher, Janet & Andrew Butcher. 2014. Sound patterns of Australian languages. In Harold Koch & Rachel Nordlinger (eds.), *The Languages and Linguistics of Australia: A Comprehensive Guide*. Walter de Gruyter GmbH & Co KG.

- Round, Erich R. 2023. Segment inventories. In Claire Bowern (ed.), *The Oxford Guide to Australian Languages* (Oxford Guides to the World's Languages), Chapter 10. Oxford, New York: Oxford University Press.
- Tabain, Marija. 2023. Articulatory and acoustic phonetics. In Claire Bowern (ed.), *The Oxford Guide to Australian Languages* (Oxford Guides to the World's Languages), Chapter 9. Oxford, New York: Oxford University Press.
- Tabain, Marija & Richard Beare. 2011. A Spectral Analysis of Stop Bursts in Pitjantjatjara. In *ICPhS, 1934–1937*.

## The New York City Metro Area Survey: language attitudes and the low back merger

Cecelia Cutler<sup>1,2</sup>, Darlene Intlekofer<sup>1,2</sup>

<sup>1</sup>Graduate Center/City University of New York, New York, USA. <sup>2</sup>Lehman College/City University of New York, Bronx, USA

This study examines the relationship between the low-back merger and attitudes towards New York City English (NYCE). The project builds on data from the New York City Metro Area Language Survey distributed via social media in 2021-2022 yielding 716 responses. Participants represent all major demographic categories and were raised and attended school in New York City and surrounding areas that make up the NYCE dialect region (Newman 2014). The questionnaire included “same or different” homophony judgments for five LOT/THOUGHT minimal pairs as well as Likert-scale responses to statements that rely on respondents’ individual interpretations of the local accent: “I like the New York City accent”; “I would be sad if the New York City accent disappears”; and “I sound like a New Yorker.”

With regard to the sociolinguistic distribution of the merger, the results replicate the findings of Haddican et al. (2016), showing there is a relationship between the low-back merger and social factors like age and ethnicity. First, there is a strong correlation between age and the low-back merger: the rate of merger is just 5% for people over 55 but rises steadily to 39% for people under 25 ( $r = .88$ ,  $p < .00001$ ). There are also differences between ethnic groups with low rates of merger among those with a long-standing presence in the city, i.e., self-identified European Americans and African Americans (5-10%), but higher rates among groups whose presence is more recent, i.e., Latinos, West Indians, East Asians, and South Asians (23-42%) (cf. Johnson 2007).

With regard to language attitudes, there are two notable findings. Perhaps unsurprisingly, there are different rates of the merger among those who express positive feelings towards the NYC accent compared with those who express negative attitudes. Specifically, there is a lower rate of merger among people who say they sound like a New Yorker and would be sad if the NYC accent were to disappear. On the flip side, there is a higher rate of the merger among people who don't think they sound like a New Yorker and wouldn't mind if the NYC accent were to disappear. These trends help to explain real- and apparent-time shifts towards the low-back merger in NYCE. Speakers with negative attitudes towards NYCE avoid local variants (Labov et al. 2013, Tamminga 2019) and/or a “classic New Yorker” register (Becker 2016). Consequently, such speakers reverse THOUGHT-raising, a stereotype of NYCE, bringing the vowel closer to LOT and feeding the low back merger. The findings point to the importance of analyzing language attitudes in order to understand the factors that drive linguistic change.

- Becker, K. (2016). “Linking Community Coherence, Individual Coherence, and Bricolage: The Co-occurrence of (r), Raised bought and Raised bad in New York City English.” In “Coherence, Covariation, and Bricolage: Various Approaches to the Systematicity of Language Variation,” edited by Frans Hinskens and Gregory R. Guy. Special issue, *Lingua* 172–73: 87–99. <https://doi.org/10.1016/j.lingua.2015.10.017>.
- Herold, Ruth. “Mechanisms of merger: The implementation and distribution of the low back merger in eastern Pennsylvania” (1990). Dissertations available from ProQuest. AAI9026574. <https://repository.upenn.edu/dissertations/AAI9026574>
- Haddican, B., Newman, M., Johnson, D. E., & Kim, F. (2016). The diffusion of the low-back merger in New York City. *New Ways of Analyzing Variation*, 45.
- Johnson, D. E. (2007). Stability and change along a dialect boundary: The low vowels of southeastern New England. University of Pennsylvania.
- Labov, W., Rosenfelder, I., & Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, 30-65.
- Newman, M. (2014). New York City English. In *New York City English*. De Gruyter Mouton.
- Tamminga, M. (2021). “Leaders of Language Change: Macro and Micro Perspectives.” In *Language Variation—European Perspectives VIII: Selected Papers from the Tenth International Conference on Language Variation in Europe (ICLaVE 10)*, Leeuwarden, June 2019, edited by Hans Van de Velde, Nanna Haug Hilton, and Remco Knooihuizen, 269–90. Amsterdam: Benjamins.

---

## Is Y'all for All? The Sociolinguistic Perception of y'all

Cassandra Davenport  
Northwestern, Evanston, USA

Y'all and Southernness are inextricably linked [1]; however, with the increasing spread in usage outside the South, y'all's indexical value continues to refigure. Tillery et al. 2000 found that the demographic patterns of this pronoun's spread indicate that y'all "has begun to lose its association with Southernness." More recently, many younger non-Southern speakers have adopted y'all as a gender-neutral alternative to the variable you guys. Due to gender inclusivity being a common motivation for late-y'all adopting individuals, this pronominal form has gained indexical associations such as progressiveness and/or queerness [2], potentially conflicting with its original associations with Southernness and African American Language (AAL). Southern accented speech [3] and AAL [4,5] have been found to impact social perception, with these speakers stigmatized as less intelligent/educated and of lower-socioeconomic status. With this new indexical value, y'all appears to inhabit an indexical field [6] within which it shifts in value based on other cues, linguistic and nonlinguistic, available to the hearer. Through y'all, we can further understand how variable choice is situated among other factors of identity to influence social perception, as well as how speakers resolve potentially competing indexicalities of a variable.

This study aims to determine if increased use of y'all for gender inclusivity can lessen its stigmatization for all speakers or simply those whose speech is already considered standard, using a Matched Guise Task to elicit listener reactions of sets of linguistic performances that differ in accent. The recordings for the stimuli will be gathered from sociolinguistic interviews, balanced for gender, with two white Southern accented speakers, two AAL speakers, two white non-Southern speakers whose accent is regionally salient (e.g., New York), and two white aregional "anywhere" speakers [3]. Excerpts from this spontaneous speech will be digitally manipulated to create matched pairs differing only in use of y'all or you guys. I will then use a combination of group interviews and survey data, allowing for the interviews to improve the fit of the survey questions to the population under study (following the methodology of [3]). For the group interviews, non-native y'all speakers will first hear the excerpts and then will be asked general questions about the speakers. Participants will then be played the same recordings but with their matched pairs, and asked to comment on how the use of y'all vs. you guys changes their perception of the speakers. The aim of the survey is to elicit covert responses, to see how variable choice in combination with perceived identity influences the conceptualization of social meaning in perceptions of spontaneous speech. I theorize that the social contribution of y'all is highly dependent on other social information, hypothesizing that perceived regionality and/or race will have a greater impact on perceived speaker identities than variable choice. Speakers without these social biases already associated with their accent are more likely to be evaluated based on variable choice, indexing progressiveness and/or queerness with y'all usage.

- [1] Tillery et al. (2000). "The Nationalism of Southernism." In: *Journal of English Linguistics*, Vol. 28, pp. 280-294.  
[2] McCurdy, B. (2022). Y'all Means All: The Changing Indexical Value of Southernism. Talk presented at *New Ways of Analyzing Variation* 50. 13-15 October, San Jose, CA.  
[3] Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change*, 21(1), 135-156. doi:10.1017/S0954394509000052  
[4] Rahman, J. (2008). Middle-Class African Americans: Reactions and Attitudes toward African American English. *Am. Speech* 83 (2), 141-176. doi:10.1215/00031283-2008-009  
[5] Baugh, J. (2015). "Speaking while Black.," in *The Oxford Handbook of African American Language*. Editor S. Lanehart (UK: Oxford University Press).  
[6] Eckert, Penelope (2008). Variation and the indexical field. *Journal of Sociolinguistics* 12(4):453-476.

---

## Variable Enregisterment and Variable Stigma: Shifting Indexes of Locality in Galician

Anton de la Fuente<sup>1</sup>, Julia Peck<sup>2</sup>

<sup>1</sup>Stanford University, Stanford, USA. <sup>2</sup>UC Berkeley, Berkeley, USA

Galician language variation is often described through the lens of contact with Spanish or based on geography. Though these are salient factors influencing variation in the region, the process of standardization of Galician — which has been argued to draw on anti-rural ideologies (Recalde-Fernandez 2021) — has reproduced the Galician urban-rural divide. As a result, non-standard features of Galician have become associated with rurality and lack of refinement (Loureiro-Rodriguez 2008). For this project, we will focus on three non-standard features of interest: Gheada (/g/ realized as /x/), Seseo (/s/~/s̺/ distinction in place of /s/~/θ/) and Rotacismo (/r/ for /s/ in specific word-final phrase medial contexts). All three of these features are considered non-standard, though they have strong geographical associations mainly with the western dialectal block of Galician (Fernandez Rei 1990). Through a series of sociolinguistic interviews that will be conducted in the town of O Grove starting in the summer of 2023, we aim to understand how the process of standardization has affected the social meanings of these variants. O Grove

was chosen because of the rapid transformation it has undergone in the last century – from a cluster of fishing and agricultural villages to a tourist hotspot. This change brought about an increase in contact with urban Galicia and its associated language ideologies — spurring generational language change that plausibly affected the use of non-standard features. We hypothesize that the more widely stigmatized variants (Gheada and Seseo) will be used less than the less enregistered one (Rotacismo) because of the overt stigma placed on the former features. If true, this might indicate that this feature’s indexical associations can be leveraged without risking the stigma associated with other non-standard features, as Goris García 2022 has found for Galician speakers on television.

Ultimately, this project aims to understand how speakers navigate fraught indexical fields (Eckert 2008). To achieve this, a two-part design will be employed: interview and conversation. First, participants from the town of O Grove will be interviewed by the first author, who has strong connections to O Grove but is not from there and has lived abroad for an extended period of time. Second, the interviewed participants will be recorded in conversation with an acquaintance who is also from O Grove, either a family member or a friend. Both the interview and the conversation will be guided by prompts that focus on the town (e.g. “How has O Grove changed since you were a child?”) and on language use (e.g. “How you feel about how people talk in O Grove?”, “What do you think about Gheada?”). Initial data will be collected in August 2023, with additional data collection trips in the future. Using a combination of discourse analysis and acoustic analysis, we will gain a better understanding of how these features are leveraged to create social meanings in a linguistically self-conscious non-standard context.

Eckert, P. (2008), Variation and the indexical field. *Journal of Sociolinguistics*, 12: 453-476. <https://doi.org/10.1111/j.1467-9841.2008.00374.x>

Recalde Fernández, Montserrat (2022): A gheada e o seseo: entre as ideoloxías da estandarización e as ideoloxías da resistencia [Tese Doutoral]. Repositorio Institucional da USC. Universidade de Santiago de Compostela <<http://hdl.handle.net/10347/28961>> [28/07/2022]

Goris García, Xoán Carlos. 2022 "Cambio de código e acomodación lingüística ante as cámaras de televisión: a gheada eo rotacismo. Unha achega ao cambio lingüístico no galego."

Dubert García, F. (2002): “Os sociolectos galegos”, *Cadernos de lingua*, 24 (2002), páxs. 5-27.

Regueira, X. L. (1999): “Estándar oral e variación social da lingua galega”, en R. Alvarez e d. Vilavedra (eds.): *Cinguidos por unha arela común. Homenaxe a Xesús Alonso Montero*. Santiago de Compostela: Universidade de Santiago de Compostela, vol. 1, páxs. 855- 875.

---

## Perceived social congruence between sources of social information about speakers

Chantal Marie L. De Leon  
Northwestern University, Evanston, USA

Linguistic expectations about a speaker, drawn from their voice, physical appearance, and biographical information, can affect how their speech is interpreted (e.g., Strand, 1999; D’Onofrio, 2019; Vaughn, 2019). Further, the influence of this social information on how speech is understood can be mediated by the perceived social congruence between the sources that provide speaker information. For example, listeners’ expectations of an individual being an L1 English speaker based on if listeners believe the individual’s voice and perceived racialized identity are “matching” can affect how well their speech is interpreted such that listeners can have worse (e.g., Rubin, 1992) or better (e.g., McGowan, 2015) comprehension. It is less clear, however, how consistent these outcomes are, since prior work implemented different voices, social primes, tasks, and focused on different participants. Extending past work, this study used a larger stimuli set in speech-in-noise and social attribute evaluations tasks, as well as novel types of social information primes, to ask: How does the perceived suitability between different sources of social information about speakers (e.g., their voice and a visual or written prime) affect how listeners understand and evaluate their speech?

129 English-speaking participants on Prolific completed a web experiment in which they transcribed sentences mixed in noise, encountering “speakers” represented in one of four ways: by (1) voices only, either L1 American English or L2 Mandarin-accented English, (2) voices with photographs that depict racially East Asian or White individuals, (3) voices with written biographies describing an individual as born and raised in or outside of the United States, or (4) voices with written first names that were normed as associated with a racially East Asian or White identity. Participants then evaluated these “speakers” on social attributes including foreign accentedness and intelligibility.

Preliminary results show that participants were more accurate in transcribing sentences in noise with L1 English voices overall compared to L2 English voices (mean accuracy averages: 73.28%; 50.38%), following past work (e.g., Munro, 1998). With L1 English voices, participants with any social prime (photo, 77.65%; written bio, 75.07%; written first name, 72.59%) were more accurate than voice-only participants (67.83%), supporting Vaughn (2019). However, participants performed worse with L2 English voices with any social prime (photo, 48.41%; written bio, 51.51%; written first name, 49.01%) compared to voice-only participants (52.74%), unlike McGowan

(2015). These findings indicate that the written primes, in addition to photos, contribute to listeners' perceptions of who speakers are and expectations about their speech. Notably, these findings contradict past work such that the speaker information cued by the social primes (i.e., East Asian versus White racialized identities or U.S. or non-U.S. nationalities) yielded similar transcription accuracies. Further analyses are needed to disentangle which source(s) of speaker information may have contributed to these patterns, and if other aspects (e.g., participant demographics) better predict responses. Additional results across and within experimental conditions for transcription accuracy and foreign accentedness and intelligibility ratings will be presented and discussed.

- D'Onofrio, A. (2019). Complicating categories: Personae mediate racialized expectations of non-native speech. *Journal of Sociolinguistics*, 23(4), 346-366.
- McGowan, K. B. (2015). Social expectation improves speech perception in noise. *Language and Speech*, 58(4), 502-521.
- Munro, M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, 20(2), 139-154.
- Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33(4), 511-531.
- Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of language and social psychology*, 18(1), 86-100.
- Vaughn, C. R. (2019). Expectations about the source of a speaker's accent affect accent adaptation. *The Journal of the Acoustical Society of America*, 145(5), 3218-3232.

---

## Variation in homorganic consonants across Albanian dialects and its implications for the history of the language: Evidence from the Malsia Madhe dialect

Lindon Dedvukaj  
The Ohio State University, Columbus, USA

The Albanian dialects are divided between Gheg, located in the geographic north and Tosk to the south. Gheg is subdivided into the areas of Northeastern Gheg, Northwestern Gheg, and Central Gheg. Within the Northwestern Gheg dialect there is a geographically and historically isolated region known as Malsia Madhe, spanning northwestern Albania and southwestern Montenegro. This area shows substantial phonological and lexical variation from the other Gheg and Tosk dialects. Linguistic analyses of the Albanian language have predominantly dealt with the now Standard Tosk dialect in addition to some varieties of the Gheg dialect. Moreover, reconstruction of the history of Albanian has been based primarily on the Tosk dialect and the more widely known Gheg varieties. The Malsia Madhe (Malsia) dialect has remained relatively unknown and has not undergone a thorough historical or comparative analysis with the other Albanian dialects.

This paper argues that the Malsia dialect provides new insight into the history of Albanian and the relationship among its dialects. The evidence comes from the variable realization of word-initial onsets in a particular subset of verbs across the dialects. In the Standard Albanian dialect (based on Tosk since the 1960's), many verbs surface with word-initial nasal-stop clusters [mb, nd, ŋg]. In Gheg and Malsia tribal zones that border the other Gheg regions, these onsets appear as simple nasals [m, n, ŋ], while in the highlands of Malsia (in Albania and Montenegro) they are realized as stops [p, b, t, d, k, g] (Table 1, Figure 1) (see Dedvukaj and Ndoci 2023, S. Demiraj 2013, Shkurtaj 1974, 1975).

The Malsia Madhe stops are the oldest forms, representing continuations of the Proto-Albanian stops (see Dedvukaj and Gehringer 2023). The other onsets arose through various processes involving the reinterpretation of structures containing the preposition \*en- immediately preceding the verb. First, the preposition \*en 'in' was reanalyzed as a preverb en- in most of Albanian, attaching to different roots and undergoing homorganic nasal assimilation. Thus, the Standard Albanian verb mbush 'fill' is derived from the sequence en- + bushgiving mbush. The other nasal variants that arose in Gheg and Malsia tribal zones that border the Gheg regions are reductions of the clusters mb > m, nd > n. and ŋg > ŋ (Standard mbush appears in Gheg as mush). By contrast, Malsia preserved the earlier preposition en 'in', preventing prefixation to verb stems and retaining the word-initial stop (Malsia bush). Parallels to the Malsia stops also appear in other ancient Indo-European cognates (see Table 2).

The variable onsets in Albanian dialects are important for understanding the development of the Albanian language. The continuation of the Proto-Albanian stops in the isolated Malsia Madhe dialect emphasizes the necessity for reexamining the languages history.

Dedvukaj, Lindon & Patrick, Gehringer. 2023. Morphological and phonological origins of Albanian nasals and its parallels with other laws. *Proceedings of the Linguistics Society of America (PLSA)* 8(1) 5508. <https://doi.org/10.3765/plsa.v8i1.5508>

Dedvukaj, Lindon & Rexhina, Ndoci. 2023. Linguistic variation within the Northwestern Gheg Albanian dialect. *Proceedings of the Linguistic Society of America (PLSA)* 8(1) 5501. <https://doi.org/10.3765/plsa.v8i1.5501>

Demiraj, Shaban. 2013. *Gjuha Shqipe dhe Historia E Saj*. (2nd Edition). Tirana, AL: Onufri.

Shkurtaj, Gjovalin. 1974. *E folmja e Hotit*. Tirana, AL: Dialektologjia Shqiptare II.



---

## Acting different(-ly): Bringing derivational morphology into variationist linguistics

Dan DeGenaro

Georgetown University, Washington, D.C., USA. University of Massachusetts Amherst, Amherst, MA, USA

There is anecdotal evidence that African American English (AAE) does not require -ly affixation on de-adjectival adverbs (e.g. *different* → *differently*) in certain contexts. For instance, Schneider & Kortmann (2004) claim that AAE prefers to omit -ly from adverbs in general. This suggests that adverbial -ly may have a zero variant in AAE. In this study, I show that this is the case, and that the zero and non-zero variants of -ly adverbs differ significantly in their respective distributions.

Using the Corpus of Regional African American Language (Kendall & Farrington 2021) as a source of data, I present a statistical analysis of adverbial -ly as it is used in AAE, focusing on defining the contexts in which it may be obligatory, variable, or ungrammatical. I investigated the contexts shown in (1).

I define -ly to be a linguistic variable in AAE, following Lavandera (1978). As AAE is known to make use of zero words and morphological objects (Green 2002), I further define a similar framework for the -ly variable and name the Overt Adverbial Morpheme (OAM) and Zero Adverbial Morpheme (ZAM) variants.

I claim that adverbs exhibit individual preferences for the OAM or ZAM forms, ranging from 100% of instances being OAM to 100% of instances being ZAM for certain adverbs. I hypothesize that this may be due to the prosodic structure of those adverbs, with most ZAM adverbs being stress-initial, though I acknowledge that other factors are likely relevant. Furthermore, I find that preverbal contexts strongly disprefer ZAM adverbs, while other contexts appear to allow for some degree of variation, as is shown in Figure 1. The reason(s) for this remain unclear and are thus a key direction for future studies. Finally, I compare a human-annotated corpus of examples drawn from CORAAL with the automatic labels produced by the ARK Tagger (Gimpel et al. 2010; Owoputi et al. 2013), a part-of-speech tagger previously used for studies involving AAE. I find that the ARK Tagger tends to erroneously label ZAM adverbs as adjectives.

This study is the first detailed analysis of OAM and ZAM adverbs in AAE (in fact, in any variety of English) to the best of the author's knowledge, and thus contributes new information to the body of variation research on this linguistic variety.

Gimpel, Kevin, Nathan Schneider, Brendan O'Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan & Noah A. Smith. 2010. Part-of-speech tagging for twitter: Annotation, features, and experiments. Tech. rep. Carnegie-Mellon Univ Pittsburgh Pa School of Computer Science.

Green, Lisa J. 2002. *African American English: A Linguistic Introduction*. Cambridge University Press.

Kendall, Tyler & Charlie Farrington. 2021. *The Corpus of Regional African American Language*. Version 2021.07. <https://oraal.uoregon.edu/coraal> [<https://oraal.uoregon.edu/coraal>].

Lavandera, Beatriz R. 1978. Where Does the Sociolinguistic Variable Stop? *Language in Society* 7(2). 171–182. <https://www.jstor.org/stable/4166996>. Publisher: Cambridge University Press.

Owoputi, Olutobi, Brendan O'Connor, Chris Dyer, Kevin Gimpel, Nathan Schneider & Noah A. Smith. 2013. Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 380–390. Atlanta, Georgia: Association for Computational Linguistics. <https://aclanthology.org/N13-1039> [<https://aclanthology.org/N13-1039>].

Schneider, E.W. & B. Kortmann. 2004. *A Handbook of Varieties of English: Morphology and syntax A Handbook of Varieties of English: A Multimedia Reference Tool*. Walter de Gruyter GmbH & Company KG. <https://books.google.com/books?id=allIAAAAMAAJ>.

---

## Comparison of the Grammaticalization of genre in Hexagonal French and Swiss French

Delin Deng

University of Florida, Gainesville, USA

Different scholars have documented the emerging discursive use of genre in native French (see, for example, Denison 2002; Mihatsch 2010; Secova 2011). Examples (1) and (2) illustrate the discursive and non-discursive use of genre in native speech.

(1) *y a plein de commissions genre euh théâtre vivant spectacle danse machin* (ESLO 2\_WC29)

'there is plenty of commissions like euh theater living show dance stuff'

(2) *mais c'est un autre genre de vie euh* (ESLO 1\_BU77)

'but there is another type of life euh'

As shown by these two examples, when used as a discourse marker (DM), a lexical item no longer belongs to the original category assigned to it. The original use of genre is a noun, as in example (2), while the DM genre in example (1) functions more like an adverb. The difference between the discursive and non-discursive use is that the removal of the former does not affect the semantic integrity of the sentence, while removing the latter does. This change from non-discursive to discursive use of a lexical item often involves the process of grammaticalization.

Over recent decades, much research has been dedicated to describing the discursive functions of genre. Not much has been known regarding its use in a social context. Only marginal studies discussed this aspect by pointing out that this might be a feature of youth speech (Secova 2011). However, we still do not know, for example, if the DM genre is fully grammaticalized in French or if it is still a change in progress. What is the supporting evidence for its grammaticalization? Is it an independent change in Hexagonal and Swiss French? What are the constraining social factors influencing this discursive use? etc.

Thus, by conducting a logistic mixed-effects regression analysis of data drawn from OFROM corpus (Avanzi et al. 2012-2020) collected in Francophone Switzerland in 2012 and ESLO 2 corpus (Baude and Dugua 2011) collected in France in 2008, this article examined the variable use of the French DM genre in the speech of 306 native speakers of Swiss French and 78 native speakers of Hexagonal French. Overall, our results supplied evidence of the grammaticalization of the genre in French. On the one hand, it supplied comparable results on the grammaticalization of the same particle in two French varieties and thus provided further evidence for the independence of the grammaticalization of a particle in different varieties of a language; on the other hand, our data also shed new light on the language variation and change of discourse markers in French and thus provided new insights into the development of the same particle in different regions being conditioned by different social factors. While the ongoing change of genre is led by female speakers in Swiss French, it was first led by male speakers, then quickly spread to both gender groups in Hexagonal French. The age-grading effect on this particle was observable in Swiss French, while it was absent from Hexagonal French.

Avanzi, M., Marie-José, B., Gilles, C., Federica, D., & Anne, J. L. (2012). 2020. Corpus OFROM–Corpus oral de français de Suisse romande. Université de Neuchâtel.

Baude, O., & Dugua, C. (2011). (Re) faire le corpus d’Orléans quarante ans après: quoi de neuf, linguiste?. *Corpus*, (10), 99-118.

Denison, D. (2002). History of the sort of construction family, paper presented at “ICCG2: Second International Conference on Construction Grammar”.

Mihatsch, W. (2010). The diachrony of rounders and adaptors: Approximation and unidirectional change. *New approaches to hedging*, 9(1), 93.

Secova, M. (2011). Discourse-pragmatic features of spoken French: analysis and pedagogical implications (Doctoral dissertation, Queen Mary University of London).

---

## Nasality variation in Greater Manchester: A multi-dimensional approach to quantifying sociophonetic nasality

Maya Dewhurst

Lancaster University, Lancaster, United Kingdom

Variation in voice quality is a socially meaningful phenomenon and plays an important role in acts of identity construction and group membership (Podesva & Callier, 2015). “Nasal voice”, specifically, has is a salient feature of speech and indexes a range of social characteristics (Bucholtz, 2010; Podesva et al., 2013). Previous sociophonetic studies tend to focus only on one correlate of nasalisation, such as A1-P0 (Tamminga & Zellou, 2015; Walker, 2020), but speakers frequently vary in the acoustic implementation of nasality, which motivates a multidimensional approach (Styler, 2017). Such an approach is now possible due to recent phonetic research on robust acoustic approaches to measuring multi-dimensional correlates of nasality (Carignan, 2021; Liu et al., 2019). The study presented here adapts the multi-correlate method used in Carignan (2021) to observe patterns of nasality variation in speech data from Greater Manchester, where anecdotal evidence suggests that nasality is a salient feature of local speech.

This study is part of a wider project that aims to create a robust method for estimating the extent of sociophonetic nasalisation using acoustic data. Eight speakers from Greater Manchester aged 15-18 produced a monosyllabic word list containing canonically oral and nasalised vowels, as well as three reading passages with increased concentrations of nasalised vowels. Recording was carried out using a custom-built nasometry device which separately records oral and nasal amplitudes (Baken & Orlikoff, 2010). Around 740 tokens per speaker were recorded, totalling 5,920 tokens for analysis. Speakers were stratified by sex assigned at birth and socioeconomic status, using factors including the Income Deprivation Affecting Children Index (IDACI; Office for National Statistics, 2019), percentage of pupils eligible for free school meals in their schools as a proxy for the school’s social class, and the highest household earner’s occupation category (employing the categories determined by Baranowski & Turton (2018) to accurately represent social class in British dialects).

Following Carignan (2021), proportional nasal amplitude (nasalance) was calculated as the ratio of oral to nasal amplitude (Fletcher et al., 1974) and was used as a proxy for the extent of velum lowering. Principle component analysis (PCA) was carried out on 22 acoustic correlates of nasalisation (including A1-P0 and A1-P1) and 13 MFCCs (Carignan, 2021; Styler, 2017), extracted at the midpoint of each vowel. The resulting PC values were then used to predict numerically coded rates of nasalance in a linear regression, resulting in a nasalisation “score” for each token, ranging from zero to one.

Results suggest that female speakers differ more greatly in their canonically oral and nasal productions than male speakers. Social class also appears to play an important role. These differences could begin to explain existing attitudes to nasality in Greater Manchester; a greater difference in nasalisation scores of oral and nasal tokens may be what leads to the perception of a more nasal voice. This will be discussed further, and future work will address this idea through perceptual studies and acoustic analysis of sociolinguistic interview data.

- Baken, R. J., & Orlikoff, R. F. (2010). *Clinical Measurement of Speech and Voice* (2nd ed.). Delmar Cengage Learning.
- Baranowski, M., & Turton, D. (2018). Locating speakers in the socioeconomic hierarchy: Towards the optimal indicators of social class. *New Ways of Analysing Variation* 47, New York University, New York City.
- Bucholtz, M. (2010). *White Kids: Language, Race, and Styles of Youth Identity*. Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511975776>
- Carignan, C. (2021). A practical method of estimating the time-varying degree of vowel nasalization from acoustic features. *J. Acoust. Soc. Am.*, 149, 911–922. <https://doi.org/10.1121/10.0002925>
- Fletcher, S. G., Sooudi, I., & Frost, S. D. (1974). Quantitative and graphic analysis of prosthetic treatment for “nasalance” in speech. *The Journal of Prosthetic Dentistry*, 32(3), 284–291. [https://doi.org/10.1016/0022-3913\(74\)90032-8](https://doi.org/10.1016/0022-3913(74)90032-8)
- Liu, L., Hulden, M., & Scarborough, R. (2019). RNN Classification of English Vowels: Nasalized or Not. *Proceedings of the Society for Computation in Linguistics*, 2(1), 318–321. <https://doi.org/10.7275/cnem-ab98>
- Office for National Statistics. (2019). *English Indices of Multiple Deprivation*. <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019>
- Podesva, R., Hilton, K., Moon, K., & Szakay, A. (2013). Nasality as non-aggressiveness: An articulatory sociophonetic study [Conference presentation]. NWAV 42, Pittsburgh, Pennsylvania, United States.  
<https://stanford.edu/~podesva/documents/nwav2013-nasality-1up.pdf>
- Podesva, R., & Callier, P. (2015). Voice Quality and Identity. *Annual Review of Applied Linguistics*, 35, 173–194.  
<https://doi.org/10.1017/S0267190514000270>
- Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America*, 142(4), 2469–2482. <https://doi.org/10.1121/1.5008854>
- Tamminga, M., & Zellou, G. (2015). Cross-dialectal differences in nasal coarticulation in American English. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS XVIII, Glasgow)*. 18th International Congress of Phonetic Sciences, Glasgow.
- Walker, A. (2020). Voiced stops in the command performance of Southern US English. *The Journal of the Acoustical Society of America*, 147(1), 606–615. <https://doi.org/10.1121/10.0000552>

---

## Heritage Speakers create synchronic variation and further diachronic change: the case of Spanish *ser/estar*

Karina Di Franco, Maria Piñango, Catalina Mourgues  
Yale University, New Haven, USA

The contexts that license Spanish *ser* vs. *estar* (English *be*) vary significantly across Spanish dialects. For example, in Venezuelan and Mexican varieties, the use of *estar* in contextual-standard contexts (contexts where the adjective’s meaning is not absolute but relative to a specific situation, as in “The shower-curtain *está* short (for this specific bathtub)”) is highly acceptable. By contrast, in Iberian and Argentinian varieties this same use of *estar*, while acceptable, has significantly lower ratings, regardless of whether a supportive context is provided (Fig.1). The Mexican/Venezuelan variation appears, not as stand-alone events, but as *innovations* along a directional process of *change*: *Estar* first entered the language (from Latin *stare*) in composition with locative predicates “The book *está/is* on the table” and, from there, has continued to expand its uses into contexts originally reserved for *ser* (e.g., *Ella es cansada* (12th<sub>century</sub>)→*Ella está cansada* (15th<sub>century</sub>) ‘She is tired’). This ongoing process of change which underpins the attested dialectal variation has been shown to be rooted in monolingual speakers’ sensitivity to a semantic difference: *Estar* demands the construal of a false alternative to the proposition.<sup>1</sup> That is, *estar* in “She *está/is* big” conveys that there is an alternative circumstance where she is not big; *ser* (“She *es* big”) is neutral in this respect. This semantic requirement of a false alternative renders *estar* more informative. This greater informativity pushes its increased use, allowing *estar* to encroach on the domain of *ser*.<sup>2,3</sup>

HSs are reported linguistic innovators due to their bilingual and socio-bicultural situations and thus a potential source of variability in their heritage language.<sup>4,5,6</sup> We hypothesize that **heritage speakers (HSs) harness *estar*’s greater informativity by allowing it in more contexts, in a predictable fashion, following**

**estar's encroachment.** We test our hypothesis by comparing copula use of Mexican-variety HSs (N=33, 27 female; ages 18-29; from US Mexican-Spanish speaking households) with monolingual counterparts.

**Study\_1.** Using a forced-choice task, we tested copula choice/production as modulated by either an *alternative*-supporting context that leads to the construal of a false alternative or an *alternative*-neutral context that affords no explicit possibility of an alternative. **Prediction\_1.** Like monolinguals, HSs will show sensitivity to *estar*'s informativity by favoring *estar* use in *alternative*-supporting contexts over *alternative*-neutral contexts. **Results\_1** shown in Fig.2 bear our prediction out. **Study\_2.** Using naturalness ratings, we tested acceptability/comprehension of *estar* with the same conditions as study\_1. **Prediction\_2.** Relative to monolinguals, HSs, as innovators, will show an even greater increase in acceptability of *estar* in *alternative*-neutral contexts, the contexts their monolingual counterparts currently disprefer. **Results\_2** in Fig.3 show that while HSs accepted *estar* sentences, they did so regardless of context, indicating that HSs are able to construe *estar*'s alternatives in a wider variety of contexts and are therefore better prepared to construe the alternative in the absence of explicit supporting context. **Conclusions: Heritage speakers show sensitivity to monolinguals' ser/estar meaning distinction and are introducing further variation by accelerating estar's ongoing, pre-established, diachronic path.**

<sup>1</sup> Deo, A., Sanchez-Alonso, S., & Pinango, M. (2016). Alternative circumstances of evaluation and the ser/estar distinction in Spanish. *Manuscript, Yale University*. URL <https://ling.auf.net/lingbuzz/003543>.

<sup>2</sup> Sánchez-Alonso, S. (2018). The cognitive sources of language change and variation: Connecting synchronic variation and diachrony in Spanish copula use (Doctoral dissertation, Yale University).

<sup>3</sup> Sánchez-Alonso, S., Piñango, M., & Deo, A. (2019). Variability in ser/estar use across five Spanish dialects: An experimental investigation.

<sup>4</sup> Silva-Corvalán, C. (1986). Bilingualism and language change: The extension of *estar* in Los Angeles Spanish. *Language*, 62(3), 587-608.

<sup>5</sup> Pascual y Cabo, D. (2013). Knowledge of *gustar*-like verbs in Spanish heritage speakers. In *12th Generative Approaches to Second Language Acquisition* (pp. 162-169). Cascadilla Proceedings Project.

<sup>6</sup> Kupisch, T., & Polinsky, M. (2022). Language history on fast forward: Innovations in heritage languages and diachronic change. *Bilingualism: Language and cognition*, 25(1), 1-12.

---

## Local social meanings and language use in a Chicago bilingual community

Jennifer Dibern

Northwestern University, Evanston, USA

Bilinguals dynamically adapt their language use in response to a variety of social cues (e.g., Zhang et al., 2013). These cues are believed to influence bilinguals' expectations for which language(s) are needed or appropriate in an interaction (e.g., Mann and De Bruin, 2022). Seminal theories of bilingual language control postulate that socially-cued expectations shape language use by modulating the activation levels of a bilingual's languages; they do not determine, however, what social information is relevant to these language control mechanisms (e.g., Grosjean, 2001). Prior work has historically focused on measuring aspects of language experience, like age of acquisition or proficiency, or relied on questionnaires that assess the impact of predetermined factors (e.g., Gullifer and Titone, 2020). Such measures do not capture localized social meanings despite work demonstrating the influence of local norms and identities on language use (e.g., Eckert, 2008). To understand bilinguals' language use more fully, we must identify what local social information is relevant to bilinguals themselves and how it is incorporated into processes of speech perception and control.

In this work, I present an analysis of data from semi-structured interviews with ten Spanish-English bilingual speakers who identified as Mexican American and lived in Chicago's Pilsen neighborhood. Using thematic analysis, I ascertained the factors that participants described as related to their language use. Specifically, I considered how bilinguals within this community ideologize their language use and what localized social meanings they believe shape their expectations for which language(s) are needed in an interaction.

Four main themes were identified in participants' discourse: 1) language selection is a fluid and highly contextualized process; 2) locality conditions language use; 3) Spanish is expected within Pilsen; and 4) Spanish helps to define and construct community. Theme 1 includes descriptions of cues to language choice as heterogeneous and context dependent. Theme 2 captures discourse emphasizing local, geographic place as an important component of the context that participants consider when assessing which language(s) to use. Theme 3 reflects commentary positioning Spanish language use throughout Pilsen's public spaces as unremarkable, safe, and expected. Theme 4 includes accounts of Spanish as a language of authentication and resistance, affirming and maintaining community boundaries in the face of ongoing gentrification of the neighborhood. Together, these findings underscore the significance of local, community-level context to

bilingual language use and, therefore, its importance to understanding how social information influences bilingual language selection and control. Furthermore, participants' discourse highlights the ways in which conventionally-used assessment categories like "social setting" fail to capture the complexity of bilingual language use. For example, even within Pilsen, participants describe different coffee shops as carrying different expectations for what languages they would use or feel were appropriate. Which friend(s) they are talking with and where they are specifically when having the conversation interact to influence the language(s) they use. Therefore, although a category like "social setting" encompasses both where and with whom a bilingual is speaking, it erases the intersectionality, fluidity, and locality of their language use.

- Eckert, P. (2008). Variation and the indexical field 1. *Journal of sociolinguistics*, 12(4), 453-476.
- Grosjean, F. (2001). The bilingual's language modes. *One mind, two languages: Bilingual language processing*, 122..
- Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, 23(2), 283-294.
- Mann, A., & de Bruin, A. (2022). Bilingual language use is context dependent: using the Language and Social Background Questionnaire to assess language experiences and test-retest reliability. *International Journal of Bilingual Education and Bilingualism*, 25(8), 2886-2901.
- Zhang, S., Morris, M. W., Cheng, C. Y., & Yap, A. J. (2013). Heritage-culture images disrupt immigrants' second-language processing through triggering first-language interference. *Proceedings of the National Academy of Sciences*, 110(28), 11272-11277.

---

**Ethnic boundaries and speaker identity: TH-stopping in the Montreal Anglo-Italian community**  
Maryn Diorio  
University of Ottawa, Ottawa, Canada

While ethnicity as a social factor has been studied extensively in sociolinguistic research (Becker, 2014; Boberg, 2014; Feinstein, 1980; Fought, 2006; Labov, 1964; Poplack & Tagliamonte, 1999; Rickford, 1985), few studies have focused on how ethnicity may manifest in variable conditioning. Rather, analyses have tended to focus on rates of variant use (through the inclusion of a social factor group), with potential differences in the conditioning of variability across ethnic lines not being as well-known.

Moreover, ethnicity has typically been quantified using metrics such as family background (Boberg, 2004; Kim & Wong, 2020; Labov, 1964) or religion (Feinstein, 1980; Meechan, 1999), which does not consider the gradation in ethnic identity held by speakers. This study addresses these lacunae through the study of a feature considered to be an ethnic marker of Italian-American speech: the variable realization of TH. In order to compare TH-realization across ethnic lines, nearly 3000 tokens of TH were extracted from 2 cohorts of Montreal speakers: 17 anglo-Italians and 10 British-origin benchmark speakers. Tokens were also coded for linguistic factors invoked in previous literature as being influential to TH-stopping. Furthermore, using previous literature from various fields (Ashmore et al., 2004; Heller, 1987; Hoffman & Walker, 2010; Liebkind, 2010; Nagy et al., 2014; Noels et al., 2014; Padilla & Borsato, 2010), ethnic identity was quantified, and each speaker from the former cohort was assigned a score representing their level of orientation towards the Montreal Italian community. Results reveal 1) that TH-stopping is far less frequent than expected, and 2) rate of TH-stopping does correlate with level of ethnic orientation in the expected direction (with speakers of higher levels of ethnic orientation displaying greater rates than their lesser-oriented counterparts) On the surface, this confirms that variability can be used by speakers to index ethnic identity. However, all speakers, regardless of level of ethnic orientation, share variable conditioning, which turns out to be heavily lexically constrained - speakers are using TH-stopping in restricted contexts involving highly frequent words, demonstrating that the variant is not deeply embedded in speaker grammar; rather, it is overlain in their speech using lexical means.

- Ashmore, R. D., Deaux, K., & McLaughlin-Volpe, T. (2004). An organizing framework for collective identity: Articulation and significance of multidimensionality. *Psychological Bulletin*, 130(1), 80.
- Becker, K. (2014). Linguistic repertoire and ethnic identity in New York City. *Language & Communication*, 35, 43-54.
- Boberg, C. (2004). Ethnic patterns in the phonetics of Montreal English. *Journal of Sociolinguistics*, 8(4), 538-568.
- Boberg, C. (2014). Ethnic divergence in Montreal English. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 59(1), 55-82.
- Feinstein, M. H. (1980). *Ethnicity and topicalization in New York City English*.
- Fought, C. (2006). *Language and ethnicity*. Cambridge University Press.

- Heller, M. (1987). The role of language in the formation of ethnic identity. *Children's Ethnic Socialization*, 180–200.
- Hoffman, M. F., & Walker, J. A. (2010). Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, 22(1), 37–67.
- Kim, J. Y., & Wong, N. (2020). (Divergent) Participation in the California Vowel Shift by Korean Americans in Southern California. *Languages*, 5(4), 53.
- Labov, W. (1964). *The Social Stratification of English in New York City* [Ph.D., Columbia University].  
<https://www.proquest.com/docview/302149177/citation/48FE7B845D474F80PQ/1>
- Liebkind, K. (2010). *Social Psychology*. In *Handbook of language & ethnic identity* (Vol. 1). Oxford University Press, USA.
- Meechan, M. E. (1999). *The Mormon drawl: Religious ethnicity and phonological variation in southern Alberta*. University of Ottawa (Canada).
- Nagy, N., Chocie, J., & Hoffman, M. F. (2014). Analyzing Ethnic Orientation in the quantitative sociolinguistic paradigm. *Language & Communication*, 35, 9–26.
- Noels, K. A., Kil, H., & Fang, Y. (2014). Ethnolinguistic orientation and language variation: Measuring and archiving ethnolinguistic vitality, attitudes, and identity. *Language and Linguistics Compass*, 8(11), 618–628.
- Padilla, A. M., & Borsato, G. N. (2010). Disciplinary and Methodological Approaches: Psychology. *Language & Ethnic Identity*, 1, 5–16.
- Poplack, S., & Tagliamonte, S. (1999). The grammaticization of going to in (African American) English. *Language Variation and Change*, 11(3), 315–342.
- Rickford, J. R. (1985). Ethnicity as a sociolinguistic boundary. *American Speech*, 60(2), 99–125

---

## Complementary measures in the assessment of change: A case study of French polar interrogatives

Nathalie Dion

University of Ottawa, Ottawa, Canada

This paper explores how five complementary features of variable systems – overall rates, conditioning, productivity, contextual dispersion and diffusion in the community – must be marshaled to provide a comprehensive characterization of change in progress. We illustrate by re-revisiting a robustly variable sector of Canadian French morphosyntax whose variants (below) are known to be in flux: the polar interrogative domain. Analyses extend the timeline of Elsig (2009)/Poplack and Elsig's (2006) diachronic analysis by an additional 25 years, bringing 2,000+ questions produced by 133 speakers to bear on developments occurring over a period of nearly a century and a half of spontaneous Québec French speech.

As-tu[P-INV] rencontré des beaux hommes là-bas? (21C.175.803)

'Did you meet good-looking men over there?'

Est-ce-que[ECQ] c'est l'année passée qu'on est allés en Europe? (21C.170.300)

'[ECQ] it's last year that we went to Europe?'

Je peux-tu[-TU] me prendre une assiette de spaghetti? (21C.110.1718)

'I can[-TU] take a plate of spaghetti for myself?'

Tu es[INT] rendue hippie comme ta mère? (21C.172.769)

'You have become a hippie like your mother?'

Rates show a continuing decline of P-INV, down from 37% to only 3% in the most recent dataset, at the expense of not only progressively-rising -TU (20%-33%-43%), but also – unexpectedly – INT, which was previously trending downward (43%-34%-48%). However, overall rates emerge as the most misleading measure of change: once dispersion is considered, they are revealed to have exaggerated the rise of -TU and drop of P-INV, and distorted INT's trajectory; they also under-represent the diffusion of variants in individual grammars.

Conditioning proved a much less volatile indicator of change. Much of it (i.e. the polarity constraint, the barring of -TU from 2nd singular questions and P-INV from 1st and 3rd person contexts) was stable, despite dramatic shifts in rate and the diverse demographic profiles of the speakers. Nonetheless, conditioning enabled the identification of emerging, strengthening, weakening and neutralized constraints – key stages in the transition from one grammar to another. The productivity measure helped characterize these changes in terms of growth or decay for the variant(s) involved.

Dispersion analyses pinpointed the rate changes that derived from imbalances in the proportion of contexts across datasets, and played a key role in evaluating each variant's breadth within a given period and over time.

Diffusion measures clarified that some rate changes (e.g. -TU's increase between periods 1 and 2, and P-INV's decline between periods 2 and 3) came from horizontal gains or losses, while others (e.g. -TU's rise between periods 2 and 3, and each of INT's increases) are the consequence of shifts within individual grammars. They also provided a granular view of how -TU expanded into a new context (2nd singular subjects).

Results underscore the need to consider more than rates and conditioning in the study of language change. Linguistic dispersion and diffusion in the community provide key insights into the mechanics of the transition period and shifts in variant productivity at different points in time.

Elsig, Martin. 2009. *Grammatical Variation Across Space and Time: The French Interrogative System*. John Benjamins Publishing.

Elsig, Martin. & Poplack, Shana. 2006. Transplanted dialects and language change: Question formation in Québec. *University of Pennsylvania Working Papers in Linguistics* 12(2): 8.

---

## Lexical frequency and idiosyncrasy in morphosyntactic variation - evidence from Spanish subject pronouns

Danielle Dionne, Daniel Erker  
Boston University, Boston, USA

Despite abundant evidence that the frequency properties of linguistic forms shape variation, (Jurafsky et al., 2001; Bybee, 2010; Purse et al., 2022), several questions remain open in this area, including: (1) What types of frequency measures best reflect the sensitivities of language users?, (2) To what extent are frequency effects similar across levels of linguistic structure?, and (3) What is the relationship between frequency measures and other constraints on variation? The current study explores these questions in the context of variable subject pronoun use in Spanish (e.g., *yo creo* vs. *creo* 'I think'), examining the presence/absence of pronominal forms with 88,000 verbs drawn from sociolinguistic interviews with 221 speakers.

While frequency measures are often formulated in terms of the overall occurrence of forms (i.e., as counts of lexical items or lemmas in a corpus), there is evidence that the contexts in which forms occur also constitute important frequency considerations, especially insofar as specific contexts are known to (dis-)favor particular variants of variables (Bybee, 2010; Brown & Shin, 2022). We therefore examine both the overall and contextual frequencies of individual verb forms, with the latter measures defined by how often a given verb occurs in contexts of (i) referential (dis)continuity and (ii) preceding pronominal presence/absence. Changes in referent from one verb to the next and also prior pronoun use (as illustrated in Example 1a) are known to favor subsequent pronoun use (Carvalho et al., 2015, i.a.).

Multivariate analyses of the data indicate that both overall and contextual frequencies play a role in shaping patterns of pronoun use (alongside several other linguistic factors). Verbs that occur more frequently in pronoun-favoring contexts have higher rates of pronoun use, as predicted. However, the effect of contextual frequency is itself modulated by overall frequency, only emerging among highly frequent forms. This finding aligns with previous claims that lexical frequency is not monotonically associated with pronoun use, but rather activates/amplifies the effects of other conditioning factors (Erker & Guy, 2012). Relatedly, and more generally, results indicate that increased frequency contributes to verb-specific diversification in pronoun rates. The most frequent verbs are the most pronominally idiosyncratic, ranging widely in comparison to infrequent forms (see Figure 1b, which plots pronoun rates in relation to (log-transformed) overall frequency). The widely divergent pronominal tendencies of high-frequency verbs contrast sharply with frequency-biased patterns of phonetic variation, in which more frequent forms are often similar in showing the reducing effects of highly-practiced articulatory/motor routines (Myers & Guy, 1997; Bybee, 2001; Labov, 2006; Bybee, 2010).

Taken together, these findings present a complex picture of the role of frequency in morphosyntactic variation. Higher overall frequency not only promotes increased variability in rates of pronoun use, it also serves to activate the effects of other constraints, including those of our contextual frequency measures. These results provide substantial evidence that language users are sensitive to an array of frequency considerations and raise the possibility that morphosyntactic and phonological variables are differentially tuned by the shaping forces of repetition, experience, and practice.

Brown, E. L. & Shin, N. (2022). Acquisition of cumulative conditioning effects on words: Spanish-speaking children's [subject pronoun + verb] usage. *First Language*, 42(3), 361–382.

Bybee, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, J. (2010). *Language, usage, and cognition*. Cambridge University Press.

Carvalho, A., Orozco, R., & Shin, N. (2015). Subject pronoun expression in Spanish: A cross-dialectal perspective. Georgetown University Press.

Erker, D. & Guy, G. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: Benjamins.

Labov, W. (2006). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics*, 34(4), 500–515.

- Myers, J. & Guy, G. R. (1997). Frequency effects in Variable Lexical Phonology. *University of Pennsylvania Working Papers in Linguistics*, 4(1).
- Purse, R., Fruehwald, J., & Tamminga, M. (2022). Frequency and morphological complexity in variation. *Glossa: a journal of general linguistics*, 7(1).

---

## Neutral Tone in Beijing: Reexamining the Utilization of the Neutral Tone in Beijing Professionals

Xiao Dong, Fengming Liu, Monica Nesbitt, Chien-Jer Charles Lin  
Indiana University, Bloomington, USA

One prominent feature of the linguistic landscape of Beijing is the wide use of neutral tones. Previous studies have found that males in Beijing use more neutral tones than females, casual speech utilizes more neutral tones, while Beijing yuppies, representing the middle and upper echelons in foreign businesses in Zhang (2005), use more full tones than state professionals to express their cosmopolitan identity (Zhang, 2005; Zhao, 2018). However, the interplay between this variation and place orientation remains understudied. Currently, Beijing's residents hold diverse perceptions of their city. As China's capital, Beijing offers abundant resources through its household registration system (Beijing hukou). Acquiring a Beijing hukou is challenging and fosters a sense of superiority among locals. However, contemporary Beijing residents face workplace discrimination due to biased perceptions of laziness and low motivation (Haweiniesita, 2021). In addition, the differences between foreign and state professionals in income and lifestyle are diminishing. In light of these changes, our study seeks to further explore the neutral tone variation in Beijing middle-class residents, specifically in relation to place orientation, to deepen our understanding of language change and the complex relationship between language, place, and identity (Labov, 1972; Carmichael, 2017; Reed, 2016).

In this study, we examined 22 middle-aged individuals from both state and private businesses (10 females; aged 25-45), comparable to Zhang's (2005) yuppies regarding class and interactions with non-locals. The sociolinguistic interviews included a word reading section with a list of 48 optional neutral-tone words and 44 forbidden neutral-tone words, with balanced tonal combinations. Participants also responded to 31 questions regarding their life experiences, work experiences, and orientation toward Beijing. Two native Mandarin speakers coded the neutral tone productions. Linear mixed-effects modeling of the word reading data indicated a significant effect of place orientation ( $p=0.017$ ). Participants with a higher orientation toward Beijing utilized more neutral tones. However, no significant effect of gender ( $p=0.384$ ) or business type ( $p=0.785$ ) was observed. Males and females exhibited similar rates of neutral tone usage (around 35%), and there was no significant difference between individuals working in private and state businesses.

Compared to Zhang's (2005) findings of neutral tone usage among Beijing foreign professionals (34.5% for females and 68% for males) and state professionals (100% for both males and females), our study reveals a lower neutral tone rate, especially for males. No significant difference between individuals employed in private and state businesses was found. These findings suggest that this linguistic variation has extended beyond specific professional domains and linguistic markets, becoming more prevalent among Beijing middle-class residents, particularly males, potentially influenced by the rise of cosmopolitan and yuppie identities across businesses. Nevertheless, individuals strongly oriented toward Beijing consistently exhibit a higher frequency of neutral tone usage, emphasizing how individuals express their sense of belonging through distinctive linguistic traits and underscoring the impact of place affiliation on language use. It also demonstrates a balanced state between cosmopolitan and local identities.

- Becker, K. (2009). /R/ and the construction of place identity on New York City's Lower East Side 1. *Journal of Sociolinguistics*, 13(5), 634–658. <https://doi.org/10.1111/j.1467-9841.2009.00426.x>
- Carmichael, K. (2017). Place Orientation Displacement and local linguistic practices: R-lessness in post-Katrina Greater New Orleans. *Journal of Sociolinguistics*, 21(5), 696–719. <https://doi.org/10.1111/josl.12253>
- Haweiniesita (2021, August 12). Are Beijingers being discriminated against when working in Beijing? Zhihu. <https://www.zhihu.com/question/478779497>
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia, Pennsylvania: University of Pennsylvania Press.
- Reed, P. E. (2016). *Sounding Appalachian: /ai/ monophthongization, rising pitch accents, and rootedness* (Doctoral dissertation, University of South Carolina).
- Zhang, Q. (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, 34(03). <https://doi.org/10.1017/S0047404505050153>
- Zhao, H. (2018). *Language variation and social identity in Beijing* (Doctoral dissertation, Queen Mary University of London)



---

Variation in case marking across Nubri valley  
Cathryn Donohue  
The Hong Kong Polytechnic University, Hong Kong

Nubri is a Tibeto-Burman language spoken by ~2000 people in remote Nubri Valley, upper Gorkha, northern-central Nepal. Nubri has typical Tibeto-Burman case-markers, but the system of usage is quite atypical, with the only frequently occurring case being that of a ‘dative’ case to mark only certain objects, known as ‘differential object marking’ (DOM) (Donohue 2018, 2022).

Differential object marking (DOM) is the phenomenon whereby a subset of ‘prominent’ object arguments is morphologically distinguished based on (typically) inherent properties such as definiteness, specificity, animacy, or information structural considerations such as topicality or focus. There are two types of DOM: The first, and most commonly reported, involves inherent properties of the object argument determining the case marking outcome (e.g. de Swart 2007; de Hoop & Malchukov 2008; Malchukov & de Swart 2009). However, there are a few languages in which it is the simultaneous evaluation of properties of both subject and object arguments that determines the case marking outcome, a phenomenon that has been referred to as ‘global’ distinguishability or relative scenario splits (Haspelmath 2018). This paper reports on differential object marking in Nubri which has this latter kind of DOM in which object marking is determined only through simultaneous evaluation of properties of animacy of the clausemate arguments as seen in (1) (Donohue 2018).

- (1) a. Nga mo tung yin.  
1.SG 3.SG.F see AUX  
‘I saw her.’  
b. Kho mo-la tung so.  
dog 3.SG.F-DAT see AUX  
‘The dog saw her.’  
c. \* Kho mo tung so.  
dog 3.SG.F see AUX  
‘The dog saw her.’

(NB. The differences in the auxiliary verbs are largely to do with Nubri’s evidentiality so I have simplified the glosses here.)

The sentences in (1) illustrate this phenomenon from Sama dialect of Nubri. In (1a), ‘I saw her’, the object of the sentence *mo* bears no case, while in (1b), ‘The dog saw her’, we see that ‘her’ *mo* is marked by the dative case *-la*. This is because the arguments in (1a) follow expected assignments of higher animate to subject, lower animate to object. In (1b) we see that the opposite is true: there is a ‘markedness reversal’ and the object must now bear case.

Table 1 illustrates the case outcomes observed with the verb ‘see’ in Nubri. Interestingly the Prok dialect from the other end of the valley has a similar, but not identical, pattern of DOM, as illustrated in Table 2.

This paper is based on primary field data. I introduce the phenomena of case variation in the DOM system between Nubri dialects as well as in the major contact languages for the region: Tibetan, Nepali, Ghale/Gurung and discuss the possible role of contact in the observed dialectal variation.

Donohue, C. 2018. Case marking in Nubri. *Nepalese Linguistics* 33: 28-33.

Donohue, C. 2022. Changing identity and linguistic practices in Nubri: Veiled language endangerment in the Nepalese Tibetosphere. In Gerald Roche & Gwendolyn Hyslop (eds), *Bordering Tibetan languages: Making and marking languages in Transnational High Asia*. Amsterdam: Amsterdam University Press, pp. 157-173.

Haspelmath, M. 2018. Role-reference associations and the explanation of argument coding splits. <http://ingbuzz/004047>

de Hoop, H. 2009. Case in Optimality Theory. In A. Malchukov and A. Spencer (eds) *Case*. Oxford: OUP, pp. 88-101.

de Hoop, H. & A. Malchukov. 2008. Case-marking strategies. *Linguistic inquiry* 39: 565-587.

Malchukov, A. & P. de Swart. 2009. Differential case marking and actancy variations. In A. Malchukov and A. Spencer (eds) *Case*. Oxford: OUP, pp. 339-355.

de Swart, P. 2007. Cross-linguistic variation in object marking. *LOT international series*, Vol 168.

---

Visible articulatory variation as a cue to sound change: Lip rounding and lip protrusion  
variability in the Mandarin sibilant merger

Baichen Du<sup>1,2</sup>, Alexandra Pfiffner<sup>1</sup>, Keith Johnson<sup>1</sup>

<sup>1</sup>University of California, Berkeley, Berkeley, CA, USA. <sup>2</sup>University of Hong Kong, Hong Kong

Mandarin traditionally has a sibilant contrast between retroflexes /ʂ, ʐ, ʐʰ/, alveolars /s, ts, tsʰ/, and palatoalveolars /ç, tç, tçʰ/. However, loss of retroflexion (neutralization to the alveolar PoA) has been reported in Taiwan Mandarin (Chang, 2012; Chiu, 2009; Chiu et al., 2020; Jeng, 2006) and in some Mandarin speakers with origins in mainland China (Hauser, 2023). Merged retroflexes are characterized acoustically by an increased spectral

center of gravity (CoG) and articulatorily by a shortened front resonating cavity. A previous ultrasound study on Taiwan Mandarin found more variability in tongue postures, which could lead to a less robust retroflex-alveolar contrast and neutralization over time (Chiu et al., 2020). However, there are two points worth noting; first, this merger is unidirectional in that it is almost exclusively found from retroflexes to alveolars. Second, a shortened front cavity can also be achieved by less protruded and rounded lips. The current study examines visual variation in the beginning stages of the Mandarin sibilant merger in mainland China. We aim to 1) document the visible articulatory cues of sibilants, 2) assess whether visual properties of retroflexes are also neutralizing with alveolars, and 3) hypothesize whether visible articulation and any variation can contribute to sound change.

We conducted an audiovisual production experiment with 30 native speakers. Subjects were seated in a sound-attenuated booth in front of a teleprompter with a camera recording 60 fps and a microphone recording at a sampling rate of 44.1 kHz. Stimuli were 45 monosyllabic Tone 1 Mandarin words with sibilant onsets and 15 fillers, each randomly repeated 3 times. Participants were told to produce each stimulus upon seeing it appear on the screen, which occurred after a numeric countdown and metronomic beep to serve as a reference point of articulatory onset. Audio and video recordings were aligned using Adobe Premiere Pro, and audio recordings were segmented and annotated in Praat (Boersma & Weenink, 2023). All videos were further segmented, and facial landmarks were extracted using OpenFace, a tool for facial behavior analysis (Baltrušaitis et al., 2018).

We focused our analysis on the vertical and horizontal apertures between the two outer sides of the lips and lip corners. Results show that retroflexes are significantly different from alveolars and palatoalveolars, with larger vertical apertures (Figure 1), smaller horizontal apertures, and more consistent radii, correlating to lip rounding and protrusion. This suggests that both acoustic and visual cues are used in this phonological contrast. However, we found robust interspeaker variation in aperture and roundness (Figure 2); some individuals are entirely merged in terms of visual cues. Future work will include CoG to assess levels of acoustic neutralization.

We hypothesize that this variability, together with variable tongue gestures, makes cues to retroflexes (but not alveolars) less stable, constituting the unidirectionality in the sound change. Essentially, listeners who previously attended to cues in both dimensions no longer have consistent access to them.

Baltrušaitis, T., Zadeh, A., Yao Chong Lim, Y. C., & Morency, L.-P. (2018). OpenFace 2.0: Facial Behavior Analysis Toolkit.

IEEE International Conference on Automatic Face and Gesture Recognition (FG 2018), 59-66.

Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer (Version 6.3.10). [Computer program].

<http://www.praat.org>

Chang, Y.-H. (2012). Variability in cross-dialectal production and perception of contrasting phonemes: The case of the alveolar-retroflex contrast in Beijing and Taiwan Mandarin. PhD Thesis, University of Illinois at Urbana-Champaign.

Chiu, C. (2009). Acoustic and auditory comparisons of Polish and Taiwanese Mandarin sibilants. *Canadian Acoustics*, 37(3), 142–143.

Chiu, C., Wei, P.-C., Noguchi, M., & Yamane, N. (2020). Sibilant Fricative Merging in Taiwan Mandarin: An Investigation of Tongue Postures using Ultrasound Imaging. *Language and Speech*, 63(4), 877–897.

<https://doi.org/10.1177/0023830919896386>

Hauser, I. (2023). Differential cue weighting in Mandarin sibilant production. *Language and Speech*. doi: 10.1177/00238309231152495

Jeng, J.-Y. (2006). The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwan Mandarin. *Journal of Humanistic Studies*, 40(1), 27–48.

---

## Metaphor and metonymy as markers of register variation

Markus Egg

Humboldt-Universität zu Berlin, Berlin, Germany

We investigate how metaphors and metonymies are used as register markers, presenting results from a new annotated German corpus. Linguistic registers are intra-individual variation caused by the situational and functional context of communication (Halliday and Hasan, 1985). Metaphors refer to entities similar to the referents of their literal interpretations (Ritchie 2013), metonymies refer to entities linked to literal referents through a salient functional relation (e.g., creator for creation, Littlemore 2015).

Previous work correlated metaphor and metonymy functions with register features (Goatly 1994, 2011, Deignan et al. 2013), identifying e.g. hierarchy between interlocutors as a decisive factor. Other work investigated metaphors in specific text types, e.g., Steen et al. (2010) and Beger (2015) analyse news, conversation, fiction, and academic discourse. They conclude that register influences the function of metaphor, e.g., informational registers predominantly use metaphor to express content. We approach the issue from the viewpoint of register, investigating how metaphor and metonymy are used as markers of register variation.

Like the VU Amsterdam Metaphor Corpus (Steen et al. 2010), our corpus (of eventually 180,000 words) includes different text types (parliament speeches, newspaper commentaries, sermons, light fiction, debates of a debating society, and TEDx talks), which represent important dimensions of register variation like literality/orality,

persuasivity, or social relations between the interlocutors. Following Steen et al. (2010), our annotation distinguishes conventionalised and non-conventionalised metaphors. We also annotated metonymies, extended metaphors (several metaphors sharing a specific metaphorical similarity), and a new kind of metaphor called potential metaphor, the deliberate combination of tokens of the same expression first with literal and then with metaphorical senses in the same text.

To be able to perform register marking, metaphors must be free choices, whose optional use can be harnessed for register marking. Metaphors whose use is necessitated by the language system cannot be employed for register purposes, e.g., the highly conventionalised spatial metaphors for temporal constellations like the fact that one time span is located inside another or follows it. Consequently, we focus on the first kind of metaphor, which includes non-conventional, extended, and potential metaphors.

Non-conventionalised and extended metaphors appear significantly more often in highly persuasive text types. Despite their oral nature, debates and TEDx talks are just as metaphorical as the literal registers, unlike the conversations investigated by Steen et al. (2010). We link this result to the different degree of informative purpose for debates and talks as opposed to conversations. Metonymy is related to hierarchy, occurring more frequently in text types whose interlocutors are on an equal level, like in Deignan et al. (2013). Also, metonymy correlates with brevity requirements – regardless of literality or orality –, in our corpus, for commentaries and debates.

As for individual text types, we found a mixed pattern for fiction, like in Reijnierse et al. (2019): It has few conventionalised metaphors but considerable numbers of non-conventionalised and extended metaphor. Sermons exhibit the highest marking: They abound with non-conventional metaphors, furthermore, extended and potential metaphors emerge as clear markers for sermons.

Beger, Anke (2015), Metaphors in psychology genres. Counseling vs. academic lectures, in B. Herrmann and T. Berber Sardinha (eds), *Metaphor in specialist discourse*, Amsterdam: Benjamins, 53–75.

Deignan, Alice, Jeannette Littlemore and Elena Semino (2013), *Figurative language, genre and register*, Cambridge: Cambridge University Press.

Halliday, Michael and Hasan, Ruqaiya (1985), *Language, context and text*, Victoria: Deakin University Press.

Goatly, Andrew (1994), Register and the redemption of relevance theory. *The case of metaphor*, *Pragmatics* 4, 139–181.

Goatly, Andrew (2011), *The language of metaphors*. 2nd ed. Routledge.

Littlemore, Jeannette (2015), *Metonymy*, Cambridge: Cambridge University Press.

Reijnierse, Gudrun, Burgers, Christian, Krennmayr, Tina, and Steen, Gerard (2019), Metaphor in communication: the distribution of potentially deliberate metaphor across register and word class, *Corpora* 14, 301–326.

David Ritchie (2013), *Metaphor*, Cambridge: Cambridge University Press.

Steen, Gerard, Dorst, Aletta, Herrmann, Berenike, Kaal, Anna, Krennmayr, Tina, and Pasma, Trijntje (2010), *A method for linguistic metaphor identification: from MIP to MIPVU*, Amsterdam: Benjamins.

---

## Exploring the Effects of Cross-Cultural Variation and Tourism on the Perception of Utah English

Zoe Eldredge<sup>1,2</sup>, Joseph Stanley<sup>1</sup>

<sup>1</sup>Brigham Young University, Provo, Utah, USA. <sup>2</sup>Salt Lake Community College, Salt Lake City, Utah, USA

This study describes the effects of tourism and cross-cultural variation in English as observed in speakers from Park City, Utah. Though Utah is frequently called Mormonland or the Mormon State (Bowie and Morkel, 2006), residents of Park City, or Parkites, live in an area with high tourism and relatively few practicing members of the Church of Jesus Christ of Latter-day Saints. The majority of residents are not originally from Park City; nevertheless, we have anecdotally noted homogeneity in the speech patterns of born-and-raised locals and move-in locals. While Utah English is not necessarily the same as Latter-day Saint (Mormon) English (cf. Baker & Bowie 2010), we expect Parkites' language to be significantly different from Utahns' from Latter-day Saint-dominant communities. Specifically, we hypothesize that Parkites have some linguistic features typically found in the Eastern United States.

In our first study, we compared production data from speakers from several cultural backgrounds and from areas with varying levels of tourism by conducting brief sociolinguistic interviews with twelve Parkites (six women and six men) from three religious groups (Latter-day Saint, Jewish, and other). They also read a wordlist targeting Utah features (e.g. t-insertion in words like fal[t]se, cord-card merger; feel-fill merger; Bowie & Morkel 2006), Latter-day Saint features (e.g. pre-lateral back vowel mergers, trap raising; Baker & Bowie 2010), Northeast features (e.g. cord-nurse merger, lack of pre-rhotic vowel mergers; Labov, Ash, & Boberg 2006), and Jewish features (e.g. trap-lot merger, quick cadence; Benheim 2020). Another eleven non-Parkite Utahns only read the wordlist. Results from this data suggest Parkites have fewer Utah and Mormon features than general Utahns: they had less feel-fill merger, lot-thought merger, t-insertion, and /t/-glottalization, and more goose-fronting and trap-raising compared to other Utahns.

In our second study, we analyze perceptions of Parkites' English via survey, distributed to 580 people, including Parkites and other Americans. Listeners heard audio from the interviews and evaluated the speech scales of friendliness, professionalism, Utahness, and other attributes. A chi-squared test indicates Parkites are more likely to be perceived as sounding friendly and professional ( $p < 0.05$ ). Speakers from Utah were generally perceived fairly positively. Supporting the acoustic data, Parkites' speech was evaluated as sounding more Midwestern and less 'Utahn' and 'Mormon' than general Utahns. Latter-day Saints were more likely to be perceived as being from the West and/or Utah than all other religious backgrounds.

Taken together, this production and perception data suggests Parkites may be indexing less Utahness and Latter-day Saintness and using features that align more with the tourism industry. This may indicate those wishing to index hospitality or congeniality may avoid the more local features so as to not alienate others and move towards something perceived as 'neutral' or pan-regional. As Parkites experience contact with a large variety of dialects yet seemingly gravitate towards Midwestern features, we propose it is not higher contact that creates these productions but purposeful feature usage. Thus, we conclude that residency in areas with cross-cultural contact via tourism may have a greater impact on speech patterns than previously identified.

Bowie, D., & Morkel, W. (2006). Desert Dialect (Utah). *American Voices: How Dialects Differ from Coast to Coast*, 144-148.

---

## Cognition and the human speech community: large-scale evidence of crosslinguistic similarities in variable subject pronoun expression

Daniel Erker<sup>1</sup>, Aria Adli<sup>2</sup>, Robert Bayley<sup>3</sup>, Gregory Guy<sup>4</sup>

<sup>1</sup>Boston University, Boston, USA. <sup>2</sup>University of Cologne, Cologne, Germany. <sup>3</sup>UC Davis, Davis, USA. <sup>4</sup>New York University, New York, USA

Scholars of linguistic variation are accustomed, indeed drawn, to heterogeneity. We anticipate and appreciate that the linguistic behavior of individuals and groups will be diverse in character. We also recognize that uncovering patterns within linguistic diversity – chiefly through demonstrations of the differential treatment of linguistic variables by language users – justifies our most fundamental empirical claims, including the identification of linguistic change, the delineation of speech communities, and the characterization of distinct sociolects, ethnolects, and dialects. What is less clear for the variationist is how to make sense of prevailing homogeneity, especially among individuals with limited to no linguistic interaction.

Labov (2010:3) grapples with this issue as it relates to the “extraordinary uniformity of the Northern Cities Shift”, asking: “How can we explain the fact that [these vowel changes] affect so many millions of people in widely separated cities who have no connection with each other?” His answer to this question amounts to a proposal of a third category of factors, Cognitive and Cultural, that, together with the Internal and Social factors respectively occupying Volumes 1 and 2 of his *Principles* series, constitute the purported shaping forces of language change. This third set of factors is argued to relate to “the capacity of the linguistic system to transmit information” and the human “ability to decode what is being said through the accurate identification of linguistic categories” (Labov 2010:4).

The present paper applies this perspective to the variable expression of subject pronouns, exploring the possibility that universal properties of human cognition, in combination with the information content of subject-pronominal forms, can help explain the “extraordinary uniformity” that we observe in a data set of over 200,000 finite verbs produced by 800+ individuals – each representing one of five languages (German, Mandarin, Persian, Portuguese, or Spanish) – in the context of sociolinguistic interviews.

Results of a mixed effects regression model indicate that the strongest predictor of overt (as opposed to null) pronoun use is whether the subject of a target verb constitutes a switch in referent from a previous verb. This constraint, Referential Continuity, is more robustly predictive than even the language of the speaker, a remarkable fact given that overall rates of pronoun use differ significantly across these five languages, i.e. the average rate of pronoun use among Persian speakers is 18% while it is 84% for German-speaking participants. These results suggest that sensitivity to referential (dis)continuity is not a community- or even language-specific property, but rather one that emerges from a general human predisposition to track and mark referents in linguistic discourse via pronominal forms (independently of language-specific settings for rates of pronoun use overall). Such findings have significant implications for theories of language variation and change, raising the possibility that at least some of the (most important) factors that give rise to patterns of variant choice are embedded not in community-specific norms, but rather in the common cognitive inheritance of human beings, as manifested in linguistic interaction.

Labov, W. (2011). *Principles of linguistic change, volume 3: Cognitive and cultural factors* (Vol. 3). John Wiley & Sons.

---

## Who Benefits?

### Bilingual Education Program Access for Immigrant Communities in New York City

Ivana Espinet, Kingsborough Community College, CUNY

Kate Menken, Queens College, CUNY

Maite T. Sánchez, Hunter College, CUNY

María Cioè-Peña, University of Pennsylvania

This presentation shares the results of our mixed-methods research examining the access of multilingual learners (MLLs, also known as English learners) to bilingual education programs in New York City (NYC). Research indicates that MLLs in bilingual education programs typically outperform their peers in English-only programming and experience numerous academic, cognitive, and socio-emotional benefits (Baker & Wright, 2021; Goldenberg & Wagner, 2015; Umansky & Reardon, 2014). In spite of the ample research base in support of bilingual education and New York state policies that mandate bilingual education be provided, only 17% of all MLLs in NYC schools are actually enrolled in bilingual education programs. The city has in fact opened hundreds of new dual language bilingual education (DLBE) programs over the past decade, yet only a very small portion of MLLs - just 7% - are enrolled in them.

Our research involved demographic and geographic information system (GIS) analysis, as well as interviews with district and school leaders. NYC is divided into 32 school districts, and our analysis of NYC Department of Education data revealed wide disparities in the availability of bilingual education for MLLs by district. Additionally, our GIS mapping of the location of DLBE programs exposed significant inequities in access to these programs by race and language. Our findings offer clear policy recommendations by (a) suggesting that school district leaders play a significant role in determining access to bilingual education, thus education reform efforts should focus on school leadership, and (b) by highlighting where bilingual education programs should be provided to ensure equitable access to them.

---

## Changing structures of covariation

Lewis Esposito

Stanford University, USA

Whether changes in progress covary across speakers has been a hot topic in recent years (e.g., Becker 2016; Beaman 2022; Haddican et al. 2022; Tamminga 2019; Waters and Tagliamonte 2017), but little work has focused on the structure of covariation patterns that do surface (e.g., how interrelated are they?). In this paper, I show that the structure of covariation (1) can surface in diverse ways at the community level, (2) can change over time, and (3) can shed light on the sources of covariation more broadly.

The empirical focus is on 8 changes in progress in Sacramento, California English: (1,2) TOO/TOE fronting, (3) TRAP lowering/backing, (4) increasing PIN-PEN Euclidean distance, (5) /s/ fronting, (6) increasing use of quotative *be like*, (7) increasing use of adjectival intensifier *really*, and (8) increasing use of creaky voice. Except for TOO/TOE, these changes are internally unrelated. Data come from sociolinguistic interview speech (collected in 2015) from an age-stratified and racially/gender-diverse sample of 100 English speakers born/raised in Sacramento.

Interviews were FAVE-aligned, and midpoint acoustic measures were extracted from segments >40 msec – vowel formants with FAVE-extract, and /s/ center of gravity with a Praat script that band-pass filtered frequencies before measurement. Formants were Lobanov normalized. Each vowel was coded as +/- creaky using Kane et al.'s (2013) automatic creak detection. Quotative verbs and adjectival intensifiers, and their internal constraints (Tagliamonte and D'Arcy 2007), were coded manually with transcripts.

To assess covariation over time, the sample was split into two demographically comparable cohorts of 50 speakers each: Generation 1 (SilentGen+BabyBoomers), and Generation 2 (GenX+Millennials). Mixed-effects regressions containing only linguistic controls were constructed for each change within each cohort, and by-speaker intercepts were extracted from each model to capture individual speaker behavior. Pearson product-moment correlations were run on the sets of intercepts to assess covariation between all pairs of changes. I control the false discovery rate by adjusting p-values through the Benjamini-Hochberg procedure.

The changes that significantly correlate are visualized in Figure 1, and clear structural differences emerge over time. Generation 1 shows three significant patterns with no interrelations among them. Generation 2, on the other hand, shows 6 significant patterns with a number of interrelations. TRAP is at the heart of these differences over time, covarying with 4 changes in Generation 2. Importantly, in Generation 2, TRAP displays more robust social

conditioning than any other change in either cohort, with effects of gender, race, birthyear, and local orientation. Visualizations (not shown) suggest that the diversity of TRAP's social associations may be the driver of Generation 2's interrelations, as different covariation patterns pick up on different sets of social correlates (e.g., TRAP + intensifier covariation is mostly tied to gender and age variation, while TRAP + /s/ is related to locality). Overall, results suggest that the "interconnectedness" of covariation structures can vary. Further, linguistic changes (and linguistic forms) tied to a wide range of shared social associations/meanings might be expected to play a role in highly interconnected structures.

- Becker, Kara. 2016. Linking community coherence, individual coherence, and bricolage: The co-occurrence of (r), raised BOUGHT and raised BAD in New York City English. *Lingua*.
- Beaman, Karen. 2022. Coherence in Real- and Apparent-Time: A Sociolinguistic variationist investigation of language change Swabia. PhD Dissertation, Queen Mary, University of London.
- Haddican, Bill, Cece Cutler, Michael Newman, and Christina Tortora. 2022. Cross-speaker covariation across six vocalic changes in New York City English. *American Speech*.
- Tagliamonte, Sali, and Alex D'Arcy. 2007. Frequency and variation in the community grammar. *Language Variation and Change*.
- Tamminga, Meredith. 2019. Interspeaker covariation in Philadelphia vowel changes. *Language Variation and Change*.
- Waters, Catherine, and Sali Tagliamonte. 2017. Is one innovation enough? Leaders, covariation, and language change. *American Speech*

---

## Sarcasm Socially Contextualized: How Socio-indexical Information Modulates the Evaluation of Pragmatic Behavior

Sophie Faircloth, Andrea Beltrama  
University of Pennsylvania, Philadelphia, USA

Interlocutors commonly deploy pragmatic reasoning to infer social meaning(s) ([1-2-3-11-14]) – e.g. they perceive speakers using precise numbers (“\$297”) as higher in Status//lower in Solidarity than those using approximate ones (“\$300”; [3]). While affording an important perspective on the link between linguistic forms and social indexicality, these studies looked at the emergence of socio-pragmatics meanings in bare, highly stylized contexts. But interlocutors typically draw social inferences against a richer backdrop of information about speakers, which go well beyond what they say in a single utterance; this raises the question of how comprehenders utilize speaker information to evaluate pragmatic moves.

We address this through the lens of Sarcastic Jabs (SJ) — e.g., uttering “great job” to comment on someone’s failure. SJs feature two distinctive pragmatic components: (i) an intention to threaten the interlocutor’s face ([10-18]); (ii) doing so by flouting the Quality Maxim ([12]). We make two predictions. Hyp1: authors of SJs should be perceived even more unfavorably than those of a Literal Jabs (LJs): besides having a malevolent intention, the speaker is forcing the interlocutor to engage in additional pragmatic reasoning to make sense of the Quality violation (see [1-2-11-14] on how markedness highlights social meanings’ salience). Hyp2: the impact of SJs on speakers’ evaluation should be modulated by the speaker’s identity – e.g., by whether using SJs aligns with the linguistic practices typical of the speaker’s persona.

We’ll test Hyp1 via a perception experiment (Exp1) showing native speakers of American English 12 dialogues (6 with SJs; 6 with LJs; Fig.1) and having them rate the speaker along 8 scales (Fig.2). Following Hyp1, SJs should lead to lower Solidarity ratings than LJs (tested with PCA on evaluative ratings + ME models on resulting underlying factors). To investigate Hyp2 we’ll explore how the evaluation of SJs is affected by whether the speaker embodies a Nerdy persona, a widely investigated social type whose salient social traits – e.g., high intellectual engagement; low sociability; resistance to fitting in ([4-7-15-16-17]) – align with those of a typical SJ user, also consistent with popular media representation (e.g. Sheldon in Big Bang Theory). One possibility (Hyp2A) is that comprehenders ascribe greater social significance to SJs when uttered by typical users – leading to lower Solidarity ratings for Nerdy speakers. Alternatively (Hyp2B) SJ might impact the speaker evaluation more prominently when at odds with the speaker’s typical pragmatic practices – leading to lower Solidarity ratings for non-Nerdy speakers. We’ll test these in Exp2, implementing a 2x2 design: Utterance (SJ vs. LJ; =Exp1) and Persona (Nerdy vs. non-Nerdy; cued visually; Fig.3). Participants will evaluate the speaker using the same scales as Exp1. Evidence in favor of Hyp 2A/2B will be revealed by the nature of the Utterance\*Persona interaction on the Solidarity ratings. Even if (contra Hyp1) SJ weren’t perceived more unfavorably than LJs, the Persona manipulation may still reveal informative differences in how SJ users are socially evaluated. We’ll also explore how participants’ responses relate to their own attitude towards sarcasm ([13]).

1. Acton, E. and C. Potts (2014). That straight talk. Sarah Palin and the sociolinguistics of demonstratives. *Journal of Sociolinguistics* 18(1), 3–31
2. Beltrama, A. (2020). Social meaning in semantics and pragmatics. *Language and Linguistics Compass*, 14(9), e12398.
3. Beltrama, A, S. Solt and H. Burnett. Context, precision and social perception: a socio-pragmatic study. *Language in Society*.

4. Bishop, J. H., Bishop, M., Bishop, M., Gelbwasser, L., Green, S., Peterson, E., ... & Zuckerman, A. (2004). Why we harass nerds and freaks: A formal theory of student culture and norms. *Journal of School Health*, 74, 235-251.
5. Bonnefon, J.F. and G. Villejoubert (2006). Tactful or doubtful?: Expectations of politeness explain the severity bias in the interpretation of probability phrases. *Psychological Science* 17 (9), 747-751.
6. Braden, S. K. (2020). "Scientists Can't Really Talk to People": Unpacking Students' Metacommentary on the Racialized and Gendered Science Nerd Trope. *International Journal of Multicultural Education*, 22(2), 87-108.
7. Bucholtz, M. (2001). The whiteness of nerds: Superstandard English and racial markedness. *Journal of linguistic anthropology*, 11(1), 84-100.
8. Campbell-Kibler, K. (2010). Perception in sociolinguistics. *Language and Linguistics Compass*. 4(6), 377-389
9. Eckert, P. (2008). Variation and the indexical field. *Journal of sociolinguistics*, 12(4), 453-476
10. Gibbs, R. W. (1986). On the psycholinguistics of sarcasm. *Journal of experimental psychology: general*, 115(1), 3.
11. Glass, L. (2015). Strong necessity modals: Four socio-pragmatic corpus studies. *University of Pennsylvania Working Papers in Linguistics* 21(2).
12. Grice, H. P. (1975). *Logic and conversation*. In *Speech acts*(pp. 41-58). Brill.
13. Ivanko, S. L., Pexman, P. M., & Olineck, K. M. (2004). How sarcastic are you? *Journal of Language and Social Psychology*, 23(3), 244-271.
14. Jeong, S. (2021). Deriving politeness from an extended lewisian model: The case of rising declaratives. *Journal of Pragmatics* 177, 183-207
15. Kendall, L. (2000). "Oh no! I'm a nerd!" Hegemonic masculinity on an online forum. *Gender & Society*, 14(2), 256-274.
16. Kinney, D. A. (1993). From nerds to normals: The recovery of identity among adolescents from middle school to high school. *Sociology of education*, 21-40.
17. Mairesse, F., & Walker, M. (2006). Words mark the nerds: Computational models of personality recognition through language. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 28, No. 28).
18. Reyes, A. (2013). Linguistic-based patterns for figurative language processing: the case of humor recognition and irony detection. *Procesamiento del Lenguaje Natural*, 50, 107-109.

---

## Moving beyond prosodic transcription in the analysis African American English intonation: a case study with BIN utterances

Alessa Farinella, Kristine M. Yu, Lisa Green  
University of Massachusetts, Amherst, USA

Autosegmental-Metrical Theory prioritizes fundamental frequency ( $f_0$ ) targets and turning points, but there is evidence that tonal shape also plays a role in perception [1]. In addition to the bias of focusing on discrete tonal events in intonational phonology, transcription conventions bring additional assumptions about the type and distribution of tonal events. In work on African American English (AAE), transcription conventions developed for Mainstream American English (MAE-ToBI) may not accurately capture the intonational phonology [2], and emphasis on  $f_0$  targets may result in overlooking important prosodic aspects. Beyond AAE, this is a broader issue when studying intonation in less well-studied varieties.

This paper focuses on two constructions found in some varieties of AAE, which are string identical but differ in their meaning and prosody: utterances containing "stressed BIN" (BIN) [3] (Fig. 1) and utterances containing the past participle (been) (Fig. 2). BIN has been described as having higher relative  $f_0$  and intensity on "BIN" compared to the past participle been and reduced  $f_0$  following "BIN" [4].

We take a data-driven approach that moves beyond pointwise  $f_0$  measurements and transcription labels. Data come from recordings of 8 speakers of AAE (5 female, 3 male) from a small, historically-segregated African American community in southwest Louisiana, previously analyzed in [4]. Sentences were elicited with a range of syntactic contexts containing the remote past marker BIN or the past participle been, resulting in about 770 tokens. Recordings were transcribed following the conventions of MAE-ToBI [5]. While [4] focused only on  $f_0$  averages, here we compare dynamic shape measurements to ToBI transcriptions.

Fig. 3 shows  $f_0$  contours by been-type. While there was some individual variation in  $f_0$  scaling, the contours are very consistent across speakers. In BIN utterances, there's a clear  $f_0$  peak on "BIN" followed by reduced  $f_0$ . By contrast, been utterances show no  $f_0$  peak on "been", but clear  $f_0$  peaks following "been." Table 1 shows the proportion of accents transcribed on "BIN/been" by been-type. Nearly all BIN utterances were transcribed as having a pitch accent (most commonly L+H\* and H\*), matching the shape analysis. Slightly over 40% of been utterances also had an accent (most commonly H\*), despite the flatter  $f_0$  on "been" in Fig. 3.

Table 2 shows the proportion of accents present in the post-"BIN/been" region by been-type. Almost all utterances contain an accent(s) following "BIN/been." These labels fail to capture the post-"BIN" compressed  $f_0$  in BIN utterances, which may contribute to BIN's description in past literature as "stressed." By relying on prosodic transcription alone, we lose information about the striking  $f_0$  range differences evident in Fig. 3. This work highlights the use of tonal shape methods to supplement transcription in order to reduce phonological assumptions and the reliance on tonal targets, especially when working on a variety like AAE [6]. We offer an alternative method

of visualizing and capturing global f0 trends that does not rely on transcription of tonal targets, which can ultimately inform phonological analysis.

- [1] Barnes, J., Veilleux, N., Brugos, A., & Shattuck-Hufnagel, S. (2012). Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology. *Laboratory Phonology*.
- [2] Burdin, R. S., Holliday, N. R., & Reed, P. E. (2022). American English pitch accents in variation: Pushing the boundaries of mainstream American English-ToBI conventions. *Journal of Phonetics*, 94.
- [3] Rickford, J. R. (1973). Been in Black English.
- [4] Green, L., Yu, K. M., Neal, A., Whitmal, A., Powe, T., & Özyıldız, D. (2022). Range in the Use and Realization of BIN in African American English. *Language and Speech*, 65(4), 958-1006.
- [5] Beckman, M. & Elam, G. A. (1997), 'Guidelines for ToBI labeling'.
- [6] Li, A., Purse, R., & Holliday, N. (2022). Variation in global and intonational pitch settings among black and white speakers of Southern American English. *The Journal of the Acoustical Soc of Am*, 152(5).

---

## Bidialectal 7-9-year-olds' evaluative labels for the standard and local varieties of Hungarian

Krisztina Feher

University of Debrecen, Debrecen, Hungary

Little is known about the linguistic preferences of young children who live in a community using a wide range of speech forms on a continuum from the standard variety to a local dialect (for a pioneering study, see Kaiser & Kasberger, 2021). There has been research on the emergence of young children's metalinguistic awareness in a bidialectal environment (Papavavlou & Phili, 2009; etc.). Yet, the development of young children's dialectal awareness of linguistic prestige has not been investigated. Furthermore, while quantitative data are usually being gathered from young children by applying experimental methods (Kinzler & DeJesus, 2013; Paquette-Smith, Bucker & Johnson, 2023; etc.), qualitative (Cornips, 2020) and mixed-methods (Choi et al, 2021) data are scarce. There is no qualitative and mixed-methods research specifically on young bidialectals' evaluations of varieties. In addition, young children's preferences toward varieties of languages such as Hungarian are understudied. There is only one study (Fehér, 2020) on Hungarian in the field.

In the above-mentioned study, both quantitative and qualitative data have been collected from Hungarian-speaking bidialectal kindergarteners at the age of 5-7. It has been demonstrated that 5-7-year-olds have already preferred the standard variety to the local dialect. However, they have not yet shown sociolinguistic awareness of their preferences. As a continuation of that study, the present paper investigated when and how awareness of the prestige of varieties along with judgmental behavior emerged and developed in the same population.

Eighty-eight 7-9-year-old first graders from the same community were invited to participate in a child-friendly matched-guise experiment followed by mini-interviews. The participants were asked to choose between two teachers telling the same tale and to explain their choices. The direction and degree of preferences were determined by statistical analysis of the choices, and evaluative labels were identified by content analysis of the explanations. Results were compared with the previous study applying the same methodology.

Like the kindergarteners, the first graders also preferred the standard variety of Hungarian to the local dialect. However, quantitative data shows that their preference toward the standard variety became stronger. A qualitative leap in development was also found between the kindergarteners and the first graders in how they labeled the varieties.

First graders had much more linguistic comments on the guises than kindergarteners. When commenting, only a few tokens and types of evaluative labels, based on the "good/bad" and "(not) beautiful" dimensions, were used by some kindergarteners. In first graders, the earlier dimensions remained the most robust, but various novel label types also showed up and were mostly used to judge the regional dialect: "funny", "weird", "not normal", etc. The diversity of their direct references also became greater. While kindergarteners used the labels with reference to "voice", "speech", "tale" and "person", first graders applied novel references such as "storytelling" and "language" and used label-reference collocations that were characteristic of each variety. Interestingly, first graders followed the stereotypical concept of NORM-speakers ("non-mobile, older, rural, male"), known from dialect geography, when they characterized the speakers of the local variety as ORM-speakers ("older, rural, male").

Fehér, K. (2020). Status-based preference of bidialectal kindergarteners: An experimental study. *Argumentum* 16, 147-172. doi:10.34103/ARGUMENTUM/2020/10

Choi, N., Cho, H-J., Kang, S. & Ahn, H. (2021). Korean children's attitudes towards varieties of English: The role of age and English learning environment. *Languages* 6(3), 133. <https://doi.org/10.3390/languages6030133> [<https://doi.org/10.3390/languages6030133>]

Cornips, L. (2020). The impact of preschool attendance on children's bidialectism in the Netherlands: Why toddlers may stop speaking a regional language (Limburgish) at home. *Language in Society*, 49(3), 333-355. <https://doi.org/10.1017/S0047404520000275> [<https://doi.org/10.1017/S0047404520000275>]



- Kaiser, I., & Kasberger, G. (2021). Children's sociolinguistic preferences: The acquisition of language attitudes within the Austrian standard-dialect continuum. In Ghimenton, A., Nardy, A. & Chevrot, J. P. eds., *Sociolinguistic variation and language acquisition across the lifespan*. Amsterdam, The Netherlands: John Benjamins, 130-160.
- Kinzler, K. D. & DeJesus, J. M. (2013). Northern = smart and Southern = nice: The development of accent attitudes in the United States. *The Quarterly Journal of Experimental Psychology* 66(6), 1146–1158. <https://doi.org/10.1080/17470218.2012.731695> [<https://doi.org/10.1080/17470218.2012.731695>]
- Papapavlou, A. & Phili, A. (2009). Childhood bidialectism and metalinguistic awareness. *Selected Papers on Theoretical and Applied Linguistics* 18, 311–320. <https://doi.org/10.26262/istal.v18i0.5450> [<https://doi.org/10.26262/istal.v18i0.5450>]
- Paquette-Smith, M., Bucker, H., & Johnson, E.K. (2023). How sociolinguistic factors shape children's subjective impressions of teacher quality. *Quarterly Journal of Experimental Psychology*, 76(3), 485 -496. doi:10.1177/17470218221094312

---

## The Social Meanings of Definite Articles with Proper Names in Spanish

Evelyn Fernández-Lizárraga  
Stanford University, Stanford, USA

In Spanish, variation in the use of definite articles with proper names (e.g., La María; henceforth DA+PN) is a robust site for social meaning. DA+PN's display extreme variability in affect, from showing affection (1) to disparaging (2) a referent. However, DA+PN's can serve other discourse functions without expressing an extreme attitude towards the referent (3-4). Following other socio-pragmatic analyses of determiners (Acton & Potts, 2014; Acton 2019, 2021), I argue that these pragmatic effects can be derived from the meanings of definite articles and proper names; DA+PN's mark a referent as both "salient" (DA) and "hearer-old" (PN). Furthermore, the marked co-occurrence of an article with a name invites inferences concerning the relation between a referent and the conversation participants. In using DA+PN's, speakers draw heavily on the common ground to convey a range of meanings regarding their referents.

To test the role of common ground, I ran an online experiment on Amazon's Mechanical Turk. Participants (N=131) rated target sentences along a sliding scale for felicity. For each critical trial, participants randomly received either a neutral context with no common ground (i.e., the listener does not know the referent) or with common ground (i.e., the listener knows the referent) (5). Each context was randomly paired with a target sentence that either did or did not contain the DA. I additionally tested the role of affect by using a sliding scale (0=negative, 1=positive) to gauge baseline affective interpretations of DA+PN's in neutral contexts (cf., Christodoulelis, 2017; Tieperman, 2020).

The results demonstrate that when a name is in the common ground, DA+PN's can be felicitous without an extreme attitudinal context. I ran two mixed effect linear regression models to test the relationship between DA presence, context condition, and slider ratings. For felicity (Figure 1), there was a main effect of DA presence ( $\beta = 0.11$ ,  $SE = 0.02$ ,  $t = 6.46$ ,  $p < 0.0001$ ) and context condition ( $\beta = 0.07$ ,  $SE = 0.02$ ,  $t = 4.73$ ,  $p < 0.001$ ), such that target sentences with No DA or in common ground contexts received higher felicity ratings. For affect (Figure 2), there was a main effect of DA presence ( $\beta = 0.03$ ,  $SE = 0.01$ ,  $t = 3.85$ ,  $p < 0.001$ ) and context condition ( $\beta = 0.02$ ,  $SE = 0.01$ ,  $t = 3.34$ ,  $p < 0.01$ ), such that target sentences with No DA or with common ground were rated more positively.

I provide a pragmatic account for how social meanings emerge as speakers leverage the referential functions of definite articles and proper names. Definite articles can construct referents as "salient" (Epstein, 2002; Becker 2021), while proper names presuppose that the hearer already knows the referent (Prince, 1992). The use of a definite article with a proper name does not predict a particular affective quality, but rather, that speakers will use DA+PN's in contexts where it is pragmatically useful to mark a referent as "salient" and "hearer-old," opening the door for many types of social evaluation.

Acton, E. K. (2019). Pragmatics and the social life of the English definite article. *Language*, 95(1), 37–65. <https://doi.org/10.1353/lan.2019.0010>

Acton, E. K. (2021). Pragmatics and the Third Wave: The Social Meaning of Definites. In E. Moore, L. Hall-Lew, & R. J. Podesva (Eds.), *Social Meaning and Linguistic Variation: Theorizing the Third Wave* (pp. 105–126). Cambridge University Press; Cambridge Core. <https://doi.org/10.1017/9781108578684.005>

Acton, E. K., & Potts, C. (2014). That straight talk: Sarah Palin and the sociolinguistics of demonstratives. *Journal of Sociolinguistics*, 18(1), 3–31. <https://doi.org/10.1111/josl.12062>

Becker, L. (2021). Articles in the World's Languages. *De Gruyter*. <https://doi.org/10.1515/9783110724424>

Christodoulelis, E. (2017, October 27). Definite Articles with Anthroponyms Across Varieties of Spanish: Quantitative Support of Conventional and Conversational Implicature. *Hispanic Linguistics Symposium*, Texas Tech University, Lubbock, TX.

Epstein, R. (2002). The definite article, accessibility, and the construction of discourse referents. *Cognitive Linguistics*, 12(4). <https://doi.org/10.1515/cogl.2002.007>

Prince, E. F. (1992). The ZPG letter: Subjects, definiteness, and information-status. *Discourse Description: Diverse Analyses of a Fund Raising Text*, 295–325.

---

**Heritage speakers' categorization of variable phonetic input in Spanish**  
Sofia Fernandez<sup>1,2</sup>, Michael Gradoville<sup>1</sup>, Hunter Langenhorst<sup>1</sup>, Sarah Henderson<sup>1</sup>  
<sup>1</sup>Arizona State University, Tempe, USA. <sup>2</sup>New York University, New York, USA

The present study investigates how heritage language Spanish speakers categorize phonetic input from a non-local variety of Spanish to which some of the participants had previous exposure and others did not. While Leeman (2018) asserts the importance of fostering dialect awareness in the heritage language classroom to acknowledge and validate language variation, to the extent that we are aware, no previous study has addressed the ability of heritage speakers to accurately perceive non-local varieties of their language.

Although the voiceless palato-alveolar fricative [ç] does not exist in most varieties of Spanish, it can be an allophone of /tʃ/ or /j/, depending on the variety where it occurs (Lipski, 1994). In some Northern Mexican varieties of Spanish, including in Chihuahua and Sonora, [ç] occurs due to deaffrication of /tʃ/ while in Rioplatense Spanish, most speakers produce [ç] as the default variant of /j/. This study explores the extent to which heritage language speakers can learn and recognize speaker-specific phoneme-allophone correspondences that do not align directly with the Spanish varieties they have been previously exposed to.

Participants were enrolled in two fourth semester Spanish as a heritage language course sections (N = 29). Participants reported having parents of Mexican origin and having regular contact with Mexican Spanish. The two instructors of the courses were from Argentina and Bolivia. Participants completed a perception task, a language proficiency test, and a language background questionnaire.

In the perception task, participants listened to 192 sentences containing a carrier phrase with an invented place name at the end, such as *Me gusta el chocolate de Yafilo* 'I like the chocolate from Yafilo'. The invented place name had either one of the target sounds (/tʃ/, /j/) or distractor sounds (/r/, /x/). The carrier phrases fell into three categories: those containing the same phoneme as the invented word, those containing the "opposite" phoneme (/tʃ/ for /j/, for example), and those lacking any of the sounds (/tʃ/, /j/, /r/, or /x/). The participants selected an orthographic response, such as Yafilo, Chafilo, Yasilo, Chasilo, based on what they heard. The sentences were spoken by four female speakers of different regional varieties: Rioplatense, Puerto Rican, Sonorenses, and Central Mexican Spanish. The data were then analyzed using a logistic mixed-effects model.

The results indicate that participants were very accurate in their phonemic identification in most combinations of speaker and phoneme ( $\geq 89.5\%$  accurate). In identifying Rioplatense /j/, produced as [ç], overall accuracy was only 10.5% with 62.1% of participants failing to identify any of these items accurately. Participants from the class with the Rioplatense instructor did not have an appreciably higher accuracy rate (11.9%) than the participants with the Bolivian instructor (10.1%). While we can tentatively assume that real words containing /j/ from Rioplatense Spanish do not greatly impede comprehension by the heritage speaker participants of this study, our findings raise the question of how much input from another variety of a language is necessary in order to accurately perceive such sounds in unfamiliar words.

Leeman, J. (2018). Critical language awareness and Spanish as a heritage language: Challenging the linguistic subordination of US Latinxs. In K. Potowski (Ed.), *Handbook of Spanish as a minority/ heritage language* (pp. 345–358). Routledge.  
Lipski, J. (1994). *Latin American Spanish*. Longman.

---

**Cross-linguistic variation in the domain of intentionality and its consequences for real-life legal outcomes**

Luna Filipovic  
University of California, Davis, USA

Does the language we speak affect what details we are likely to provide about them in our spoken or written accounts? Linguistic research in the last 20 years has confirmed that this may indeed be the case (e.g. Fausey & Boroditsky 2011; Filipović 2007; Slobin 2003, 2016). And if we are bilingual or multilingual, would the choice of language for the description of an event, or our strongest language, lead us to focus on certain event features and omit or neglect others that we could have focused on if we had used our other language(s)? The research presented in this paper is about effects of cross-linguistic variation in an applied context of high social relevance, namely the context of the law and law enforcement. It is an interdisciplinary study that taps into the cognitive domain of intentionality (i.e. is something done on purpose vs accidentally), crucial for the law and expressed differently in different languages. Previous research has found some languages, e.g. Spanish (Filipović 2007, 2013; Ibarretxe-Antuñano 2012; Ariño-Bizarro 2023) have different verbs and constructions for drawing intentionality distinctions

that speakers consistently use (e.g. *Lo rompió* – He broke it vs. *Se le rompió* – meaning something like “it happened to him that it broke”, which does not really exist in English). Monolingual speakers of this language as well as bilinguals who are Spanish-dominant consistently focus on this crucial distinction across modalities - in the verbalisation of events, in memory and in their gestures. Some other languages, e.g. English, do not draw this distinction with consistency and they habitually leave this information ambiguous. For instance, it was found that expressions such as *He broke it* and *It broke* are used interchangeably in descriptions of both intentional and non-intentional actions because they are underspecified with respect to intentionality (Filipović 2013, 2018). Further information would be needed in English to clarify which meaning was intended, and such clarification is typically missing because it is not obligatory.

The goal of this talk is to examine how and why serious miscommunication may occur as a result of a) cross-linguistic semantic variation in terms of the means that two languages have for expressing different types of intentionality, b) the frequency with which the relevant intentionality distinctions (i.e. intentional vs. non-intentional) are drawn in each, and c) the accuracy of rendering the relevant information in cross-linguistic (bilingual) communication. The dataset consists of authentic transcribed police interrogations with victims, witnesses and suspects from different jurisdictions in the state of California, USA, 100 cases in total, 50 monolingual and 50 bilingual. The study is both quantitative and qualitative, showing that: i) there is a significantly higher occurrence of miscommunication in bilingual than monolingual police interrogations and ii) there are qualitatively different sources of miscommunication (linguistic and cultural). The research findings are discussed in the context of equality in access to justice for non-native speakers and concrete proposals are made for applications of the findings in academic education and professional development.

- Ariño-Bizarro, A. (2023) Estudio psicolingüístico y tipológico de la causividad en español. University of Zaragoza, PhD dissertation.
- Fausey, C., & Boroditsky, L. (2011) Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin and Review*, 18, 150–157.
- Filipović, L. (2007) Language as a witness: Insights from cognitive linguistics. *Speech, Language and the Law* 14 (2), 245-267.
- Filipović, L. (2013) Constructing causation in language and memory: Implications for access to justice in multilingual interactions. *International Journal of Speech, Language and the Law*, 20, 1–19.
- Filipović, L. (2018) Speaking in L2 but thinking in L1: Language-specific effects on memory for causation events in English and Spanish. *International Journal of Bilingualism*, 22, 180–198.
- Ibarretxe-Antuñano, I. (2012) Placement and removal events in Basque and Spanish. In A. Kopecka & B. Narasimhan (Eds.), *The events of putting and taking. A crosslinguistic perspective* (pp. 123–143). Amsterdam: John Benjamins.
- Slobin, D. I. (2003) Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–191). Cambridge, MA: MIT Press.
- Slobin, D. I. (2016) Thinking for speaking and the construction of evidentiality in language contact. In M. Güven, D. Akar, B. Öztürk, & M. Keleşir (Eds.) *Exploring the Turkish linguistic landscape: Essays in honor of Eser Erguvanlı-Taylan* [Studies in Language Companion Series 175] (pp. 105–120). Amsterdam: John Benjamins.

---

## Hedges and apologies in ChatGPT responses to African-American English and Mainstream U.S.

English

Eve Fleisig

University of California, Berkeley, Berkeley, USA

Large language models associate African-American English (AAE) with hate speech and stereotype Black users (Sap et al., 2019; Sheng et al., 2021), suggesting that model outputs do not prioritize AAE speakers. This work examines whether ChatGPT engages in different negative politeness strategies (Brown and Levinson, 1987) when responding to AAE versus Mainstream US English (MUSE) inputs, focusing on its use of apologies and hedges. I sampled AAE and MUSE tweets (100 of each,  $\geq 5$  words, hashtags/@mentions removed) from Blodgett et al. (2016)’s corpus. Tweets were sent as inputs to ChatGPT’s underlying gpt-3.5-turbo model; responses were coded following Table 1’s schema.

Conditional apologies (Figure 1). Unconditional and conditional illocutionary force-indicating devices (IFIDs) occur at similar rates in responses to both varieties. However, expressing regret was more common in responses to AAE (49%, vs. 31% for MUSE,  $p < .05$ ).

Explanations (Figure 2) and hedging (Figure 3). Explanations within apologies responding to MUSE tended to mention self-deficiency (21%), such as model limitations, more often than appeals unrelated to model weaknesses (12%), such as stating that something could not be understood. However, responses to AAE referenced self-deficiency less often (14%) than unrelated excuses (25%). The difference in frequency of excuses unrelated to self-deficiency between AAE and MUSE is significant ( $p < .05$ ). Use of approximators (e.g., “sometimes” or “about”)

also occurs less often in responses to AAE (3% vs. 17% for AAE,  $p < .01$ ). Differences for other hedging/apology types were not significant.

The responses suggest that ChatGPT uses apologies that are less threatening to its own “face” and only partially satisfy the addressee’s face when the input is in AAE. ChatGPT uses significantly fewer approximators in responses to AAE, suggesting that it appears more assertive and less cautious in response to AAE inputs. Although ChatGPT expresses regret more often in response to AAE, explanations in apologies responding to AAE are less likely to reference model limitations, which casts blame on the model alone, and more likely to state that there was a communicative failure, which implicates both model and user. Thus, there appears to be less commitment to satisfying the user’s negative face relative to protecting the model’s “face” when the user writes in AAE.

- Blodgett, S. L., Green, L., & O’Connor, B. (2016). Demographic Dialectal Variation in Social Media: A Case Study of African-American English. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 1119-1130).
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language use*. Cambridge Univ. Pr.
- Fraser, B. (2010). Pragmatic competence: The case of hedging. *New approaches to hedging*, 1534.
- Olshtain, E. (1989). *Apologies and Remedial Interchanges*. Berlin: Mouton de Gruyter.
- Sap, M., Card, D., Gabriel, S., Choi, Y., & Smith, N. A. (2019). The risk of racial bias in hate speech detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 1668-1678).
- Sheng, E., Chang, K.-W., Natarajan, P., & Peng, N. (2019). The woman worked as a babysitter: On biases in language generation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*.

---

## Some early observations of an incipient merger of /i/ and /y/ in Gothenburg Swedish

Julia Forsberg<sup>1</sup>, Johan Gross<sup>2</sup>

<sup>1</sup>Stockholm University, Stockholm, Sweden. <sup>2</sup>University West, Trollhättan, Sweden

This study investigates an incipient change in the vowel system of Gothenburg Swedish, where a possible outcome is a merger of the two front vowels /i/ and /y/ to the horizontal centralized [i]. The /i/-vowel has historically been described as “extra” fronted in the traditional Gothenburg dialect (Björseth 1957). However, studies have shown that the vowel now has a more centralized quality among adolescents who live in or go to school in the city center (Author 2 2018; Author 2 & 1 2020).

The current study is based on two investigations. In the first investigation, acoustic measurements from recordings of 19 speakers (16-19 years old) are analyzed to locate the position of the two vowels in the vowel system. In total, 3059 vowel tokens covering the whole vowel system were analyzed (409 tokens of /y/ and 296 of /i/). Acoustical values were normalized using the Lobanov algorithm and used as dependent variables in a MANOVA with the two vowels as independent variables. The statistical test showed no significant differences between the two vowels with regards to the first and second formant, and that the quality of both thus can be described as somewhat lowered and retracted [i].

The second investigation focuses on the perceptual differences between the two vowels through a perceptual experiment of the recorded tokens used in the first investigation. 22 words were chosen based on the sound quality, and the vowels were then manually transcribed. This procedure resulted in four categories: (i) where /i/ is produced as [i]; (ii) where /i/ is [i]; (yy) where /y/ is [y], and; (yi) where /y/ is [i]. The data was collected through a web form, where 203 informants classified the vowels appearing in the original word context and in isolation. Results show that informants are consistent in their classifications of the two /i/ variants as < i >, while (yy) is harder to classify and (yi) is most often classified as < i >.

The overall conclusion points in the direction of an incipient merger of the two vowels where the resulting quality will be the unrounded, closed, horizontally centralized vowel [i]. As the sociolinguistic factors shown in previous research to be important are constant across the group of speakers, and all can be considered native speakers of Gothenburg Swedish, the merger can be explained using the merger-by-approximation-model (Lennig 1978), where two phonemes gradually move toward each other and the resulting form will be an “in-between” form. These results also indicate that the language internal motivation for the merger is the lack of perceptual strength of lip rounding in the disappearing phoneme, and the consequent perceptual similarities between the horizontally centralized [i:] and /y:/.

Björseth, B. (1957). *Göteborgsspråket*. Stockholm: Norstedts.

Gross (2018). Segregated vowels: Language variation and dialect features among Gothenburg youth. *Language Variation and Change*, 30(3), 315–336.

Gross & Forsberg (2020). Weak Lips? A Possible Merger of /i:/ and /y:/ in Gothenburg. *Phonetica*, 77(4), 268-288.

Lennig, M. (1978). *Acoustic measurement of linguistic change: The modern Paris vowel system* (PhD dissertation), University of Pennsylvania.

---

## Seeking a life without shame in the Haitian linguistic market: Understanding popular resistance to Kreyòl mother tongue education

William Scott Frager

University of Oxford, Oxford, United Kingdom

This study reports on ethnographic research conducted in Haiti in early 2022 concerning why Haitian non-elites resist mother tongue education (MTE) in Kreyòl Ayisyen (Haitian Creole). It explores how the Haitian linguistic market requires Standard French in many key domains, engenders powerful feelings of shame in Kreyòl-monolinguals, and influences popular opinion of MTE.

It has been 70 years since UNESCO stated that it is “axiomatic” that “education is best carried on through the mother tongue of the pupil” but that “between [this] proposition and its realization many complex and difficult problems arise” (UNESCO 1953, 6). Haiti at present is an exemplar of the tension between these two observations. Since being reinstated in November 2021, the Haitian Minister of Education, Nesmy Manigat, has proposed (though largely not yet realized) ambitious MTE reforms. These reforms aim to give a central role in primary and secondary education to Kreyòl, the only language of over 90% of Haitian citizens (Dejean 2006; Hebblethwaite 2012), in lieu of Standard French, the traditionally dominant language of education, largely spoken by only a small group of elites. While studies continue to affirm the advantages of MTE for increasing learning outcomes, reducing dropout rates, and building an emotionally and intellectually solid foundation for later second language acquisition (GCE 2013; Ouane and Glanz 2011), stakeholders should take heed of Haiti’s past experience with such initiatives. In the 1980s, the Haitian Ministry of Education attempted a similar MTE reform, popularly known as the Bernard Reform (DEN 1982). Ultimately failing in most of its objectives, the most puzzling source of resistance came from the monolingual majority - the intended beneficiaries of the policies - described as fearing their children to remain trapped in the “Kreyòl ghetto” (Prou 2009, 49). Recent studies show teachers and parents continue to hold negative views toward Kreyòl-based education (Jean-François 2006; DeGraff and Stump 2018).

To explore the nature of this resistance to MTE, I draw on ethnographic observations in central and southeast Haiti and 16 in-depth in-person interviews with non-elite Haitians from L1 Kreyòl households. Informed by the sociology of Pierre Bourdieu (1991), the interviews explored the relationship between the participants’ exposure to language at home, the “linguistic correction” they received in school, and their experiences as adults with language use in the larger Haitian linguistic market. Systematic qualitative analysis indicates that while participants have mixed positive and negative attitudes towards French, they ultimately share a common recognition of the sociolinguistic “game” of Haiti and the grave consequences of not playing it well. In particular, participants expressed shame for being unable to understand and use French when required, such as in education, business, and courtship.

Through an analysis of these findings, the research aims to help policymakers currently involved in Haitian MTE reform understand both a) how proper education sector-led policy can act on the larger Haitian linguistic market and make a more linguistically equitable society and b) what constraints but also opportunities the larger linguistic market currently places on potential avenues of language-in-education reform.

Bourdieu, Pierre. 1991. *Language and Symbolic Power*. Harvard University Press.

DeGraff, Michel, and Glenda S. Stump. 2018. “Kreyòl, Pedagogy, and Technology for Opening up Quality Education in Haiti: Changes in Teachers’ Metalinguistic Attitudes as First Steps in a Paradigm Shift.” *Language* 94 (2). Linguistic Society of America: e127–57. doi:10.1353/lan.2018.0030.

Dejean, Yves. 2006. *Yon Lekòl Tèt Anba Nan Yon Peyi Tèt Anba*. FOKAL.

DEN. 1982. “Decree of March 30th, 1982 i.e. Décret Organisant Le Système Educatif Haïtien En Vue d’Offrir Des Chances Egales à Tous et de Refléter La Culture Haïtienne.”

GCE. 2013. “Mother-Tongue Education: Policy Lessons for Quality and Inclusion.”

Hebblethwaite, Benjamin. 2012. “French and Underdevelopment, Haitian Creole and Development: Educational Language Policy Problems and Solutions in Haiti.” *Journal of Pidgin and Creole Languages* 27 (2): 255–302. doi:10.1075/jpcl.27.2.03heb.

Jean-François, Lesly. 2006. “Attitudes Des Édicateurs Envers Français et Le Créole: Le Cas d’Haïti.” Dissertation, Baton Rouge, LA: Louisiana State University and Agricultural & Mechanical College.

Ouane, Adama, and Christine Glanz, eds. 2011. “Optimising Learning, Education and Publishing in Africa: The Language Factor: A Review and Analysis of Theory and Practice in Mother-Tongue and Bilingual Education in Sub-Saharan Africa.” UIL/ADEA.

Prou, Marc. 2009. “Attempts at Reforming Haiti’s Education System: The Challenges of Mending the Tapestry, 1979-2004.” *Journal of Haitian Studies* 15 (1/2): 29–69.

UNESCO. 1953. “The Use of Vernacular Languages in Education.” In *Monographs on Fundamental Education VIII*. Paris, France: UNESCO.

---

## Getting socialized: Variation and change in the passive in Canada

Karlién Franco<sup>1</sup>, Sali Tagliamonte<sup>2</sup>

<sup>1</sup>KU Leuven, Leuven, Belgium. <sup>2</sup>University of Toronto, Toronto, Canada

According to recent research, the GET passive, got arrested, rather than BE, was arrested, has been increasing in frequency since 1850 (Hundt 2001). However, according to some studies, the two variants remain differentiated by semantic nuances like verb adversativity (arrested vs. awarded) and whether or not an animate agent is expressed, e.g. got arrested by the police (e.g. Leech et al. 2009:144-165, Quirk et al. 1985: 167-171). Other research assumes that the variants are interchangeable and attributes the choice of variant to social factors (Sneller & Fisher 2015, Weiner & Labov 1983). Recent studies on vernacular spoken data from Tyneside English (England) (Fehringer 2022) and from Victoria English (Canada) (Allen 2022) tested both linguistic and social factors and found that both play a role. In this paper, we aim to contribute new insights by analyzing GET vs. BE in vernacular English in Ontario Canada from 355 individuals born between 1884 and 1999. To assess dispersion, we focus on three community types: a metropolitan area, a small rural community, and a city far from the major urban centre.

We extracted all forms of BE and GET + participle using procedures from Schwarz (2015:158-159), leaving aside non-central passives and coding for features tested in earlier studies (e.g. Bohmann et al. 2023, Fehringer 2022), including adversativity (using the procedure of Bohmann et al. 2023), explicitness of agent and animacy, as well as broad social factors (community, perceived speaker gender, education level, occupation level and year of birth). We used conditional inference tree analysis and mixed-effects logistic regression to assess the effect of these factors on the choice of GET versus BE, while controlling for main verb and individual.

The results confirm that the GET passive is increasing over time consistent with reports in the literature and with recent results from British Columbia (Allen 2022) and there is a notable acceleration among individuals born after 1976. In contrast to the two recent studies on vernacular speech data, linguistic features play only a minor role in the multivariate analysis: only agent explicitness is significant. Instead, two main social factors exert a significant effect: the GET variant is favoured by men and by speakers with less education. In addition, the cities, regardless of location and size, have higher frequencies of GET than the small town. Taken together, the evidence suggests: 1) change-in-progress; 2) semantic bleaching of earlier contextual constraints (cf. Hundt 2001) and 3) geographic dispersion. We conclude that the rise of the GET passive in Canada is a change from below that has spread according to the cascade model of language change (Labov 2003). We will explore further what may have propelled GET forward in the late 20th century whether a zeitgeist effect (cf. Fruehwald 2017) or a more general process of colloquialization.

- Allen, Caroline. 2022. Get over time: a longitudinal variationist analysis of passive voice in contemporary English. Doctoral dissertation. University of Victoria.
- Bohmann, Axel, Julia Müller, Mirka Honkanen & Miriam Neuhaussen. 2023. A large-scale diachronic analysis of the English passive alternation. In Beatrix Busse, Nina Dmrukcić & Ingo Kleiber (eds.), *Language and Linguistics in a Complex World*. Berlin/Boston: De Gruyter. 31-55.
- Fehringer, Carol. 2022. The get-passive in Tyneside English. *English World-Wide* 43(3), 330-356.
- Fruehwald, Joseph. (2017). Generations, lifespans, and the zeitgeist. *Language Variation and Change* 29(1), 1-27.
- Hundt, Marianne. 2001. What corpora tell us about the grammaticalisation of voice in "get"-constructions. *Studies in Language* 25(1), 49-87.
- Labov, William. 2003. Pursuing the cascade model. In Peter Trudgill, David Britain & Jenny Cheshire (eds.), *Social Dialectology: In Honour of Peter Trudgill*. Amsterdam: John Benjamins. 9-22.
- Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London/New York: Longman.
- Schwarz, Sarah. 2015. Passive voice in American soap opera dialogue. *Studia Neophilologica* 87(2), 152-170.
- Sneller, Betsy & Sabriya Fisher. 2015. When GET got noticed: The emerging salience of GET-passives. *University of Pennsylvania Working Papers in Linguistics* 21(1), article 34.
- Weiner, E. Judith & William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19(1), 29-58.

---

## Individual variation in the processing of filler-gap dependencies in L2 Italian

Pamela Franciotti

CUNY, Graduate Center

This study investigates the well-attested subject/object asymmetry in the processing of filler-gap dependencies (e.g., relative clauses, wh-questions) in L2 Italian. The study aims to establish whether L2 learners at different levels of proficiency experience a delay in the processing of object-extracted

dependencies as attested cross-linguistically in L1 acquisition and adult processing. Two self-paced reading experiments, measuring both response times (RTs) at critical regions in the sentence and post-reading accuracy of responses to comprehension statements are presented to a group of L1 English-L2 Italian learners at different levels of proficiency, as well as to a control group of native Italian speakers. Experiment 1, testing subject/object RCs (e.g., *Il professore che ha chiamato lo studente/lo studente ha chiamato è uscito dalla biblioteca* ‘The professor who has called the student/the student has called left the library’) found that L2 learners and NSs alike have longer RTs at the region of matrix V (*è uscito/left*). Experiment 2, testing subject/object who-questions (e.g., *Chi ha/hanno chiamato gli studenti?* ‘Who has called the students/have the students called?’) found that L2 learners experience longer RTs in object wh-questions at a later region compared to the NSs controls, the region following the post-verbal subject. In both experiments, an initial analysis conducted with L2 participants categorically grouped (i.e., intermediate vs. advanced) masked the complete picture of the L2 processing patterns, which only emerged in an individual-level analysis treating proficiency score as a continuous variable. Overall, this study provides evidence in support of the hypothesis that L2 learners adopt the same parsing strategies in the L2 as native speakers do in their L1. Moreover, the study suggests that the variability observed between L1 speakers/L2 learners with respect to the locus of the asymmetry across RCs and wh-questions is likely due to the processing in the wh-question constructions of both the filler-gap and the VS word order dependencies. Most importantly, the study emphasizes the need to consider individual variation when analyzing data in L2 research.

---

### Word-gender and voice-gender perceptions in LGBTQ+ vs. cishet listeners

Valerie Freeman, Athena Kirby, Tori Courouleau  
Oklahoma State University, USA

Gender is a culturally-dependent construct that often categorizes people, social roles, and associated concepts or terms into a pre-defined set of discrete bins. However, recent conceptualizations of human social gender as a continuum rather than a binary dichotomy may influence listeners’ perceptions of both people and utterances. Some work has identified combinations of acoustic-prosodic features that contribute to listeners’ perceptions of speakers’ gender along a masculinity-femininity continuum (e.g., Houle et al. 2022; Merritt & Bent 2022; Munson et al. 2007), but little work has examined lexical effects. English has very limited grammatical or morpho-phonemic gender, allowing speakers to assign conceptual or semantic gender to words via culturally relevant associations or gender-role stereotypes. These associations may be influenced by individual experience, with listeners’ perceptions varying in relation to their exposure to greater or lesser variation in gender expression. Thus, this study examines relationships between word-gender associations, voice-gender perception, and experience with gender diversity.

In an online survey, college-aged participants (18-24) rated 20 English words as “very feminine”, “somewhat feminine”, “no/any gender”, “somewhat masculine” or “very masculine” to collect baseline data on the words’ conceptual gender. In a separate phase, different participants listened to five people of various gender identities reading the words and were asked to rate the words (not the speakers) using the same gender scale. Text and audio ratings were compared to show influences of perceived voice-gender on conceptual word gender, with ratings of gender-neutral words determining baseline voice-gender perceptions. Participants were divided into high and low “gender-diversity experience” groups; initially these comprised LGBTQ+ vs. cishet participants, with later procedures adding a questionnaire specific to their familiarity with and acceptance of gender diversity (e.g., trans and non-binary people and concepts).

Initial data included about 30 text raters (none LGBTQ+) and 30 listeners (about 1/3 LGBTQ+). Cishet listeners’ word-gender ratings were clearly stratified by perceived speaker gender (Figure 1), but LGBTQ+ listeners’ ratings were remarkably consistent across speakers (Figure 2). They rated only a few words differently when said by feminine voices compared to masculine or gender-neutral voices. That is, LGBTQ+ listeners were largely able to ignore voice-gender when asked to rate word-gender, while cishet listeners were strongly influenced by the speakers’ perceived genders.

This ability to “listen past” voice-gender may stem from familiarity with and acceptance of people identifying and presenting themselves on a gender continuum (in appearance, social behavior, and speech/voice). Listeners experienced with gender diversity may process lexical content in the task more directly, with less cognitive interference from the speech/voice patterns that inexperienced listeners may find distracting or unexpected (e.g., as a departure from male/female prototypes; Merritt 2022). Such gender-variety exposure should include LGBTQ+ allies as well, so current procedures include a questionnaire on attitudes and experience with trans and non-binary people and issues. Analysis of these listener factors as predictors of word-gender vs. voice-gender separation will be

presented in addition to the current results. If expectations hold, studies like this support increasing exposure to gender diversity.

- Houle, N., Goudelias, D., Lerario, M. P., & Levi, S. V. (2022). Effect of anchor term on auditory-perceptual ratings of feminine and masculine speakers. *Journal of Speech, Language, and Hearing Research*, 65(6), 2064-2080.
- Merritt, B. M. (2022). Perceptual representation of speaker gender. Doctoral dissertation, Indiana University.
- Merritt, B., & Bent, T. (2022). Revisiting the acoustics of speaker gender perception: A gender expansive perspective. *Journal of the Acoustical Society of America*, 151(1), 484-499.
- Munson, B. (2007). The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation. *Language and Speech*, 50(1), 125-142.

---

## A New Way of Analyzing Vowel Space Area

Josef Fruehwald

University of Kentucky, Lexington, USA

The calculation of speakers' vowel space areas (VSA), whether for studying infant directed speech (Kuhl et al. 1997; Cristia and Seidl 2014), cross-dialectal differences (Fox and Jacewicz 2017), sound change in progress (D'Onofrio, Pratt, and Hofwegen 2019), or sociolinguistic signs (Pratt 2023), usually involves the selection of representative vowel categories as the points of the vowel space polygon. In this paper, I propose an alternative method that combines probability density estimates and spatial methods originally developed for GIS to calculate speakers' VSAs without selecting any point measurements. I apply this technique to data from the Philadelphia Neighborhood Corpus (PNC), and find a similar diachronic vowel space compression to D'Onofrio, Pratt, and Hofwegen (2019).

A downside to choosing representative vowel categories to define a vowel space polygon is that the same set of vowels may not best characterize the vowel space from speaker to speaker, or dialect to dialect. Moreover, the use of point vowels may overemphasize some areas of the vowel space that are less commonly used, for example, in the case of back vowel fronting. Polygons defined with probability density estimates, on the other hand, will conform in shape relative to overall utilization of the vowel space, and don't require the selection of representative vowel categories. In Figure 1 (a), the 90% probability polygon contains 90% of the speaker's vowel data, with a polygon based on point vowels overlaid for comparison.

For every speaker's data in the PNC, I extracted the 90% probability density polygon using modified functions from the `{ggdensity}` package (Otto and Kahle 2023) for both their Nearey Normalized F1 and F2, as well as simple  $\log(\text{Hz})$ . These were then converted to spatial polygons, and their areas calculated using functions from the `{sf}` package (Pebesma 2018). In order to ensure any estimated effects of gender were not due to vocal tract length (VTL) differences, I also estimated every speaker's VTL following Johnson (2020), and fit a bayesian mediation model in `{brms}` (Bürkner 2017).

Figure 1 (b) shows that VSA estimates across speakers are numerically identical between Nearey normalized values versus simple  $\log(\text{Hz})$ . This is because, as a single factor scaling method (Barreda 2021), it merely shifts the position, but not the scaling of vowel spaces in the log-space. Figure 1 (c) shows that there is an effect of VTL on VSA. When properly controlling for the effect of gender on VTL and of VTL & gender on VSA in a mediation model, a reliable effect of VTL remains, but there is no remaining direct effect of gender on VSA. Figure 1 (d,e) shows that there is a downward trend on VSA over time, specifically for those with some higher education.

This density-based approach to calculating VSA can be applied as-is to more complex data sets, such as vowel trajectory data, without needing to choose any measurement point to stand-in for the point vowels, with F1 and F2 Hz measurements log-transformed.

- Barreda, Santiago. 2021. "Perceptual Validation of Vowel Normalization Methods for Variationist Research." *Language Variation and Change* 33 (1): 27–53. <https://doi.org/10.1017/S0954394521000016>.
- Bürkner, Paul-Christian. 2017. "Brms: An r Package for Bayesian Multilevel Models Using Stan" 80. <https://doi.org/10.18637/jss.v080.i01>.
- Cristia, Alejandrina, and Amanda Seidl. 2014. "The Hyperarticulation Hypothesis of Infant-Directed Speech." *Journal of Child Language* 41 (4): 913–34. <https://doi.org/10.1017/S0305000912000669>.
- D'Onofrio, Annette, Teresa Pratt, and Janneke Van Hofwegen. 2019. "Compression in the California Vowel Shift: Tracking Generational Sound Change in California's Central Valley." *Language Variation and Change* 31 (2): 193–217. <https://doi.org/10.1017/S0954394519000085>.
- Fox, Robert Allen, and Ewa Jacewicz. 2017. "Reconceptualizing the Vowel Space in Analyzing Regional Dialect Variation and Sound Change in American English." *The Journal of the Acoustical Society of America* 142 (1): 444–59. <https://doi.org/10.1121/1.4991021>.
- Johnson, Keith. 2020. "The F Method of Vocal Tract Length Normalization for Vowels." *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11 (1): 10. <https://doi.org/10.5334/labphon.196>.



- Kuhl, Patricia K, J E Andruski, I A Chistovich, L A Chistovich, E V Kozhevnikova, V L Ryskina, E I Stolyarova, U Sundberg, and F Lacerda. 1997. "Cross-Language Analysis of Phonetic Units in Language Addressed to Infants." *Science* 277: 684–86.
- Otto, James, and David Kahle. 2023. "Ggdensity: Interpretable Bivariate Density Visualization with 'Ggplot2'." <https://CRAN.R-project.org/package=ggdensity>.
- Pebesma, Edzer. 2018. "Simple Features for r: Standardized Support for Spatial Vector Data" 10. <https://doi.org/10.32614/RJ-2018-009>.
- Pratt, Teresa. 2023. "The Vowel Space as Sociolinguistic Sign." *Journal of Sociolinguistics*, 1–20. <https://doi.org/10.1111/josl.12617>.

---

## Word-initial /p/ in Khalkha Mongolian: Variation in connected speech

Corrina Fuller

UCLA, Los Angeles, USA

**Background:** This paper explores the variable realization of word-initial /p/ in Khalkha Mongolian through the lens of connected speech. While phoneme /p/ has two phonologically-conditioned allophones word-internally ([p] and [w]), word-initial /p/ does not alternate in careful speech. In connected speech, however, not only is word-initial /p/ variably realized as [p] or [w], it is at times deleted altogether, yielding a "zero" realization. In this study, I offer a characterization of the nature of the variable realization of word-initial /p/ through careful investigation of a small corpus of naturalistic speech (four Khalkha-language TEDx Talks). Then, I model the variation illuminated within the corpus data using a Maximum Entropy grammar (MaxEnt; Goldwater & Johnson, 2003), a probabilistic version of Optimality Theory (Prince & Smolensky, 1993) that generates probability distributions over all candidates.

**Claims:** I argue that the variation in the realization of word-initial /p/ is patterned, and that processes related to ease of articulation in connected speech motivate lenition (in the spirit of Kaisse's (1985) "fast speech rules"), while processes particular to certain morphosyntactic configurations condition reduction (Kaisse's "external sandhi rules").

**Results:** The corpus data show a strong preference for the [p] realization, followed by a slight preference for the zero realization, with [w] being the least preferred variant of word-initial /p/ (Figure 1). All four speakers follow this trend, demonstrating that the observed variation is patterned and generalizes beyond the idiosyncrasies of any individual speaker. The results also reveal a massive split between content and function words, with the less-faithful realizations of /p/ ([w] and zero) occurring almost exclusively in function words (Figure 2). However, the function words do not all behave as one. Figure 3 features the distribution of /p/ in functional types with token frequencies greater than 30.

**Model:** The variation observed in word-initial /p/ in function words can be captured using a MaxEnt grammar (summarized in Figure 4). Each input has three candidate output forms, one for each potential realization of word-initial /p/. Constraints in this grammar apply only to the realization of word-initial /p/. Markedness constraints favoring the [w] and zero realizations reflect the articulatory pressure to lenite in different phonological environments, while lexically-indexed faithfulness constraints represent morpheme-specific resistance to this pressure. Finally, a constraint encoding a structurally-conditioned listed allomorph of auxiliary *pai-* captures a case of auxiliary reduction. This allomorph of *pai-*, *pai-2*, appears only in a specific morphosyntactic configuration and is almost categorically realized as zero.

**Conclusions:** This study situates connected speech as the locus of variation in the realization of word-initial /p/ in Khalkha Mongolian. The lens of connected speech affords insights into the interplay of articulatory pressures affecting lenition and the morphosyntactic configurations that select phonetically-reduced allomorphs, both of which affect Khalkha word-initial /p/. The core finding of this study is that the variation concerning word-initial /p/ is patterned. Crucially, this patterning is systematic, but not categorical, reflecting Weinreich et al.'s (1968) postulation of "ordered heterogeneity" as the norm for all linguistic systems and the heart of the variationist enterprise.

Goldwater, S. & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model.

In Spenader, Jennifer & Eriksson, Anders & Dahl, Östen (eds.) *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120. Stockholm: Stockholm University, Department of Linguistics.

Kaisse, E. M. (1985). *Connected Speech: The interaction of phonology and syntax* (First Edition). Academic Press.

Prince, A., and Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar*. Technical report, Rutgers University Center for Cognitive Science. Oxford: Blackwell.

Weinreich U., Labov W. & Herzog, M.I. (1968). *Empirical foundations for a theory of language change*. Austin, University of Texas Press.

---

## Rapid and introspective processing of sociolinguistic associations of (ING) in context

Xin Gao<sup>1</sup>, Lacey Wade<sup>2</sup>

<sup>1</sup>University of Pennsylvania, USA. <sup>2</sup>University of Kansas, USA

### Introduction

Linguistic forms are associated with social information to index a wide range of social meanings. However, implicit and explicit socio-linguistic associations do not always align [1,2], leading to proposals that they are handled by different systems: i.e., rapid, implicit, automatic processing on the one hand, and slower, deliberative, introspective processing on the other [3]. To complicate matters, these associations are often context-dependent [4,5].

In this study, we follow up on previous studies by (1) employing maximally similar tasks that target (a) differences in the amount of deliberation permitted for a response and (b) whether strength of association is measured implicitly and explicitly and (2) investigating the role of social context in drawing both implicit and explicit sociolinguistic associations. By employing maximally parallel paradigms in implicit and explicit tasks, we anticipate observing significantly correlated association strengths if the rapid and introspective processes involve the same or overlapping mechanisms; conversely, if the associations and/or context effects of the two tasks do not consistently align with each other, it would serve as compelling evidence in support of different systems.

### Methods

We investigate sociolinguistic associations between (ING), where a verbal suffix can be produced as either -ing or -in, and two social characteristics: “nerdy” and “chill” (Figure 1) following [4]. We predict -ing will be more readily associated with the “nerdy” persona (linked to education and precision) and this association will be stronger in a classroom context (compared to a pub context, Figure 2).

Task 1 assesses more rapid, implicit sociolinguistic associations. Participants (N=53) were instructed to choose which speaker was more likely to have produced each item as quickly as possible, and the strength of this association was measured by reaction time. Task 2 assesses more deliberative associations and was set up almost identically to Task 1, but with no time pressure. Participants moved a slider bar to evaluate which character was more likely to be the speaker, where the slider bar number served as an explicit measure of association strength.

### Results

(ING) significantly influences both the responses in the implicit task and evaluation in the explicit task in the expected direction: words with the -ing are more likely to be assigned to the “nerdy” character. The role of context differs across the two tasks: in the implicit task, context neither significantly affects the response nor the difficulty of reaching a decision (i.e., response time); however, in the explicit task, the classroom setting yields a higher overall “nerdy” evaluation ( $p < 0.000$ , Figure 3), suggesting participants are more attuned to contextual effects in explicit association processing. Finally, there is no correlation between individuals’ speed of time-pressured choices in the implicit task and the strength of association evaluation in the explicit task for “nerdy” responses ( $p = 0.412$ , Figure 4, left panel), but there is a marginal correlation for “chill” responses ( $p = 0.049$ , Figure 4, right panel). We interpret our findings as evidence that implicit and explicit associations may to some extent draw on distinct processing mechanisms.

[1] McKenzie, Robert M., and Erin Carrie. “Implicit–explicit attitudinal discrepancy and the investigation of language attitude change in progress.” *Journal of Multilingual and Multicultural Development* (2018).

[2] Squires, Lauren. “Processing grammatical differences: Perceiving versus noticing.” *Awareness and control in sociolinguistic research* (2016).

[3] Campbell-Kibler, Kathryn. “Towards a cognitively realistic model of meaningful sociolinguistic variation.” *Awareness and control in sociolinguistic research* (2016).

[4] Hilton, Katherine, and Sunwoo Jeong. “The role of context in sociolinguistic perception.” *Linguistics Vanguard* (2019).

[5] Beltrama, Andrea, and Florian Schwarz. “Imprecision, personae, and pragmatic reasoning.” *Semantics and Linguistic Theory* (2021).

---

## Transmasculine Entrapped Vocality: Identification and Identity

Jayce Garner

Pomona College, Claremont, USA

“Entrapped FTM vocality,” coined by Constansis (2008), refers to a voice quality found in transmasculine or assigned female at birth (AFAB) nonbinary individuals who have undergone masculinizing hormone replacement therapy (HRT). Some of these individuals’ voices, as they deepen from hormonal exposure, acquire a “permanently hoarse” quality that lacks the harmonics of a cisgender male voice (Constansis 2008). Researchers aren’t certain what causes entrapped vocality. Many vocal coaches agree with Constansis, who hypothesizes that it occurs when the vocal folds thicken faster than they lengthen. Zimman (2012) suggests that entrapped vocality is a result of the

voice's F0 lowering while the speaker maintains high vowel formants associated with individuals who are socialized as women. Regardless of what causes entrapped vocality, transmasculine individuals are acutely aware of the phenomenon. Many forum posts online ask what causes the “trans guy voice” and “what gives trans men’s voices that distinctive sound from cis men” (flannelonflannel n.d., McCallum n.d.). Transmasculine people recognize entrapped vocality in themselves and others.

As far as I am aware, no study has investigated transmasculine perceptions of entrapped vocality, and whether or not it is used to recognize other transmasculine individuals. This study proposes a three-part investigation to better understand the sociolinguistic effects and perceptions of this laryngeal phenomenon. The first part would be a listening task where participants of all gender identities hear recordings of the same phrase, where some voices are cisgender men and some voices are AFAB transgender individuals on HRT. Of the latter category, some voices would have entrapped vocality and some would not. Participants would be asked to rate the likelihood that each speaker was transgender. Afterwards, participants would complete a survey asking about their gender identity, the transgender individuals in their lives and their interactions with queer communities. Finally, transmasculine individuals who complete the task and survey would participate in sociolinguistic interviews, where they would be asked about their voices, their perceptions of others’ voices, and their thoughts on entrapped vocality.

I expect that, because transmasculine people recognize entrapped vocality, they will label the voices with entrapped vocality as transmasculine, and that cisgender individuals and transfeminine individuals may be less accurate in their labeling. I also expect that individuals who engage in queer communities or consume content from transmasculine creators may be more accurate in their labeling. This study aims to fill a gap in the research, not only by examining a voice quality and sociolinguistic variable that remains unstudied, but also by being one of the first studies to, prior to the administration of the listening task to participants, develop a criteria for identifying what is and is not entrapped vocality on a spectrogram. It would be incredibly worthwhile to the development of this project to receive advice from leading scholars in phonetics, sociolinguistics and trans linguistics through NWAV’s Project Launch program. On top of revising the project’s methodology, it would be useful to hear what phoneticians and scholars in trans linguistics would advise when identifying entrapped vocality.

Constansis, A. (2008). The Changing Female-to-Male (FTM) Voice. *Radical Musicology*, 3.

Darwin, H. (2017). Doing Gender Beyond the Binary: A Virtual Ethnography. *Symbolic Interaction*, 40(3), 317–334.

<https://doi.org/10.1002/symb.316>

flannelonflannel. (2019, August 2). The “trans guy voice” [Reddit Post]. R/FTMMen.

[www.reddit.com/r/FTMMen/comments/cky4yt/the\\_trans\\_guy\\_voice/](http://www.reddit.com/r/FTMMen/comments/cky4yt/the_trans_guy_voice/)

McCallum, Z. (n.d.). Zachary McCallum’s Answer to What gives trans men’s voices that distinctive sound from cis men? Quora. Retrieved May 28, 2023, from <https://www.quora.com/What-gives-trans-men-s-voices-that-distinctive-sound-from-cis-men>

Zimman, L. (2010). Female-to-Male Transsexuals and Gay-Sounding Voices: A Pilot Study. *Colorado Research in Linguistics*. <https://doi.org/10.33011/cril.22.1.3>

Zimman, L. (2012). *Voices in Transition: Testosterone, Transmasculinity, and the Gendered Voice among Female-to-Male Transgender People*. University of Chicago.

---

## Heritage Spanish writing complexity and its relationship to proficiency and writing genre

Alberta Gatti, Daniela Castillo, Syelle Graves

The Graduate Center, New York, USA

Heritage languages (HL) are characterized by variation across domains, and yet are traditionally compared to what the field calls non-heritage “baselines,” languages which have inherently less variation. This comparison can result in reducing HLs to “imperfect replicas of their source grammars” (Laleko & Kisselev, 2021, p. 2). An alternative approach is complexity theory, which analyzes language development as a dynamic system. Complexity research (e.g., Kisselev et al., 2021; Kuiken & Vedder, 2019; Park, 2017) has shown correlations between syntactic complexity and heritage speakers’ (HS) standardized proficiency levels (following ACTFL, American Council on the Teaching of Foreign Languages, 2012). However, no study we have seen has investigated the complexity-to-proficiency relationship while also controlling for the genre/level of the writing sample (i.e., the type of elicited writing). Finally, existing work on HS writing using the standardized ACTFL proficiency measure (Gatti & O’Neil, 2017; 2018) does not incorporate the more fine-grained complexity measure, creating two gaps in the literature.

This presentation shares results of a study on complexity and accuracy in the writing of Spanish heritage language learners (HSs enrolled in HL courses). We used a tool we modeled after the WPT (Writing Proficiency Test), which, following ACTFL standards, elicits four samples from each subject, at four levels of prompt difficulty: Intermediate, Advanced, combined Advanced/Superior, and Superior. Samples were double-rated by certified ACTFL raters, leading to a single ACTFL writing proficiency rating of each subject.

For this study, we analyzed the Intermediate-proficiency subjects: 92 total, 33 Intermediate Mid (IM), and 59 Intermediate High (IH). The first three of the four samples of each subject were then separately hand-coded by two L1 Spanish coders, who calculated syntactic complexity (finite clauses ÷ total t-units) and accuracy (error-free t-units ÷ total t-units). They made three passes to ensure intercoder reliability. The scores were averaged by writer proficiency level (IM and IH) and by prompt difficulty.

Results show a statistically significant relationship between complexity, accuracy, and proficiency sublevel (IM or IH), as well as between complexity, accuracy, and the difficulty level of the prompt that elicited the writing (a finding unattested in the literature).

\* An independent t-test was conducted on the mean complexity and accuracy rates between the IMs and IHs; see Tables 1 and 2.

\* A paired-samples t-test was conducted on the means of each of the three prompt samples, with all subjects (at both proficiency levels) combined.

\* All t-test results were significant at the .05 level, two-tailed.

First, both complexity and accuracy increased from IM to IH writers, something that speaks to the usefulness of the ACTFL proficiency scale. Second, across all subjects, as the prompt increased in difficulty, (1) the complexity score accordingly increased, while (2) the accuracy score decreased, something that speaks to the usefulness of complexity measures. Interestingly, this accuracy decrease illustrates “linguistic breakdown” (following ACTFL, 2012), known to occur when learners are pushed beyond their proficiency limits.

We tentatively conclude that establishing (1) standardized proficiency of subjects and (2) writing type is critical for accurately investigating complexity in language-learner writing.

Cole, M. (2021). *Sociolinguistics and Language Shift in Ecuador: A Case Study of Palatals*. [Dissertation, Indiana University]. ProQuest Dissertations and Theses Global.

García Mouton, P., & Molina Martos, I. (2012). The /k/-/j/ merger (yeísmo) in central Spain: advances since the ALPI. *Dialectologia*. Special issue, III, 23-42.

Godenzzi, J. C. (2013) Resistencias al yeísmo en los Andes: Convergencia lingüística y expresión identitaria. In Gómez, R., & Martos, I. M. (Eds.), *Variación yeísta en el mundo hispánico* (pp. 295-309). Iberoamericana Editorial Vervuert.

Gómez, R., & Martos, I. M. (Eds.). (2013). *Variación yeísta en el mundo hispánico*. Iberoamericana Editorial Vervuert.

Lipski, J. M. (1989). Spanish yeísmo and the palatal resonants: Towards a unified analysis. *Probus*, 1(2), 211-223.

Lipski, J. M. (1996). *El español de América*. Madrid: Cátedra.

Navarro, T. (1993). Nuevos datos sobre el yeísmo en España. In R. Páez Patiño & I. Chaves Cuevas (Eds.), *Thesaurus: Boletín del Instituto Caro y Cuervo: Muestra antológica 1945-1985*, I: Lingüística; II: Historia de la literatura, filología y análisis literario; III: Historia cultural, cultura popular, poesía latina, discursos. Inst. Caro y Cuervo.

Orzechowski Dias, E. C., & Seara, I. C. (2014). Descripción fonético-acústica de las variantes del yeísmo presentes en el español de rioplatenses y de colombianos de la región de la costa atlántica y del valle del cauca. *Arco/Libros*, 399-418.

Rost Bagudanch, A. (2013). La transcripción fonética en estudios dialectales: propuestas en el caso del yeísmo. *Revista de filología española*, 93(1), 165-192.

Rost Bagudanch, A. (2017). Variation and Phonological Change: The Case of yeísmo in Spanish. *Folia Linguistica: Acta Societatis Linguisticae Europaeae*, 51(1), 169–206.

---

## Variability of language evaluation in the American North & South as described by the other region

Patrick Gehringer  
University of Kentucky, Lexington, USA

Sociolinguistic research has conducted multiple investigations of “the South.” Some perceptual studies explore external imitations of the South (Evans 2002), while others look for the perceived geographic scope and social evaluation of the South (Zelinsky 1980, Preston 1996, 1997, 2018), demonstrating how language in the American South manifests and how it is evaluated, both internally (by Southerners) and externally (usually by Northerners). But there are few to no accounts of southern descriptions of the north, and it is unclear if “the North” has any consistently enregistered features or salient geographic bounds among southerners.

I surveyed 40 speakers of North American English, 33 of which self-identified as being “southern,” while the remainder of participants identified as “not-southern” for comparison. Data collection is ongoing. Participants were given 5 verbal response questions investigating their evaluations of the north and its language. From this I have established some commonalities among southerners’ definitions and evaluations of “the North” (forthcoming).

I am comparing these findings with previous sociolinguistic literature about northerners imitating or describing the South. I found that participants often described the northern sounding region to be smaller than the general geographic or cultural region of “the North”, which is consistent with Michiganders identifying a broad southern region with linguistic areas of interest (Preston 1996), or with the discrepancies between phonetic surveys and how people classify areas like Texas with the South (Ibid., Zelinsky 1980, Labov, Ash, Boberg 2006, Preston 2018).

I then compared the evaluations of northerners by the southerners in my study to those of the opposite direction in other literature, e.g. the Michiganders in Preston 1999. These Michiganders define their local north in terms of accentless prestige before focusing on solidarity terms such as “friendly,” “casual,” or “down-to-earth,” mixed with positive evaluations of their own intelligence. These same solidarity categories dominate their evaluations of Southern American English, followed by a description of slow, accented speech spoken by nice but unintelligent people. The southerners of my study evaluated northerners as accented and unfriendly, and brusquely formal, with few comments about prestige. When prestige is mentioned, several express a belief that “northern correctness” is artificial, while southern speech has a more natural state. There are also several comments about northerner’s non-southernness, indicated by comments about urbanity, individualism, and foreignness. This description is essentially the inverse of the northern description of the south (but makes little comment on intelligence), and is different from the Michiganders’ self description, who, while putting less emphasis on being friendly and down-to-earth had still believed that their language had those attributes after celebrating their self proclaimed language prestige and correctness.

In conclusion, while northerners at the turn of millennium saw themselves almost as approachable as their accented and dumber southern counterparts, today’s southerners seem to see themselves linguistically and culturally on par with northern intelligence and prestige, while simultaneously viewing northern language and society as a separate or even foreign entity to their own, where the people and their speech are as quick and harsh.

- Agha, Asif. 2005. Voice, Footing, Enregisterment. *Journal of Linguistic Anthropology*. (15.1), 38-59.
- Evans, Betsy E. 2002. An Acoustic and Perceptual Analysis of Imitation. *Handbook of Perceptual Dialectology*. Volume 2.
- Johnstone, Barbara, Jennifer Andrus, Andrew E. Danielson. 2016. Mobility, Indexicality, and the Enregisterment of “Pittsburghese.” *Journal of English Linguistics*. (34.2), 77-104.
- Labov, William, Sharon Ash, and Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter.
- Preston, Dennis R. 1996. Where the worst English is spoken. In *Focus on the USA*, edited by Edgar Schneider, 297–360. Cambridge: Cambridge University Press.
- Preston, Dennis R. 1997. The South: The touchstone. In *Language variety in the South revisited*, edited by Cynthia Bernstein, Tom Nunnally, and Robin Sabino, 311–51. Tuscaloosa: University of Alabama Press
- Preston, Dennis R. 1999. A language attitude approach to the perception of regional variety. In *Handbook of perceptual dialectology*, vol. 1, edited by Dennis R. Preston, 359–73. Amsterdam: John Benjamins.
- Preston, Dennis R. 2018. [ˈsʌdðə-n], [ˈsʌdən], [ˈsʌdɪən], [ˈsʌdŋ], [sʌːn], etc. *What We/They Think/Thought It Is/Was/Will Be*.
- Zelinsky, Wilbur. 1980. North America’s vernacular regions. *Annals of the Association of American Geographers*. 70.1–16.

---

## Regional Variation of Phonological Features in African American Language

Anissa Gladney

University of California, Los Angeles, Los Angeles, USA

To determine how features of African American Language (AAL) differ cross regionally, this research will analyze the rate of four features in the 6 locations available in the Corpus of Regional African American Language (CORAAL; Kendall & Farrington 2020). In recent years, research has been done to investigate regional differences in AAL and whether there has been oversimplification in the discussion and analysis of the variety (Wolfram 2007; Jones 2015). Certain AAL dialect markers were thought to override the kinds of regional boundaries associated with Mainstream American English (MAE) dialects (Wolfram & Schilling-Estes 2015: 232). Because of this, this research will also investigate the ways in which AAL regional differentiation may pattern to MAE regional boundaries as given by Labov (Labov et al. 2008).

African American Language refers to the set of English varieties that African Americans - as the descendants of U.S. chattel slavery - speak (Lanehart & Malik 2015: 3). This umbrella term includes standard and vernacular forms as well as those used in specific regions or distinct groups within the African American community (Wilkerson 2015: 214). There are many points of contention within AAL research, but the focus of this research is the uniformity hypothesis (Thomas 2007; Green 2002; King 2020).

To legitimize AAL as a dialect, researchers centered the features that were the most markedly different from MAE, but this “reduced the complexity...to a subset of the entire community” (Bucholtz 2003: 402). This simplification was integral in legitimating the dialect, but brought about the prevalence of the uniformity hypothesis. According to the uniformity hypothesis, African American speakers in the U.S., regardless of region, speak the same prototypical version of AAL (Labov et al. 1968; Legum et al. 1971; Labov 1972; Baugh 1983; Wolfram & Schilling-Estes 2015). Previous research focused mostly on working-class, young men, and this research has shaped what is considered AAL and who speaks it (Labov 1972). The majority of research on AAL assumes uniformity without questioning its validity.

Preliminary results have focused on /l/-vocalization which shows statistically significant different distribution in two of the locations. The regression results shown in Figure 1 suggest that given all data, speakers in Princeville are statistically significantly more likely to vocalize /l/ than speakers in Washington D.C.

In addition to the different rates of /l/-vocalization in those locations as well as differences in phonological conditioning. Most notably, there is a significantly higher amount of /l/-vocalization in onset position in Princeville, North Carolina (Rowe et al. 2018) in comparison to Washington D.C (Kendall et al. 2018).

The regression in Figure 2 shows that there is a statistically higher amount of /l/-vocalization in coda, onset, and complex onset position in Princeville relative to Washington D.C. This highlights the need to look more closely at onset position as a location for /l/-vocalization as it has been overlooked in previous work on /l/-vocalization in AAL (van Hofwegen 2010). This also suggests a difference in /l/-velarization that appears to differ regionally as well.

- Baugh, J. (1983). *Black Street Speech: its history, structure and survival*. Austin: University of Texas Press.
- Bucholtz M. (2003). Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics*, 7(3): 398–416
- Green, L. (2002). *African American English: A Linguistic Introduction*. New York: Cambridge University Press.
- Jones, T. (2015). Toward a Description of African American Vernacular English Dialect Regions using “Black Twitter”. *American Speech*, 90(4): 403-40.
- Kendall, T., & Farrington, C. (2020). *The Corpus of Regional African American Language*. Version 2020.05. Eugene, OR: The Online Resources for African American Language Project. [ <http://oraal.uoregon.edu/coraal> ]
- Kendall, T., Quartey, M., Farrington, C., McLarty, J., Arnson, S., & Josler, B. (2018). *The Corpus of Regional African American Language: DCB (Washington DC 2016)*. Version 2018.10.06. Eugene, OR: The Online Resources for African American Language Project.
- King, S. (2020). From African American Vernacular English to African American Language: Rethinking the Study of Race and Language in African Americans’ Speech. *Annual Review of Linguistics*, 6(1), 285–300.
- Labov, W. (1972). *Language in the Inner City: Studies in the black english vernacular*. University of Pennsylvania Press.
- Labov, W., Ash, S., & Boberg, C. (2008). *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Mouton de Gruyter.
- Labov, W., Cohen, P., Robins, C. & Lewis, J. (1968). *A Study of the non-standard English of Negro and Puerto Rican speakers in New York City. Volume I: Phonological and Grammatical Analysis: The use of language in the speech community*. U.S. Department of Health, Education and Welfare.
- Lanehart, S. L., & Malik, A. (2015). “Language Use in African American Communities: An Introduction.” *Oxford Handbook of African American Language*. Ed. Sonja Lanehart. Oxford: Oxford University Press, 1-19.
- Legum, S., Pfaff, C., Tinnie, G., & Nichols, M. (1971). *The speech of young black children in Los Angeles*. Inglewood, CA: Southwest Regional Laboratory.
- Rowe, R., Wolfram, W., Kendall, T., Farrington, C., & Josler, B. (2018). *The Corpus of Regional African American Language: PRV (Princeville, NC 2004)*. Version 2018.10.06. Eugene, OR: The Online Resources for African American Language.
- Thomas, E. R. (2007). Phonological and Phonetic Characteristics of African American Vernacular English. *Language and Linguistics Compass*, 1(5), 450–75.
- van Hofwegen, J. (2010). Apparent-time evolution of /l/ in one African American community. *Language Variation and Change*, 22(3), 373-396.
- Wilkerson, R. (2015). “African American English in the Mississippi Delta.” *Oxford Handbook of African American Language*. Ed. Sonja Lanehart. Oxford: Oxford University Press, 201-218.
- Wolfram W. (2007). Sociolinguistic folklore in the study of African American English. *Language and Linguistics Compass*, 1(4):292–313.
- Wolfram, W., & Schilling-Estes, N. (2015). *American English: Dialects and Variation*, 3rd Edition. John Wiley & Sons.

---

## Investigating lexical variation in a mixed language: A variationist corpus-based analysis of Lánnang-uè conjunctions and prepositions

Wilkinson Daniel Wong Gonzales

The Chinese University of Hong Kong, Hong Kong, Hong Kong

Over the course of many decades, numerous studies have consistently demonstrated the significance of sociolinguistic factors in shaping language variation, including the variations observed in multilingual practices (Weinreich et al. 1968; Eckert 2005). This fact is reinforced by the abundance of recent research conducted worldwide. However, there is still much that remains unknown regarding the impact of sociolinguistic factors on multilingual practices specifically within the Philippines. For example, do the factors that constrain variation in multilingual practices in other parts of the world apply to languages from this region?

In this presentation, I analyze the variation in a low-resource, previously undocumented mixed language in the Philippines called Lánnang-uè – a variety that systematically derives linguistic elements from Hokkien (Southern Min), Mandarin, English, and Tagalog. Specifically, I focus on the patterns of variation in two lexical categories: conjunction and prepositions – two categories that show higher rates of variation compared to other features in

Lánnang-uè. Using a mix of quantitative (i.e., corpus-based, computational) and qualitative (i.e., ethnographic) approaches, this analysis investigates the impact of four factors – age, sex, self-reported language proficiency in the source languages, and language attitudes – on the variation observed. I pre-processed, machine-tagged, and statistically analyzed conjunction and preposition data from the Lannang Corpus (LanCorp) – a self-compiled 375,000-word corpus of Lánnang-uè, acquired from 135 Lánnang-uè speakers. I also analyzed metalinguistic commentary from these speakers in an attempt to provide a more holistic explanation for potential sociolinguistic patterns.

The findings indicate that variation in the use of conjunctions and prepositions can be explained by at least one of the four enumerated sociolinguistic factors, corroborating previous work on Lánnang-uè and other research on related contact varieties in East Asia (Hansen Edwards 2018; Starr and Balasubramaniam 2019; Lee 2014). However, I also found that the effects of age, sex, language proficiency, and attitudes varied depending on many context-specific factors (e.g., degree of awareness, stylistic practices unique to a particular social group) (Dodsworth 2005). An instance of this can be observed in my findings, where it became evident that the sociolinguistic factors under scrutiny solely influence the variation seen in conjunctions, while having no discernible impact on prepositions. This discrepancy could potentially be attributed to Lannang language users possessing an awareness of the variation associated with conjunctions and their social meanings (such as local-ness), as supported by various metalinguistic comments made by speakers in the study (see examples below). However, they may not possess the same level of awareness regarding variation in prepositions, as evident in the absence of preposition-related metalinguistic comments across all speakers.

O, dí khuà, ‘pero’ nanamân.

“Here, look, pero ‘but’ again.” <CLIN-19-117:20240>

Dî buetsuê kong sêh “KASÍ guâ ti tsiá”.

“You can’t just say KASI gua ti tsia ‘BECAUSE I am here’ (to the Mainlanders).” <CLIN-19-117:20325>

I discuss the sociolinguistic patterns uncovered in my presentation in light of cognitive, sociolinguistic, and contact linguistics theories, and conclude by briefly identifying potential avenues for future research.

Dodsworth, Robin M. 2005. Linguistic variation and sociological consciousness. Columbus, Ohio: The Ohio State University ph.d. dissertation.

Eckert, Penelope. 2005. Variation, convention, and social meaning. Paper presented at the Annual Meeting of the Linguistic Society of America, Oakland, CA.

Hansen Edwards, Jette G. 2018. TH variation in Hong Kong English. *English Language and Linguistics* 23.439–468. doi:10.1017/S1360674318000035.

Lee, Nala Huiying. 2014. A grammar of Baba Malay with sociophonetics considerations. Honolulu, HI: University of Hawai’i at Mānoa ph.d. dissertation.

Starr, Rebecca Lurie.; and Brinda Balasubramaniam. 2019. Variation and change in English /r/ among Tamil Indian Singaporeans. *World Englishes* 38.630–643. doi:10.1111/weng.12357.

Weinreich, Uriel.; William Labov.; and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. *Directions for historical linguistics*, ed. by Winfred P. Lehmann and Yakov Malkiel, 100. Austin: University of Texas Press.

---

## Ageing in style: Towards disentangling style shifting and lifespan change

James Grama, Anne-Marie Moelders, Lea Bauernfeind, Isabelle Buchstaller  
Sociolinguistics Lab, University of Duisburg-Essen, Essen, Germany

Lifespan change and style shifting do not occur independently (Buchstaller 2015; Wagner 2021). Instead, speakers modulate their linguistic behaviour according to considerations that are situated at the intersection of cohort and lifespan-specific marketplaces (see Rickford and Price 2013; Tetreault 2017; Riverain-Coutlee and Harrington to appear). Research suggests that stylistic presentation of self is contingent on a wealth of factors that intersect with speaker age. These have been shown to include the conversational topic and the type of speech activity (Labov 1972; Rickford and McNair-Knox 1994), accommodation to interlocutors, cohort and community (e.g., Bell 1984; Tetreault 2017; Wagner 2021), as well as the speaker’s orientation to the social meaning of linguistic resources (see papers in Eckert and Rickford 2002). A persistent challenge for panel research is how to disentangle bona fide intra-speaker change from momentary stylistic shifts in linguistic choices, due in large part to a lack of larger panel datasets that allow us to unpack stylistic effects.

This talk addresses the relationship between stylistic shifts and intra-speaker change by drawing on a unique panel corpus. The data consists of 17 English speakers from the North-East of England, aged between 27 and 53 at the time of first recording, and recorded again between five and seven years later. We investigate variability in two

variables differentiated by level of linguistic structure and socio-indexical meaning: the face vowel and the first-person possessive (Ipos). To unpack the effect of style shifting, each recording includes a comparable range of socio-situational contexts. Stylistic variation is conceptualised along three different dimensions: topic (an axis of careful-casual; see Rickford and McNair-Knox 1994; Labov 2001), addressee (following Bell's 1984 audience design model), and elicitation method (read vs. spontaneous speech, a la Labov 2001). Operationalising these different measures allows us to develop a multifaceted model of style shifting, and to hone in on the relationship between synchronic variation and diachronic change.

Analyses reveal a complex relationship between ageing and style. First, the range of stylistic shifts across all dimensions is smaller than intra-speaker change across time. Second, a single cohort – the 30s age group – exhibits the most stylistic variation in both variables, albeit in opposing directions, with face undergoing lifespan change, and Ipos undergoing retrograde change. By contrast, older speakers (40 and above) seem to have developed relatively stable stylistic ranges with more limited shifts across topic, addressee, and elicitation method. For the oldest 50+ speakers, stylistic shifts are only evident in stigmatized Ipos when talking with an unfamiliar interlocutor, the most formal of the contexts considered. These findings are discussed in terms of extant models of variation and change across the lifespan that rely on marketplace pressures (cf. Bourdieu & Boltanski 1975). Our analysis paints a differentiated picture of the relationship between momentary style shifts and more permanent intra-speaker change across the entire adult lifespan. By honing in on the relationship between variation (at one time point) and change (between time points), the present study informs models of linguistic lability (see Rickford 2021; Wagner 2021).

- Bell, Alan. 1984. "Language Style as Audience Design". In *Sociolinguistics: A Reader and Coursebook*, N. Coupland and A. Jaworski (eds). New York: St Martin's Press Inc. 240–250.
- Bourdieu, Pierre, & Boltanski, Luc. 1975. "Le fétichisme de la langue." *Actes de la recherche en sciences sociales* 1(4):2–32.
- Buchstaller, Isabelle. 2015. "Exploring linguistic malleability across the life span: Age-specific patterns in quotative use." *Language in Society* 44(4):457–496.
- Eckert, Penelope. 2019. "The individual in the semiotic landscape." *Glossa: A Journal of General Linguistics* 4(1):14.
- Eckert, Penelope, & Rickford, John (eds.). 2002. *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 2001. The anatomy of style-shifting. In *Style and Sociolinguistic Variation*, P. Eckert and J. Rickford (eds). Cambridge: Cambridge University Press. 85–108.
- Rickford, John, & McNair-Knox, Faye. 1994. "Addressee- and Topic-Influenced Style Shift: A Quantitative Socio-Linguistic Study". In *Perspectives on Register: Situating Register Variation within Sociolinguistics*, D. Biber and E. Finegan (eds). Oxford: Oxford University Press. 235–276.
- Rickford, John, & Price, Mackenzie. 2013. "Girlz II women: Age-grading, language change, and stylistic variation." *Journal of Sociolinguistics* 17(2):143–179.
- Rickford, John. 2021. "Stylistic Variation in Panel Studies of Language Change". Karen V. Beaman, Isabelle Buchstaller (eds.), *Language Variation and Language Change Across the Lifespan: Theoretical and Empirical Perspectives from Panel Studies*. New York and London: Routledge, 77-97.
- Tetreault, Chantal. 2017. "Ethnographic Perspectives on Panel Studies and Longitudinal Research." In *Panel Studies of Variation and Change*, S.E. Wagner and I. Buchstaller (eds). New York: Routledge. 235–255.
- Wagner, Suzanne Evans. 2021. "What's the point of panel studies?" In *Language Variation and Language Change Across the Lifespan: Theoretical and Empirical Perspectives from Panel Studies*, K. Beaman and I. Buchstaller (eds). New York and London: Routledge. 267–280.

---

## Vowels, affect, and (dis)affiliation in a non-binary community

Chantal Gratton

Stanford, Stanford, USA. Queen Mary University of London, London, United Kingdom

Holistic changes in the vowel space can index things like regional character types (Pratt and D'Onofrio 2017) and how speakers feel about an interlocutor or an interaction (this author), such as feeling like they 'click' with each other. However, it remains unclear whether the degree of centralization of one's vowel space (vocalic peripherality) is a resource that speakers dynamically vary throughout an interaction. Building on discourse analytic work on common ground (e.g., Fox Tree and Schrock 2002; Sebba and Tate 1986), affect (e.g., Hoey 2014; Koshik 2003), and alignment (e.g., Heinemann 2008; Reed 2020), I argue that vocalic peripherality is yet another resource speakers make use of in interaction to position themselves relative to interlocutors and take evaluative affective stances.

I focus on a close-knit community of non-binary individuals and how their vocalic realizations within and across dyadic interactions with a cis-gender friend and then a non-binary one cue interpersonal and affective meaning. Overall vocalic patterns are analyzed using mixed-effect linear regression models. Models for peripherality use Euclidean distance between DeltaF (Johnson 2020) normalized vowel tokens and their respective speaker's centroid as the dependent variable (N=7,388). Additionally, models using raw hertz values of F1 or F2 are applied for each vowel class to determine their degree of participation. For all models, the fixed effect independent variables include



log duration, alignment, and an interaction between affect and interlocutor type (cis friend, enby friend). Speaker, vowel class (only in the Euclidean distance model), word, preceding segment, and following segment are treated as random intercepts. By combining quantitative analyses of vocalic variation with stance coding and a discourse analytic approach, this study develops a method for capturing moment-to-moment vocalic variation and its interactional implications.

The results reveal a significant interaction between evaluative affect and interaction type ( $\beta = -0.035$ ,  $p < .05$ ). When interacting with cis-gender friends, speakers tend to use more dispersed vowels when expressing negative stances, whereas more centralized vowels are employed with enby friends. Moreover, speakers utilize more dispersed vowels for utterances involving diverging alignment compared to converging alignment ( $\beta = 0.016$ ,  $p < .001$ ). The study's mixed-methods approach further uncovers two distinct vocalic variation patterns. Centralization is employed when building solidarity and commiserating with other non-binary individuals, emphasizing the sharedness of important and emotional experiences. Dispersion, on the other hand, serves primarily to mark negative affective stances towards personally and emotionally significant topics, as well as to disaffiliate from cis-gender friends during disagreements. Notably, extreme dispersion signals particularly divergent and negative stances towards cis-normative beliefs and cis-friends' ignorance of the hardships faced by transgender individuals.

I propose that through shifts in vocalic centralization and dispersion, speakers can index affiliation with or disaffiliation from an interlocutor, respectively. Together, these findings suggest that vocalic peripherality is indeed a resource that speakers use in interaction to shape their interpersonal relationships and convey affective experiences. This raises important questions regarding the relationship between vocalic variation and the more emotionally embodied facets of affect and positioning.

- Fox Tree, Jean E, & Schrock, Josef C. 2002. Basic meanings of you know and I mean. *Journal of Pragmatics*, 34 (6), 727–747.  
 Heinemann, Trine. 2008. Questions of accountability: yes–no interrogatives that are unanswerable. *Discourse Studies*, 10 (1), 55–71.  
 Hoey, Elliott M. 2014. Sighing in interaction: Somatic, semiotic, and social. *Research on Language and Social Interaction*, 47 (2), 175–200.  
 Koshik, Irene. 2003. Wh-questions used as challenges. *Discourse Studies*, 5 (1), 51–77.  
 Pratt, Teresa, & D'Onofrio, Annette. 2017. Jaw setting and the California Vowel Shift in parodic performance. *Language in Society*, 46 (3), 283–312  
 Reed, Beatrice Szczepek. 2020. Reconceptualizing mirroring: Sound imitation and rapport in naturally occurring interaction. *Journal of Pragmatics*, 167, 131–151  
 Sebba, Mark, & Tate, Shirley. 1986. You know what I mean? Agreement marking in British Black English. *Journal of Pragmatics*, 10 (2), 163–172.

## A layering approach to DP2 agreement

Brian Gravely<sup>1,2</sup>, Daniel Greeson<sup>3</sup>

<sup>1</sup>Emory University, Atlanta, USA. <sup>2</sup>University of Southern Maine, Portland, USA. <sup>3</sup>Stony Brook University, Stony Brook, USA

**INTRO:** In this abstract, we present a layering approach to verbal agreement in (pseudo-)partitive constructions in Spanish of the type [DP1 [*de* DP2]] ('DP1 of DP2') in which To agrees with DP2, not DP1 (often called *ad sensum* agreement). In a survey of  $n=117$  native Spanish speakers, we found that some common lexical items commonly used in the DP1 slot have all but grammaticalized, akin to *a lot* in English, showing grammaticality only when DP2 takes agreement on the verb. We claim that this pattern is possible in a *layering* approach (also called *remerge*, *renumeration*; cf. Thoms (2019) and Johnson (2003), respectively) in which only DP2 is merged in the base-generated position of the subject and able to agree with To, after which it receives additional structure as the derivation progresses.

**OVERVIEW:** *Ad sensum* agreement has often been considered a semantic-based agreement phenomenon in which DP1 of the structure [DP1 [*of* DP2]] serves as a quantifier-type element that has been semantically bleached (Brems 2008, Traugott 2008).

(1) **Alot / heap / ton** of kids **\*is / are** out of school today

In Spanish, significant variability has been cited with respect to verbal agreement of these partitive constructions, where the verb may bear singular or plural agreement (2) (Real Academia Española 2009, Pérez-Jiménez & Demonte 2017, Gupton & Howe 2023). This may be seen with common DP1 candidates, both definites and indefinites, such as *un montón* ('a mountain'), *la mayoría* ('the majority'), and *un puñado* ('a fist').

(2) Un <b>montón</b> de personas <b>marchó/</b>	<b>marcharon</b> temprano
<b>a mountain</b> of people.PL <b>leave.PST.3SG</b>	<b>leave.PST.3PL</b> early
<b>La mayoría</b> de sus amigos no la <b>conocía /</b>	<b>conocían</b> bien
<b>the majority</b> of her friends NEG CL <b>know.IMPF.3SG</b>	<b>know.IMPF.3PL</b> well

‘The majority of her friends didn’t know her well.’

**Un puñado** de alumnos **suspende** / **suspenden** el primer examen  
**a fist** of students **fail.PRS.3SG fail.PRS.3PL**  
the first exam ‘A bunch of students fail the first exam.’

*Ad sensum* agreement is distinct from ‘agreement attraction’ (Wagers et al. 2009) in the sense that the former is a viable grammatical output, whereas the latter is considered a form of erroneous agreement during processing. Moreover, ‘agreement attraction’ phenomena rely on linear adjacency between the verb and the agreeing DP (3a), whereas *ad sensum* occurs even in constructions in which these two syntactic elements are not linearly adjacent to one another (3b).

(3) \*The way to [their hearts] are through food  
**Comieron** la tarta **un montón** de [los chicos] de la fiesta  
**eat.PST.3PL** the cake **a mountain** of the kids of the party  
‘A bunch of kids from the party ate the cake.’

**DATA:** We surveyed  $n=117$  native Spanish speakers from Spain, Peru, and Mexico (age range 23- 40 years old) in a 40-question grammaticality judgement task where they were asked to rate each construction via a 5-point Likert scale (‘bad’ to ‘good’). We tested four ( $n=4$ ) common DP1 candidates for their involvement in agreement: *un montón* (‘a mountain’), *la mayoría* (‘the majority’), *un puñado* (‘a first’), and *un huevo* (‘an egg’). We tested for a combination of three syntactic aspects: (i) preverbal vs. postverbal position of the target subject, (ii) transitive vs. intransitive verbal predicates, (iii) partitive vs. pseudo-partitive target subjects. The results of our findings were as follows. When selecting a plural DP2 resulting in *plural agreement* on the verb, there was no significant difference in grammaticality based on position of the targeted subject (preverbal score=4.957, postverbal score=4.786), the transitivity of the verb (intransitive score=4.829, postverbal score=4.573), or the type of partitive complement (pseudo-partitive score=4.743, partitive score=4.957). Moreover, all DP1 candidates resulted in scores that reflect strong acceptability judgements. In contrast, when selecting a plural DP2 resulting in *singular agreement* on the verb, these syntactic contexts had average scores of 1.214 for preverbal and 1.239 for postverbal subjects, 1.186 for transitive and 1.197 for intransitives, and 1.248 for pseudo-partitives and 1.145 for partitives. Unlike what has been claimed in the literature (Real Academia Española 2009, Pérez-Jiménez & Demonte 2017, Gupton & Howe 2023), these results show that there is a group of DP1 candidates in Spanish that no longer permit singular agreement when DP2 is plural regardless of any of the aforementioned syntactic conditions.

**THEORETICAL PROPOSAL:** In order to account for constructions in which DP2 bears agreement in these (pseudo-)partitive constructions, we claim that DP2 is merged alone in the corresponding base-generated position (Spec,vo for transitives; complement of Vo/ $\sqrt{ROOT}$  in unaccusatives). As in Thoms (2019a,b) and Thoms & Heycock (2022), only a portion of the spelled-out subject is present when To is merged (in our case, DP2; in their non-partitive cases, *nP*). When the subject moves to Spec,To (cf. 2), we claim that DP2 undergoes a layering process in which DP1 and its prepositional complement (*de*) are merged *after* the probe on To has scanned its c-command domain and agreed with DP2. Based on the data shown above, this ensures plural agreement when DP2 is plural.

For situations of DP2 agreement from postverbal subjects, we claim that there are two possible avenues. The first assumes that all postverbal subjects move to a designated position between *vP* and *TP*, what Ordóñez (2007) labeled *SubjP*. Upon moving to the specifier of this projection, To is merged and finds DP2 in its search space. Upon agreeing with it, the additional structure of DP1 plus the preposition is merged via sideward movement (Nunes 1995, de Vries 2009, a.o.). As in the preverbal-subject constructions, the  $\phi$ -features of DP1 are never a target for agreement on To. A second possibility follows the motivation for Thoms’ layering approach, the *renumeration* procedure as described in Johnson (2003). Johnson claims that a syntactic constituent, upon creation, may be *renumerated* or put back through the Numeration of a given Lexical Array (Chomsky 2000) and re-inserted into the derivation as a new syntactic unit. Specific to Johnson’s system of renumeration is the notion that renumerated constituents need not move. One key to his proposal, however, is that sub-extraction from renumerated constituents is illicit, a fact that is borne out in the case of (pseudo-)partitives:

(4) \*[De cuántas personas]<sub>i</sub> estaban en la fiesta [un montón  $t_i$ ]<sub>j</sub>?  
of how.many persons be.IMPF.3PL in the party a mountain  
Intended: ‘How many people were there a bunch of at the party?’

Although not present in the survey detailed above, the effects of these constructions predict not only the distinction between singular and plural but also 1<sup>st</sup>- and 2<sup>nd</sup>-person DP2s from 3<sup>rd</sup>-person DP2s as in (5).

(5) [Un montón de **nosotros**] **fuimos** / \*fue a ver-lo  
a mountain of **we go.PST.1PL** go.PST.3SG to see.INF-CL  
‘A bunch of us went to see him.’

- Johnson, K. 2003. Towards an Etiology of Adjunct Islands. *Nordlyd* 31(1).
- Ordóñez, F. 2007. Cartography of postverbal subjects in Spanish and Catalan. In S. Baauw, M. Pinto, & F. Drijkoningen (eds.) *Romance Languages and Linguistic Theory 2005*. Amsterdam: John Benjamins.
- Thoms, G. 2019a. Antireconstruction as layering. In *Proceedings of NELS 49*, Cornell University.

---

## Caregiver Narratives: Variation in the Input and Child African American Language

Lisa Green, Chloe Ostiguy, Harmony Donald  
University of Massachusetts Amherst, Amherst, USA

This study investigates narratives produced by adults in an African American Language (AAL)-speaking community in the Southern US as a means of contributing to research on the relationship between caregiver input and child AAL. Narratives produced from sentence repetition and story retell were collected from 36 three- to five-year-old children and their caregivers from 25 to 69 years of age. In separate sessions, children and caregivers were instructed to repeat lines of a story produced by an experimenter while viewing pictures illustrating the events. All 60 main verbs in the story were inflected with third person singular marking, including equal representation for allomorphs [s] (e.g., sticks), [z] (e.g., stays), and [ɪz] (e.g., wishes). After the repetition task, the participants were instructed to retell the story, keeping their stories as close as possible to the original story they repeated. While there was low-level variation in verb endings in the repetition task, such that caretakers repeated 94% third person singular inflected verb forms (e.g., she gets/goes/washes) and used only 6% unmarked forms (e.g., she get), there was high-level variation in the retell task, in which caretakers produced 27% overtly marked 3rd person singular verbs, 32% unmarked verbs, and 41% verbs with past morphology (e.g., went, knocked) (See examples 1-5). Comparing these results to the child data, we found that the forms produced by the three- to five-year-old AAL-speaking children in the community were not carbon copies of those of the adults in that the child speakers produced over 50% unmarked verbs in the repetition task (Newkirk-Turner and Green 2016).

The discourse and aspectual properties of caregiver narrative retells were also analyzed. Smith (2003) explained that as narrative text progresses, narrative time advances with bounded situations. In addition, sentences in narratives express bounded events with an event verb constellation and the perfective aspectual viewpoint conveyed by the simple PAST verb form. The caregivers used verbs with variable morphological marking, not just PAST morphology, to move narratives along. Caregivers also drew on different rhetorical strategies and tense/aspect marking, such as past perfect resultant state (e.g., had dən (6)), that did not move the narratives along but served to elaborate on events in the past before the narrative past.

In recent research on child AAL, Kohn, Wolfram, Farrington, Renn, and Van Hofwegan (2021) concluded that mothers' and children's vernacularity scores differ in children's early adolescent years but are statistically similar at all other points across the lifespan. There continues to be more research on school age and adolescent AAL speakers; however, child AAL from birth to five years is understudied, especially from the perspective of caregiver input and the development of variation and the acquisition path. The narrative data in this research goes beyond addressing frequency of production of variable morphological forms and considers discourse and rhetorical strategies that provide insight into caregiver language and target AAL that serve as input for the acquisition and development of variation in morphological marking and tense/aspect properties of the linguistic variety.

Kohn, Mary, Walt Wolfram, Charlie Farrington, Jennifer Renn, and Janneke Van Hofwegan. 2021. *African American Language: Language Development from Infancy to Adulthood*. Cambridge: Cambridge University Press.

Newkirk-Turner, Brandi and Lisa Green. 2016. Third Person Singular -s and Event Marking in Child African American English, *Linguistic Variation* 16, 103-130.

Smith, Carlota. 2003. *Modes of Discourse: The Local Structure of Texts*. Cambridge: Cambridge University Press.

---

## Developing a child-centred approach to understand children's explicit sociolinguistic awareness: how 9-11 year old children in a London suburb perceive social meaning in talk.

Ella Gregory, Bronwen G Evans  
University College London, London, United Kingdom

This paper presents findings from an ongoing project investigating how children develop the ability to extract socio-cultural meaning from talk. We focus on children aged 9-11 yrs, all attending the same local primary school in Hampton, a suburban area of London where the population are largely speakers of Southern Standard British English (SSBE). Of particular interest is whether more can be understood about children's sociolinguistic awareness at this crucial lifestage – the transition from primary to secondary school - using child-friendly approaches (Clark & Moss, 2017; O'Kane, 2008).

In Study 1, 11 children (3 female; 2 bilingual) completed semi-structured sociolinguistic interviews. Children listened to short audio clips of four talkers. Three had accents that are regularly heard in the local community; SSBE, Indian English, and Popular London English (PLE). The other was a speaker of Multicultural London English (MLE), which is heard in London more broadly. Discussion was organized around questions such as ‘Where is this person from?’, ‘What job might they have?’ and ‘What might they be like?’. Thematic analysis (Braun and Clarke, 2006) revealed that children consistently associated certain indexical meanings with the most familiar accents, SSBE and Indian English. For Indian English in particular, children consistently identified a persona which reflected not only their personal experience but also the cultural stereotypes associated with this variety (e.g. “maybe he works in one of those corner shops”). For MLE and PLE however, the children’s ideas were more varied and overall their responses were more uncertain

These results are consistent with the idea that children rely on personal experiences when identifying accents and forming judgments about unfamiliar speakers due to having less experience with different talkers and varieties (Paquette-Smith et al., 2019). However, it is possible that by using an approach primarily designed for adults, we may have underestimated children’s knowledge. Consequently, in Study 2, we used an ethnographic approach alongside a new, child-centred task (Clark & Moss, 2017). Following 6 weeks of informal fieldwork, sociolinguistic interviews were conducted with 42 children (9-11 years, 20 female, 9 bilingual) in groups of 3. Prior to the interview, children completed a task at home (cf. Llamas, 1999). Children were asked to listen to an audio clip of a speaker of one of the accents (SSBE, PLE, or MLE; one each per group) and to draw or write their thoughts about the speaker. They were asked to bring this with them to the interview where it formed the basis of the discussion.

A thematic analysis uncovered similar patterns to those from Study 1: children had the most consistent responses for SSBE, the most familiar accent, while responses for PLE and MLE were more varied. However, responses for PLE and MLE were more detailed than in Study 1, illustrating a more in-depth knowledge of cultural stereotypes (e.g., “single mum”). These findings mirror adult behaviours more closely than previous research indicates (Kinzler and DeJesus, 2013) and demonstrate the importance of incorporating qualitative, child-friendly approaches in sociolinguistic work with children.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

Clark, A., & Moss, P. (2017). *Listening to Young Children, Expanded Third Edition: A Guide to Understanding and Using the Mosaic Approach*. Jessica Kingsley Publishers. <http://ebookcentral.proquest.com/lib/ucl/detail.action?docID=4784949>

Kinzler, K. D., & DeJesus, J. M. (2013). Northern = smart and Southern = nice: The development of accent attitudes in the United States. *Quarterly Journal of Experimental Psychology*, 66(6), 1146–1158. <https://doi.org/10.1080/17470218.2012.731695>

Llamas, Carmen (1999). A new methodology: Data elicitation for social and regional language variation studies. *Leeds Working Papers in Linguistics and Phonetics* 7:95–118 [https://www.latl.leeds.ac.uk/wp-content/uploads/sites/49/2019/05/llamas\\_1999.pdf](https://www.latl.leeds.ac.uk/wp-content/uploads/sites/49/2019/05/llamas_1999.pdf)

O’Kane, C. (2008). The Development of Participatory Techniques: Facilitating Children’s Views about Decisions Which Affect Them. In *Research With Children* (2nd ed.). Routledge.

Paquette-Smith, M., Buckler, H., White, K. S., Choi, J., & Johnson, E. K. (2019). The effect of accent exposure on children’s sociolinguistic evaluation of peers. *Developmental Psychology*, 55(4), 809–822. <https://doi.org/10.1037/dev0000659>

---

## Exploring the Deficiencies of Large Language Models in Using AAL: An Interdisciplinary Study

Jessi Grieser<sup>1</sup>, Nicholas Deas<sup>2</sup>, James Shepard<sup>3</sup>, Shana Kleiner<sup>4</sup>, Desmond Patton<sup>4</sup>, Elsbeth Turcan<sup>2</sup>, Kathleen McKeown<sup>2</sup>

<sup>1</sup>University of Michigan, Ann Arbor, USA. <sup>2</sup>Columbia University, New York, USA. <sup>3</sup>University of Tennessee, Knoxville, USA. <sup>4</sup>University of Pennsylvania, Philadelphia, USA

In December 2022, the launch of OpenAI’s ChatGPT, the natural language processing model which interacts with users in a conversational format, created a new pipeline of access to NLP tools to the general public. While previous, smaller language models have been well-documented as presenting biased results when dealing with minoritized language varieties, especially African American Language (AAL), little work, has investigated the possible biases in Large Language Models (LLMs), such as ChatGPT, which are often trained on data in which language varieties like AAL are significantly underrepresented. This work, an interdisciplinary collaboration between computer scientists, linguists, and social work researchers, aims to interrogate the limits of these LLMs with respect to their ability to interpret AAL.

Four research assistants who self-identified as proficient AAL users used their own language knowledge to produce what we termed “counterparts” in White Mainstream English (WME, see Baker-Bell 2021) to 346 intonation units containing features of AAL taken from publicly available sources: the Twitter AAE corpus (Blodgett et al., 2018), CORAAL (Kendall and Farrington 2021), and “Country Club Threads” containing

contributions from Black Redditors on r/BlackPeopleTwitter. We then compared the output produced by six different LLMs on the same task, as well as the reverse task of taking WME and outputting AAL. In adjacent work (Authors 2023), we found significant differences in the models' ability to generate AAL, and also found that WME output from the models was judged as failing to preserve the meaning of AAL input.

In this study, we compare the successes and failures of the models against one another, and also examine which AAL features especially cause these failures. We identify three specific categories of features which the models find especially difficult: AAL specific lexical items, AAL-specific verbal aspects, and existential it. For example, an AAL input of "Nerves Been Too Bad To Do Anything Else!" was rendered as "My nerves are too bad to do anything else," removing the remote continuous meaning of been in the input and changing it to present tense.

Further, in examining the training data, we generally find larger performance gaps between WME and AAL in the social-media driven corpora from Twitter and Reddit, than in the data from CORAAL, suggesting that the exaggerated performative patterns of AAL common on social media affect the ability of a model to adequately handle AAL. We additionally discover performance gaps between WME and AAL are larger the larger the size of the model, suggesting that the models may encode a bias toward WME as model size increases. Finally, we compare the frequency of use of specific AAL features in model outputs and data using existing feature detection models (Masis et al., 2022).

This work indicates the presence of significant biases toward WME in these LLMs which render them mostly ineffective at both interpreting and producing AAL, and point toward the possibilities for specific ways in which the models' input and output should be interrogated before they are broadly used in socially significant settings.

Baker-Bell, April. 2020. *Linguistic justice: Black language, literacy, identity, and pedagogy*. Routledge.

Blodgett, Su Lin, Johnny Wei, and Brendan O'Connor. 2018. Twitter Universal Dependency parsing for African-American and mainstream American English. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1415–1425, Melbourne, Australia. Association for Computational Linguistics

Kendall, Tyler and Charlie Farrington. 2021. The corpus of regional african american language.

Masis, Tessa, Anissa Neal, Lisa Green, and Brendan O'Connor. 2022. Corpus-Guided Contrast Sets for Morphosyntactic Feature Detection in Low-Resource English Varieties. In *Proceedings of the first workshop on NLP applications to field linguistics*, pages 11–25, Gyeongju, Republic of Korea. International Conference on Computational Linguistics.

---

## /o/! They're j/u/st about the same!: Vowel Shift in Heritage and Homeland Seoul Korean

Laura Griffin

University of Toronto, Toronto, Canada

The back vowel space in Seoul Korean is currently undergoing a chain shift with /u/ fronting and /o/ raising (Figure 1) (Han and Kang 2013, Kang 2016). This vowel shift appears to be age-graded, with younger female speakers producing more fronted and raised realizations of /u/ and /o/ (Han and Kang 2013). This shift has been largely uninvestigated in diaspora Korean populations. I compare spontaneous speech of homeland and heritage Seoul Korean speakers and show that heritage Korean speakers participate in the vowel shift, sharing a similar vowel space to homeland speakers.

Interview data comes from the Heritage Language Variation and Change Corpus (Nagy 2009, 2011): 8 first generation and 8 second generation heritage speakers in Toronto and 10 homeland speakers in Seoul. Vowels with duration >5 ms and <200 ms (n = 60,082) and their formant values were extracted from Praat (Boersma and Weenink 2014) by script with 5,941 tokens of /o/ and 2465 tokens of /u/. Data was not normalized due to low token numbers per speaker.

Linear mixed effect models in R using RStudio (RStudio Team 2020, R Core Team 2022) tested for effects on F1 and F2 values with speaker as a random intercept and generation (Homeland, Gen1, Gen2), gender, preceding and following manner and place of articulation as factors. Age and phonological factors are not significant main effects for F1 or F2 values, suggesting that the vowel shift is stable. Gen1 and Gen2 speakers produce less peripheral realizations of /u/ and /o/. Generation is not a significant effect for F1 value of /o/, suggesting that the chain vowel shift began with /o/ in the homeland and that this shift has been completed. Figure 2 presents a comparison of average F1 values of /o/ across gender and generation, showing that heritage speakers pattern closely to their homeland counterparts.

Heritage speakers of both genders have overall lower F2 of /u/ than of their homeland counterparts, but only Gen2 female speakers and Gen1 male speakers have significantly lower realizations (p < 0.05). Figure 3 presents a comparison of average F2 values of /u/ compared to homeland speakers. Gen2 female speakers have an average F2 value of 1423 Hz while homeland speakers have an average F2 value of 1541 Hz. Male Gen1 speakers have an average F2 value of 1192 Hz compared to an average homeland value of 1267 Hz, which may be negligible.

These results show that heritage Korean speakers participate in the chain shift, exhibiting /o/-raising and a slightly lower degree of /u/-fronting. Generation and age are generally not significant factors, indicating that the

vowel space is stable in spontaneous speech. Heritage Korean speakers do not have more conservative vowel spaces and instead pattern similarly to homeland speakers, suggesting that heritage speakers receive enough input to participate in the chain shift. The popularity of Korean media may contribute to the amount of input received. First and second generation speakers' lesser degree of /u/-fronting demonstrates that the shift may not be completed but is still progressing.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 6.3.09) [Computer Software].

Han, J. I., & Kang, H. (2013). Cross-generational change of /o/ and /u/ in Seoul Korean I: Proximity in vowel space. *Phonetics and Speech Sciences*, 5(2), 25-31.

Kang, Y. (2016). A corpus-based study of positional variation in Seoul Korean vowels. In *Japanese/Korean Linguistics 23*, eds. Michael Kenstowicz, Ted Levin & Ryo Masuda, 3-22. Stanford: CSLI Publications.

Nagy, Naomi. 2009. *Heritage Language Variation and Change in Toronto*. Available online:

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.591.912&rep=rep1&type=pdf>

Nagy, Naomi. 2011. A Multilingual corpus to explore variation in language contact situations. *Rassegna Italiana di Linguistica Applicata* 43: 65-84.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Yoon, T.-J. & Kang, Y. (2012). A Forced-Alignment-Based Study of Declarative Sentence-Ending 'da' in Korean. In *Proceedings of the 6th International Conference on Speech Prosody 2012*.

---

## Exploring the Virtual Linguistic Landscape: An Analysis of Pronouns of Address in Hispanic and Lusophone Universities

Daniel Guarin, Larissa Oliveira-Cardoso

Temple University, Philadelphia, USA

Today's world is a complex and dynamic system where language plays a central role in shaping social connections and interactions. As technology advances and sociocultural and political issues continue to evolve, language undergoes constant adaptation and change. This makes language a challenging field of study that requires innovative approaches to understanding the intricate relationships between language, society, and technology (Arnaut, 2015; Budach & de Saint-Georges, 2017; Vertovec, 2022). In this context, recent studies have emphasized the importance of virtual spaces as determinants of linguistic behavior. The Virtual Linguistic Landscape (VLL) provides a new perspective for examining complex systems like pronouns of address (PoA) in Spanish and Portuguese.

Studying language in society requires an appreciation of the influence of space, including virtual spaces like the VLL, on language use. PoA is one of the most challenging aspects of Spanish and Portuguese due to its contextual and socially bound variations, which makes it challenging to establish a general theoretical framework for studying PoA (Lara-Bermejo, 2021, p. 7; Ton, 2019, p. 29). In this research, we aim to explore how Higher Education Institutions (HEI) in the Hispanic and Lusophone world use PoA in their VLL and measure how this virtual usage reflects oral practices.

Our study adopts a quantitative variationist approach to analyze the VLL of HEIs from 15 countries (10 Hispanic and 5 Lusophone). We collected data from official websites and social media pages of these institutions, where signs that use any singular second person pronoun in either Spanish or Portuguese were counted. A total of 600 tokens were collected (400 from the Hispanic world and 200 from the Lusophone world) during the second semester of 2022 using a stratified sampling method.

Unlike research on speech variation where social variables like age, gender, sex, and socioeconomic class are interpreted as independent variables, our research, inspired by the Variationist Analysis of Linguistic Landscape Studies (VaLLs) (Soukup, 2020), proposed country as an independent variable to measure how location impacts PoA choice in the VLL of these institutions.

Our data analysis involved creating contingency tables and running chi-squared tests. We found that in Spanish-speaking universities, *tú* was the most commonly used pronoun (73.5%), while in Portuguese-speaking universities, *você* was the most commonly used pronoun (69%). We also discovered that location impacts pronominal variation, which suggests that linguistic patterns are more stable in virtual interactions, despite the significant variability that pronouns present within countries.

In summary, our research highlights the importance of examining virtual spaces like the VLL in understanding complex linguistic phenomena like PoA. By adopting a variationist approach and treating country as an independent variable, we found that location impacts PoA choice in the VLL of HEIs. Our

findings provide new insights into how Spanish and Portuguese pronouns are being used in virtual interactions and how this usage reflects oral practices.

Arnaut, K. (2015). Superdiversity: Elements of an Emerging Perspective. In K. Arnaut, J. Blommaert, B. Rampton, & M. Spotti (Eds.), *Language and Superdiversity*. Routledge.

Budach, G., & de Saint-Georges, I. (2017). Superdiversity and Language. In S. Canagarajah (Ed.), *The Routledge Handbook of Migration and Language* (pp. 63–78). Routledge. <https://doi.org/10.4324/9781315754512>  
[<https://doi.org/10.4324/9781315754512>]

Lara-Bermejo, V. (2021). Introduction. Politeness and pronouns of address in the Iberian Peninsula (1700-1950). *Revista Internacional de Linguística Iberoamericana*, 38(7), 7–9.

Soukup, B. (2020). Survey area selection in Variationist Linguistic Landscape Study (VaLLS): A report from Vienna, Austria. *Linguistic Landscape: An International Journal*, 6(1), 52–79. <https://doi.org/10.1075/ll.00017.sou>  
[<https://doi.org/10.1075/ll.00017.sou>]

Ton, T. N. L. (2019). A literature review of address studies from pragmatic and sociolinguistic perspectives. In B. Kluge & M. I. Moyna (Eds.), *It's not all about you: New perspectives on address research* (1st ed., pp. 23–46). John Benjamins. <https://doi.org/10.1075/tar.1.01ton> [<https://doi.org/10.1075/tar.1.01ton>]

Vertovec, S. (2022). *Superdiversity: Migration and Social Complexity*. Routledge. <https://doi.org/10.4324/9780203503577>

---

### Subject pronouns and communicative function in five languages

Gregory Guy<sup>1</sup>, Karen Beaman<sup>2</sup>, Aria Adli<sup>3</sup>, Rafael Orozco<sup>4</sup>

<sup>1</sup>New York University, New York, USA. <sup>2</sup>Eberhard Karls Universität Tübingen, Tübingen, Germany. <sup>3</sup>Universität Koeln, Cologne, Germany. <sup>4</sup>Louisiana State University, Baton Rouge, USA

Functionalist approaches to variation and change postulate that avoidance of referential ambiguity is a systemic constraint; in Kiparsky's (1982) 'distinctness condition' this arises from "a tendency for semantically relevant information to be retained on the surface." An appropriate site for investigating this position is variable subject pronoun expression (SPE). Functional arguments seek to explain different SPE rates within and between languages: more overt pronouns (vs. null subjects) should be found in contexts and languages where verbal inflections are referentially ambiguous. Prior studies show mixed results: Hochberg (1986) finds overt pronouns in Puerto Rican Spanish favored with ambiguous person/number inflection, while in Andalusian Spanish Ranson (1991:149) finds that "subject pronouns are not used systematically... to compensate for ambiguous verb forms." The present paper tests the functionalist position using comparative evidence from a corpus comprising over 150K tokens from five languages: Spanish, Portuguese, German, Persian, and Mandarin.

Verbal inflection for person/number – prominent in Indo-European languages – is the SPE constraint most relevant to functional conditioning. Unambiguous forms like Sp/Port 1pl -mos (cantamos 'we sing') should permit more nulls, while more overt pronouns should occur with more ambiguous inflections, such as -a, indexing both 3sg and 2sg-formal (canta 'you/he/she sings'). In Swabian German 2sg has unique marking (-sch), motivating low SPE rates, while other forms have ambiguity: 1sg and 3sg are equivalent (-Ø), and -et marks all plurals. In Persian all person/number forms are distinctive except 2sg-formal, implying no functional motive for SPE. Mandarin, with no verbal inflection, presents the biggest challenge to functional explanation.

Our results are inconsistent with these predictions of the functional hypothesis. Pronoun rates do not correlate well with inflectional ambiguity (Table 1). Pronoun rates are indeed lower for plural forms in Portuguese and Spanish, which are more distinctive, but also in Persian (where all desinences are distinctive) and even Mandarin, suggesting some motivation other than inflectional ambiguity, perhaps lower discursive salience for plural referents. Portuguese, Spanish, and Swabian have relatively low rates in ambiguous 3sg. In Swabian, distinctively inflected 2sg has low SPE, but 3sg and 3pl also have lower SPE than inflectionally equivalent first-person forms. In several languages some person/number effects reflect social practice, like politeness norms and the T/V distinction.

Functional arguments are also advanced to explain dialect differences in Portuguese and Spanish. For example, Brazilian Portuguese (BP – 66%) has higher SPE than European Portuguese (EP – 33%) along with structural changes creating more referential ambiguity, such as loss of 2sg inflection, and pronominal a gente 'we' taking 3sg inflection. Similarly, Caribbean Spanish (CS) has high SPE plus consonant deletions increasing ambiguity. But BP and CS show high pronoun rates in all person/number forms, not just those with greater ambiguity. Broader comparative perspectives also undermines this account: Persian has slightly more distinctive inflection than EP, but much lower SPE (18% vs 32%), and

Mandarin, with no inflection, has 40% SPE, much lower than inflected BP and Swabian. Thus, cross-language differences must be due to factors other than function.

Hochberg, Judith. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62:609-621

Kiparsky, Paul. 1982. *Explanation in phonology*. Dordrecht: Foris.

Ranson, Diane. 1991. Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change* 3:133-152.

---

## Acquisition of Variation: Acquiring Constraints on Variable Clitic Climbing in Children

Emily Herman

Pennsylvania State University, State College, USA

**Introduction.** Variable clitic placement (VCP) in Spanish is a well-documented case of linguistic variation. In Spanish, pronominal clitics can refer to an object in the discourse or express the reflexivity of an action. In [finite + non-finite verb] constructions, pronominal clitics can attach either pre-verbally (proclisis; *Ella lo quiere comprar* ‘She it wants to buy’) or post-verbally (enclisis; *Ella quiere comprarlo* ‘She wants to buy it’). Previous research has indicated that a variety of factors can influence the proclitic or enclitic position, such as the finite verb and animacy of the referent [2, 5]. Although most corpus research on this topic has been done on adult speech [5, 6] experimental work has shown that by the age of 4;0, children have acquired the lexical and semantic constraints on VCP in Spanish [2]. However, most of the previous research on children has been cross-sectional experimental data [2, 4], and not much research has been conducted on longitudinal child corpus data.

The present study aims to gain a more complete understanding of the acquisition of VCP. Given the tendency of children to overextend one variation over another with other variable phenomena [see 1], one would expect that children would overextend proclisis or enclisis in variable contexts; previous research has indicated that this is not the case [2, 4]. The question is then: when do children acquire the sociolinguistic variables that constrain proclitic and enclitic position?

**Methods.** Tokens were extracted from five corpora in the CHILDES database (Aguirre, Ornat, Linaza, SerraSole, Nieva, and Vila) and one corpus from PhonBank (Llinas Ojea). Each child was monolingual between the ages of 1;10 and 4;11. The context of each corpora ranged from play (with their adult caregivers) to routines (such as taking a bath, cooking dinner, etc.). All instances of VCP contexts were extracted from the child tokens. The first 300 adult tokens were also extracted from each corpora to compare child and caregiver variable clitic placement. There were 800 child tokens and 1891 adult tokens. Tokens were coded for: clitic, clitic type, VCP, finite verb, non-finite verb, negation, question, number of clitics, animacy, and clitic doubling.

**Results & Discussion.** All speakers showed an overall preference for proclisis in variable contexts. The finite verb was found to be a significant factor in the rate of proclisis, as well as the type of clitic. Some finite verbs were found to have a significant effect on enclisis rates in adults, while a subset of these verbs were significant factors for children. Direct object clitics had a higher rate of enclisis than indirect and reflexive pronouns in both children and adults. Additionally, the longitudinal data provide a broad picture of variable clitic placement acquisition. The results indicate that the acquisition of VCP happens at an early stage in language, and the constraints are acquired gradually and a closer look at Maria and Irene’s development provide support for Shin and Miller’s 2022 model of variation acquisition [3].

[1] Miller, K. (2013). Variable Input: What Sarah Reveals About Nonagreeing Don’t and Theories of Root Infinitives. *Language Acquisition*, 20(4), 305–324. [2] Requena, P. E. (2020). A Usage-Based Perspective on Spanish Variable Clitic Placement. *Languages*, 5(3), 33. [3] Shin, N., & Miller, K. (2022). Children’s Acquisition of Morphosyntactic Variation. *Language Learning and Development*, 18(2), 125–150. [4] Eisenchlas, S. (2003). Clitics in Child Spanish. *First Language*, 23(2), 193–211. [5] Davies, M. (1995). Analyzing Syntactic Variation with Computer-Based Corpora: The Case of Modern Spanish Clitic Climbing. *Hispania*, 78(2), 370. [6] Schwenter, S. A., & Torres Cacoullos, R. (2014). Competing constraints on the variable placement of direct object clitics in Mexico City Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 27(2), 514–536.



---

## Stylized performance of “mock Berber” in a Moroccan Stand-Up comedy talent show

Atiqa Hachimi<sup>1</sup>, Gareth Smail<sup>2</sup>

<sup>1</sup>University of Toronto, Toronto, Canada. <sup>2</sup>College of Charleston, Charleston, USA

In 2011, Amazigh (commonly known as Berber) became a second official language next to Arabic in Morocco after decades of activism for Berber linguistic and cultural rights. With this newly found legitimacy, modest use of the language started manifesting on widely pop cultural shows on prime-time state TV to a nation-wide audience composed of a non-Berberophone majority and a Berberophone minority. In this paper, we examine “mock Berber”, i.e., a performance of stylized Amazigh accent in Moroccan Arabic, by a young Amazigh comedian, Zakaria Ouarssam, who won the 2019 season of the Moroccan Stand-Up competition. The talent show aired on prime time on the First national TV channel (Al Oula) and continues to garner millions of views on YouTube, offering a rich sociolinguistic and metacommunicative context for understanding the creative and strategic uses of regional, social, and stylistic variation in contexts of Berber-Moroccan Arabic language contact. As a bilingual lower-class youth from Khenifra—a small city in the Middle Atlas—Zakaria’s successful comedic performance, we argue, sheds new light on the representation of a demographic and local identity that are marginalized both within a broader national culture, as well as the Amazigh movement itself, even as it reproduces certain linguistic hierarchies and ideologies and normalizes a way of being from which he has been excluded.

Our study is based on four prime shows, each lasting approximately two and a half hours, including not only Zakaria’s onstage live performances and interactions with the judges and the studio audience, but also his interactions backstage with fellow contestants, coaches, and show animators, as well as interviews with the contestants’ family members. Specifically, we investigate the linguistic practices and embodied actions undertaken by Zakaria onstage and off-stage with an eye on the wide range of indexical meanings animated through his playful voicing of different Amazigh characters and personas, and the stances he takes vis-à-vis these voices, but also examine the evaluation of his performances by different audiences.

Stylized regional Moroccan Arabic accents have a long history of mediatization in Morocco, particularly as comedy. The analysis shows how Zakaria’s stylized performances, by virtue of their strategic inauthenticity, contribute to the construction and valorization of an “accentless” Moroccan Arabic via a mass mediated show, a process which is as much ideological as it is linguistic. At the same time, the analysis of the different audiences’ uptake of his performance reveals that although his stylized mocking is no doubt offensive to some Amazigh identity activists, it is also celebrated, wildly popular, and taken up by many on social media as an expression of Amazigh pride. Overall, the study has broader implications for the role of performance in ongoing sociolinguistic change in Morocco (Hachimi 2022), insofar as it “provides a frame that invites critical reflection on communicative processes” (Bauman and Briggs 1990:61).

Bauman, Richard, and Charles L. Briggs. 1990. Poetics and performance as critical perspectives on language and social life.

*Annual Review of Anthropology* 19, 59-88.

Chun, W. Elaine. 2004. Ideologies of legitimate mockery: Margaret Cho’s revoicings of mock Asian. *Pragmatics* 14:2/3, 263-289.

Coupland, Nikolas. 2001. Dialect stylization in radio talk. *Language in Society* 30:3, 345-375.

Hachimi, Atiqa. 2022. “It’s Cool to Sing Moroccan”: Ideologies of slang and contested meanings of Arabic popular music on social media. *International Journal of the Sociology of Language* 278, 107-131.

---

## Conversational Comfort and Linguistic Insecurity with Respect to Gender Identity, Sexual Orientation, and Transness

Shannon Harasta

Michigan State University, East Lansing, USA. Southern Illinois University, Carbondale, USA

Linguistic insecurity has been framed as individuals’ awareness of a “correct” speech form, and the gap between that form and what they believe they actually produce (e.g., Labov, 1966; Thompson and Baker, 1984; Preston, 2013). As the straight white man can be seen as a sort of cultural default in the American context (see Kiesling 2018), speech that marks speakers as anything other than that norm may trigger these feelings of insecurity. Deviation from expected norms may be of particular concern for non-straight people or transgender people for whom there is a possibility that their speech may out them (see Daniele et al., 2020; Zimman, 2017), which in turn may have consequences, both in terms of social ramifications and personal safety (see Brumbaugh-Johnson & Hull, 2018; Orne, 2011). This study seeks to examine linguistic insecurity while paying special attention to gender identity, sexual orientation, and transness. Survey data was collected from 232 respondents. Respondents were asked their level of comfort in various speech settings, what concerns they have about their own speech, and what negative comments they have received about their speech. Free response answers were coded for themes, while Likert data

were analyzed using chi-squared, t-tests, and ANOVA. Analysis showed that the comfort of speakers across social settings was linked to gender, sexuality, and transness. Men (n=67) had higher comfort scores than women (n=146) who had higher scores than those with other gender identities (n=12); straight people (n=126) had higher scores than non-straight people (n=88); and cisgender people (n=207) had higher scores than transgender people (n=20). All differences were statistically significant ( $p < .05$ ) except for that between women and people with other gender identities. While people across groups were similarly likely to think that their speech needed improvement, non-straight and transgender people were significantly more likely to report receiving negative comments about their speech than straight or cisgender people. Patterns can also be seen regarding what specific concerns people have about their speech. The most common concerns mentioned include confidence (n=27), stuttering (n=23), clarity (n=18), and pronunciation (n=17), with confidence and stuttering being overrepresented among women. As specific speech patterns become enregistered (see Johnstone, 2016) to different groups (e.g., sexy baby voice, gay voice) speakers may carry the feeling that these enregistered items deviate from the “correct” standard form, causing them to feel linguistic insecurity. Additionally, these results indicate that when speakers deviate from a norm (male vs. other, straight vs. other, cisgender vs. other) they feel less comfort speaking than members of more hegemonic categories.

- Brumbaugh-Johnson, Stacey M. & Kathleen E. Hull. 2018. Coming out as transgender: Navigating the social implications of a transgender identity. *Journal of homosexuality* 66.8: 1148-1177.
- Daniele, Maddalena, Fabio Fasoli, Raquel Antonio, Simone Sulpizio & Anne Maass. 2020. Gay voice: Stable marker of sexual orientation or flexible communication device?. *Archives of Sexual Behavior* 49: 2585-2600.
- Johnstone, Barbara. 2016. Enregisterment: How linguistic items become linked with ways of speaking. *Language and Linguistics Compass*, 10(11), 632-643.
- Kiesling, Scott F. 2018. The “Gay Voice” and “Brospeak”: Toward a Systematic Model of Stance.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
- Orne, Jason. 2011. ‘You will always have to “out” yourself’: Reconsidering coming out through strategic outness. *Sexualities* 14.6 (2011): 681-703.
- Owens, Thompson W. & Paul M. Baker. 1984. Linguistic insecurity in Winnipeg: Validation of a Canadian index of linguistic insecurity. *Language in society* 13.3: 337-350.
- Preston, Dennis R. 2013. Linguistic insecurity forty years later. *Journal of English Linguistics* 41.4: 304-331.
- Zimman, Lal. 2017. Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society*, 46(3), 339-370.

---

## Individual language variation across discourse types: a study of 112 idiolects

Annina Heini, Krzysztof Kredens, Piotr Pezik  
Aston University, Birmingham, United Kingdom

Forensic authorship analysis is based on two assumptions: (i) that language users can have unique linguistic styles, or quantifiable ‘idiolects’, and (ii) that features characteristic of those styles are likely to recur with a relatively stable frequency in an individual’s linguistic output (Coulthard, 2004). Hundreds of style markers and a variety of author-matching techniques have been proposed over the years, with some studies reporting true-positive attribution rates for the less complex closed-set tasks in the region of 90 per cent (e.g. Grieve, 2007; Koppel et al., 2013; Wright, 2017). However, the issue with those studies has been their tendency to use data from just one discourse type, and the few existing cross-genre studies are typically limited to two genres and are based on only a small number of authors (e.g. Kestemont et al., 2012; Stamatatos, 2013).

Meanwhile, a forensically useful author identification system needs to be able to capture stylistic similarities also between texts created in different contexts and for different purposes and audiences. As expert witnesses, forensic linguists often compare disputed documents of one genre with known-authorship documents of another; they might, for example, be asked to compare a set of text messages with business emails. To help establish a protocol to opine on cross-genre authorship problems, we thus need an understanding of individual stylistic variation across, rather than within, discourse types. For our study, 112 participants have shared with us natural language samples from eight discourse types. We have collected emails, text messages, university essays, oral interview data, oral image description data, digital data of Google search behaviour, business memos, and handwritten texts. Each participant’s dataset comprises a wide range of genres but also communication channels, contexts, and language input modes. The individual datasets consist of roughly 10,000 words each, amounting to a total corpus size of over a million words.

We have used stylometric classification tools to measure within-author and between-author variability drawing on samples from six of the discourse types. The results indicate low levels of individual stability, with clustering instead occurring according to discourse types. We identify a set of apparently idiolectal features that have ‘survived’ the genre effects, and we offer a sociolinguistically-based interpretation of the results, with a special focus on their implications for forensic authorship analysis.

- Coulthard, M. (2004) Author Identification, Idiolect, and Linguistic Uniqueness. *Applied Linguistics* 25(4). 431-447.
- Grieve, J. (2007). Quantitative Authorship Attribution: An Evaluation of Techniques. *Literary and Linguistic Computing* 22. 251-270.
- Kestemont, M., Luyckx, K., Daelemans, W. & Crombez, T. (2012) Cross-genre authorship verification using unmasking. *English Studies* 93(3). 340-356.
- Koppel, M., Schler, J. & Argamon, S. (2013) Authorship Attribution: What's Easy and What's Hard? *Journal of Law and Policy* 21(2). 317-331.
- Stamatatos, E. (2013) On the robustness of authorship attribution based on character n-gram features. *Journal of Law and Policy* 21(2). 421-439.
- Wright, D. (2017) Using word n-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem. *International Journal of Corpus Linguistics* 22(2). 212-241.

---

## How do you say pasta?: Foreign (a) in Canada

Claire Henderson

McGill University, Montreal, Canada

The way /a/ is adapted into English from foreign loanwords has long formed a split: Americans typically use back (a) (/ɑ/ or merged /ɑ/-/ɔ/) while Canadians have traditionally used front (a) (/æ/) in words like drama or pasta (Avis, 1973; Boberg, 2000). However, the use of back (a) has been increasing in Canada (Boberg, 2000; Boberg, 2009). In particular, Boberg (2000) found that adoption of back foreign (a) was higher in Windsor than Toronto, suggesting there is diffusion across the Windsor-Detroit border. The current study looks at Canadian foreign (a) pronunciation in Toronto and the Niagara border region using questionnaire and acoustic data. The main goal was to see whether back (a) use is continuing to increase and whether the border plays a role in its diffusion.

The questionnaire asked participants (N=803) to match their pronunciation to the sound in two unambiguous words (e.g., passable or possible for pasta). The acoustic study recorded participants (N=66) reading a word list to analyze (a) and its relation to /æ/ and /ɑ/-/ɔ/. The participants were also asked for opinions on Canadian and American speech. Data were collected on both sides of the border to investigate the possibility of cross-border diffusion.

The results show continued increase of back (a) in Canada with small increases in backness in apparent time. Niagara is overall ahead of Toronto: across variables, the mean (a) in Toronto (fig.1) is significantly further forward and overlaps more with /æ/ than Niagara (fig.2) which overlaps more with /ɑ/-/ɔ/. Toronto (a) also has greater variance, suggesting more variation between /æ/ and /ɑ/-/ɔ/. However, the pattern also varies by word. The questionnaire shows that Canadians in both regions primarily use back (a) with one exception: pasta has back (a) in Niagara and front (a) in Toronto. This split is confirmed by the acoustic results with pasta's mean significantly further forward in Toronto. The acoustic data also reveal instances of both front and back (a) in both regions: certain words (e.g., Iran, pajamas, bratwurst) continue to favour front (a) while others (e.g., drama, llama, façade) clearly favour back (a) (fig.3). For words with back (a), the Niagara means tend to be further back suggesting more speakers have adopted the back pronunciation.

Overall then, the pronunciation of foreign (a) in Canada is increasingly favouring back (a) and seems to show diffusion across the border, with Niagara more advanced than Toronto. However, it remains unclear if the diffusion is actually from Buffalo or a more general effect of exposure to American speech, particularly since the quality of back (a) in Canada tends to align more with back, merged /ɑ/-/ɔ/ rather than the fronted, unmerged /ɑ/ found in Buffalo. Additionally, many Niagara participants who used /ɑ/-/ɔ/ in pasta incorrectly thought that Buffalo uses /æ/ and associate this pronunciation with negative views of tensed /æ/. These results therefore provide further evidence for the increase of back (a) in Canada while also raising questions about how this feature is diffusing across the border.

- Avis, W. S. (1973). The English language in Canada. In T. A. Sebeok (ed.), *Current Trends in Linguistics* (Vol. 10): Linguistics in North America (pp. 40-74). Mouton.
- Boberg, C. (2000). Geolinguistic diffusion and the U.S.-Canada border. *Language Variation and Change*, 12, 1-24.
- Boberg, C. (2009). The emergence of a new phoneme: Foreign (a) in Canadian English. *Language Variation and Change*, 21(3), 355-380.

---

(DH)ere's more than one way to strengthen a fricative:  
Voiced interdentals in a diverse sample of Californians

Alexia Hernandez, Bran Papineau, Marie Tano, Tony Velasquez, Hannah Zhang, Robert Podesva  
Stanford University, Stanford, USA

Studies on (DH), one of the most widely studied variables in English due to its ubiquity across varieties (e.g. Labov et al., 1968; Rose, 2006; Mendoza-Denton, 2008; Grieser, 2013), have traditionally relied on auditory coding and classified its realization as either a stop or a fricative. In this paper, we use acoustic methods to make finer-grained distinctions among these variants. In particular, we note the prevalence of released fricatives (RF; frication noise followed by a release burst) and observe their distinctive social patterning among Latinx speakers compared to stops. We argue that differently fortified realizations of (DH) can index different social meanings.

All instances of words with (DH) (N=14,148) were extracted from 190 sociolinguistic interviews with speakers from across California (84 Latinx, 25 Black, 81 White), with a maximum of 3 tokens/lemma/ speaker. The sample was balanced in terms of age and gender and included speakers with varying education levels. Based solely on spectrographic and waveform analysis, the manner of articulation for each (DH) token was categorized as deleted, fricative (Figure 1A), RF (1B), or stop (1C). Each token was coded for several linguistic (word position, lexical frequency, preceding phonological environment, stress) and social (birth year, gender, race/ethnicity, field site, education level, bilingualism) factors. Two mixed-effects logistic regression models (full sample, Latinx speakers only) were stepped up to identify factors influencing the occasion of stops, another two models for RFs.

Both variants' full models show significant effects of word position and race/ethnicity: (DH) is fortified word-initially more than word-medially, and Latinx speakers fortify (DH) more than White or Black speakers (Figure 2). Importantly, we find that the social patternings of stopped realizations and RFs differ for Latinx speakers. The Latinx stop model shows an interaction between education and bilingualism: The least educated Spanish-English bilinguals show the highest proportion of stop realizations (Figure 3). By contrast, the Latinx RF model shows only that women are more likely to use RF than men (Figure 4).

Even though (DH)-stopping is not exclusive to Latinx varieties of English, Latinx speakers' greater use of both stops and RFs here suggests that fortition in general could be available as a resource for indexing Latinx identity. However, other social distinctions can be conveyed through the use of stops vs. RFs. Stops' correlation with lower education levels (among bilinguals) suggests that they are ideologized as less standard. An RF realization, on the other hand, might represent a way to strengthen (DH) without committing to a stop and its association with non-standardness. It's noteworthy that women, canonically described as prototypical users of standard forms of stable variables, use RFs more than men, and that education did not structure variation patterns for RFs.

Our results suggest that previous approaches that have relied on auditory methods may have conflated stops with RFs (some of which we perceived as stops). Future work should consider phonetic detail even for variants that have seemingly categorical realizations, since finer-grained distinctions can have distinct social indexicalities.

- Grieser, Jessica. (2013). Locating style: Style-shifting to characterize community at the border of Washington, DC. *University of Pennsylvania Working Papers in Linguistics*, 19(2), 10.
- Labov, William, Paul Cohen, Clarence Robins, and John Lewis. (1968). *A study of the non-standard English of Negro and Puerto Rican speakers in New York City*. Report on cooperative research project 3288. New York: Columbia University.
- Mendoza-Denton, Norma. (2008). *Homegirls: Language and Cultural Practice Among Latina Youth Gangs*. Hoboken: Wiley.
- Rose, Mary. (2006). *Language, place and identity in later life*. PhD Thesis Linguistics, Stanford University.

---

The present indicative and confidence about the future:  
Presumed settledness in French and across Romance

Mark Hoff  
Queens College CUNY, New York City, USA

Argentine Spanish, Brazilian Portuguese, Italian, and Hexagonal French allow for tense/mood variation in future-framed adverbial clauses such that the present indicative (PI) competes with the standard variant, which differs across languages, as in (1).

Author (2019, 2020) provides qualitative and quantitative data to show that in Argentine Spanish, Brazilian Portuguese, and Italian, this variation is conditioned by presumed settledness (Kaufmann 2002); that is, PI is both chosen significantly more in forced-choice tasks and rated significantly more favorably in acceptability judgment tasks when contextual clues indicate that the speaker perceives the future eventuality in the subordinate clause (in 1, her arrival) as guaranteed in light of her beliefs at speech time. The notion of presumed settledness facilitates cross-linguistic comparison, despite differing standard forms in Argentine Spanish, Brazilian Portuguese, and Italian, by highlighting the shared pragmatic function of PI in all three languages as an efficient means for speakers to signal

confidence about future eventualities. However, the form-function isomorphism differs across Romance; whereas Argentine Spanish and Brazilian Portuguese neatly restrict PI to contexts of very high confidence, in Italian this preference is maintained yet relaxed, resulting in less perfect form-function symmetry (Poplack & Dion 2009). The present study offers quantitative data to demonstrate that: a) Hexagonal French allows PI in future-framed adverbials, a fact heretofore unacknowledged in the vast literature on French future expression; b) PI use is conditioned by settledness but is also possible in contexts of less-than-absolute confidence (cf. Sankoff 1988); and thus c) French constitutes an intermediate point on a cline of PI restriction in comparison to these other languages.

Data come from an online contextualized forced-choice task completed by 277 native French speakers. Presented with 8 target items (see 2 and 3) and 20 fillers, each containing a preceding context and bolded carrier sentence, participants chose between PI and future indicative (FI) to fill in a blank in both main and subordinate clauses. Stimuli were maximally distinct in terms of explicit indicators of the likelihood of occurrence, temporal distance, and temporal specificity of the future eventuality in the subordinate clause. Results of mixed-effects logistic regression in R with a random slope for settledness condition by participant and a random intercept for verb show that PI is chosen significantly more ( $p < .001$ ) in subordinate clauses expressing future eventualities that are presumed settled (70% choice of PI), whereas FI is clearly preferred for non-settled eventualities (75% choice of FI). Appreciable interspeaker variation was observed, though it did not map clearly onto sex, age, region, or level of education in this sample.

These experimental findings build upon and challenge previous, corpus-based characterizations of French future reference (Poplack & Turpin 1999, Poplack & Dion 2009) which emphasize the importance of morphosyntactic factors (e.g. polarity) and minimize the role of semantic/pragmatic distinctions. These results also provide further evidence for the appropriateness of presumed settledness to account for tense/mood variation across multiple languages and syntactic contexts (Author & Other 2023) in a way that highlights the utility of such a contrast for speakers' communicative goals.

Author. 2019. Details omitted for blind reviewing.

Author.. 2020. Details omitted for blind reviewing.

Author & Other. 2023. Details omitted for blind reviewing.

Kaufmann, Stefan. 2002. The presumption of settledness. *CLS* 38. 313-328.

Poplack, Shana & Danielle Turpin. 1999. Does the futur have a future in (Canadian) French? *Probus* 11, 1. 133-164.

Poplack, Shana & Nathalie Dion. 2009. Prescription vs. praxis: the evolution of future temporal reference in French. *Language* 85, 3. 557-587.

Sankoff, David. 1988. Sociolinguistic and syntactic variation. In Frederick Newmeyer (Ed.), *The Cambridge Survey*, 140-161. Cambridge: Cambridge University Press.

---

## Attention to speech (varieties) or attention to task?

### Locating styles in Arabic and Francoprovençal

Jonathan Kasstan<sup>1</sup>, Uri Horesh<sup>2,3</sup>

<sup>1</sup>University of Westminster, London, United Kingdom. <sup>2</sup>Achva Academic College, Palestine. <sup>3</sup>University of Essex, Colchester, United Kingdom

What theory undergirds our choice to elicit certain kinds of data in certain kinds of ways? We explore this question through the prism of experience gained in eliciting and analysing speech samples in communities where English is not a dominant language, and where bi/multilingualism might interact along the stylistic dimension typically tapped into by attention-to-speech (Labov 1971).

We consider data from two genealogically and typologically different varieties: Francoprovençal as spoken in the Lyonnais mountains (France) and the Canton of Valais (Switzerland), and Arabic varieties spoken in Palestine. Unlike in monolingual English-speaking communities, in these settings we observe that styles that would typically be characterised as 'formal' and 'casual' are instead located on different planes and can involve criss-crossing different language varieties altogether. Table 1 illustrates how, for Arabic, some of the phonological variables previously analysed apply to one or the other variety, with only partial overlap.

How continuous then is the formal~informal 'continuum' in such language settings? For native speakers of Palestinian Arabic, two and sometimes three systems interact: Palestinian Arabic, Standard Arabic, and, for many, Modern Hebrew. In contrast, native speakers of Francoprovençal are now all French-dominant, and there are few remaining domains where their L1 is used. Here, attention-to-speech-based tasks can elicit stylistic variants that are in some cases shared with French, and in others not. One example includes an allophonic rule of lateral palatalisation before obstruents in Francoprovençal, which is not present in French (e.g., /kja/ vs. /kle/ 'key'), a variable characterised as salient for speakers. Production studies have reported the distribution of this variable to be inverted (see Figure 1), i.e., more attention to speech triggers more vernacular (Francoprovençal) variants; less attention to speech triggers a French-like variant where the allophonic rule is neutralised.

We synthesise these observations and consider:

1. to what extent stylistic variation reported in these varieties is a reflex of adapting classic Labovian methods to very different linguistic ecologies (i.e., task-based effects)
2. how must we adjust and adapt elicitation methods in order to discern micro-level stylistic variation within each of the language systems.

While macro-level shifting also occurs, linguistic elements from each of the linguistic systems inevitably finding their way into the others, we focus on analysing stylistic variation and shifts internal to each system. Returning to Arabic, it is important to acknowledge that formal speech occurs in virtually all of the regional dialects. For instance, while the prescribed norm of delivering an academic lecture is in Standard Arabic, adherence to this norm is declining. Some of the data we have collected or otherwise analysed represent formal styles of Palestinian (not Standard) Arabic. In some cases, we have data from the same speaker in different styles: e.g., casual data elicited in a sociolinguistic interview, alongside monologues recorded in more formal settings.

On the basis of our observations, we recommend that new benchmarks should be decided upon, based on empirical research and experimentation with methods, to establish what constitutes formal and informal style in bi/multilingual speech communities.

Labov, William. 1971. Some principles of linguistic methodology. *Language in Society* 1: 97–120.

---

## Perceptions of "Southern" Utah English

Chad Huckvale, Joey Stanley  
Brigham Young University, Provo, USA

Utahns generally have negative attitudes towards some features of Utah English (Savage 2014, Eddington & Brown 2021, Stanley 2023). However, while previous work has analyzed perceptions of specific linguistic features, in this study we focus on descriptors like “southern” and how they are used to describe language in specific regions of Utah.

Experiment 1 is a classic draw-a-map study (Preston 1989, Bucholtz et al. 2007) wherein participants were asked to mark areas on a map of Utah where people speak differently. 193 surveys were completed, primarily by university students from a university in Utah. Based on the labels on these surveys, participants are generally aware of certain features of Utah English, like Utahns’ pronunciation of mountain (Eddington & Savage 2012; Stanley & Vandierneet 2018; Eddington & Brown 2021). However, the most common labels, with 128 references, were those such as “rural,” “hick,” “Southern,” and “cowboy.” These labels were most common in regions outside of Utah’s population centers, i.e. the Wasatch Front (Salt Lake City, Provo, and Ogden areas; Figure 1).

To understand why certain regions of Utah were designated as “Southern” by participants, despite Utah’s lack of shared features with Southern United States English (SUSE), Experiment 2 uses a slightly altered version of an existing research method, which we call the “perceptual audio survey” (similar to “voice-placing task” (Cramer & Montgomery 2016:10) or “dialect classification task” (Chartier 2020)). This method invites participants to listen to recordings of native English speakers and identify where in Utah this speaker is likely from. Crucially though, we used speakers from throughout the United States (13 speakers in total). While participants were not explicitly told that the speakers were from Utah, but this was greatly implied. Surveys were distributed through various social media channels, resulting in 159 responses. Generally, we find that speakers with features of SUSE were more likely to be placed in rural areas (Figure 2). The data also shows a divide between speakers age 20–49 and speakers age 50+: listeners were more likely to place younger speakers in urban regions of Utah and older speakers outside of urban centers. This is in line with what was found in Seoul, Korea by Shin et al. (2020).

The results of these two experiments reveal some alignments and misalignments between perceptions of Utah English. On one hand, rural areas are associated with southern speech and vice versa, suggesting that Utah English objectively lacks most linguistic features traditionally associated with SUSE, those features are indexical of the rural parts of Utah. Meanwhile, while many map-drawers labeled both the Wasatch Front and St. George (a rapidly-growing city in southwest Utah) as “California,” listeners usually placed actual Californian speakers in the Wasatch Front and not St. George, further suggesting that imagined perceptions of St. George do not align with linguistic perceptions.

Overall, this study gives new insight into the concept of perceived Southernness in some regions of rural America (Hall-Lew & Stephens 2012, Podesva et al. 2015) and reveals potential biases and ideologies held by Utahns regarding language variation.

Bucholtz, Mary, Nancy Bermudez, Victor Fung, Lisa Edwards & Rosalva Vargas. 2007. Hella Nor Cal or Totally So Cal?: The Perceptual Dialectology of California. *Journal of English Linguistics*. SAGE Publications Inc 35(4). 325–352. <https://doi.org/10.1177/0075424207307780> [<https://doi.org/10.1177/0075424207307780>].

- Chartier, N. 2020. Old stereotypes "live free or die": addressing the evaluation problem of non-rhoticity in New Hampshire. [Doctoral dissertation, University of Washington]. ProQuest Dissertations Publishing.
- Cramer, Jennifer & Chris Montgomery. 2016. *Citiescapes and Perceptual Dialectology: Global Perspectives on Non-Linguists' Knowledge of the Dialect Landscape*. Boston, UNITED STATES: De Gruyter, Inc. <http://ebookcentral.proquest.com/lib/byu/detail.action?docID=4426407>. (23 May, 2023).
- Eddington, D. & M. Savage. 2012. Where Are the Moun[ ]ns in Utah? *American Speech* 87(3). 336–349. <https://doi.org/10.1215/00031283-1958345>.
- Eddington, David Ellingson & Earl Kjar Brown. 2021. A Production and Perception Study of /t/-Glottalization and Oral Releases Following Glottals in the United States. *American Speech* 96(1). 78–104. <https://doi.org/10.1215/00031283-8620501>.
- Hall-Lew, Lauren & Nola Stephens. 2012. Country Talk. *Journal of English Linguistics*. SAGE Publications Inc 40(3). 256–280. <https://doi.org/10.1177/0075424211420568>.
- Podesva, Robert J., Annette D'Onofrio, Janneke Van Hofwegen & Seung Kyung Kim. 2015. Country ideology and the California Vowel Shift. *Language Variation and Change*. Cambridge University Press 27(2). 157–186. <https://doi.org/10.1017/S095439451500006X>.
- Preston, Dennis R. 1989. *Perceptual Dialectology: Nonlinguists' Views of Areal Linguistics*. Berlin/Boston, GERMANY: De Gruyter, Inc. <http://ebookcentral.proquest.com/lib/byu/detail.action?docID=3042196>. (16 December, 2021).
- Savage, David Matthew. 2014. *How We Feel about How We Talk: A Language Attitude Survey of Utah English*. United States -- Utah: Brigham Young University M.A. <https://www.proquest.com/docview/2446960747/abstract/3585BA95D908480DPQ/1>. (22 November, 2021).
- Shin, W., Lee, H., Shin, J., Holliday, JJ . 2020. The Potential Role of Talker Age in the Perception of Regional Accent. *Language and Speech*.
- Stanley, Joseph A. 2023. Utahns sound Utahn when they avoid sounding Utahn. <https://joeystanley.com/blog/lisa-and-ads-2023>.
- Stanley, Joseph A & Kyle Vanderniet. 2018. Consonantal Variation in Utah English. 16.

---

Wiggly lifespan change in a crisis – Contrasting reactive and proactive identity construction  
Stefánsdóttir Lilja, Ingason Anton  
University of Iceland, Reykjavík, Icela

We examine individual lifespan change in Icelandic and argue that style shift during personal crises is crucially tied to the agency of the individual involved: A reactive response leads to stylistic downshift whereas a proactive response leads to upshift. We investigate an MP who reactively lays low until the crisis ends, in which case style temporarily shifts toward less formal use.

**Background:** While some linguistic lifespan change is systematically connected to community change (e.g. Sankoff and Wagner 2006; 2011), reasons for change can be individual-specific; tied to personal histories. For Neil of 7-up, “abandonment of the Short-U pattern [was] tied to the locus of his painful adolescence” (Sankoff 2004). While David Attenborough’s /ó/-tapping remains relatively stable, his stylistic range shrinks as he gains status (MacKenzie forthcoming). The Icelandic finance minister, Sigfússon, acquired a more formal style during the economic crisis of 2008–11, as seen in his rate of Stylistic Fronting (SF), hypothesized to reflect long term style predictors like his Linguistic Market Value (Sankoff and Laberge 1979; Stefánsdóttir and Ingason 2018; Figure 1).

**Variable:** SF is optional movement of elements to a phonologically empty Spec,TP in Icelandic, see (1–2). SF reflects formal style (Maling 1980), was historically stable (Figure 2), but manifests a more complicated usage pattern since the 20th century (Figure 3), demonstrating how syntactic variables can gain social salience (cf. Sneller and Fisher 2015).

**Personal crisis of an MP:** Here, we consider SF use of another Icelandic MP. He has experienced dramatic highs and lows in his career, including switching parties in 2012 and being intoxicated on an airplane in 2014 (Jóhann Óli Eiðsson 2015; Bjarki Ármannsson 2015), resulting in media storms and calls for his resignation. After keeping a low profile, the MP eventually left Parliament. However, after another election, the MP was re-elected and became a minister. Later, he changed the title of his office to reflect his focus on the welfare of children. As minister, he opened up about his troubled childhood and his popularity rose. Since then, every poll has shown him to be the MP whom the public trusts the most (e.g. Maskína 2022, 2023; Magnús Geir Ólafsson 2021). This is a story of a crisis and redemption, and interestingly, there is a correlation between the fluctuation of his use of SF and the highs and lows of his political career; notably his 2013–14 low-point corresponds to his lowest SF use (Figure 4).

**Implications — Reactive Downshift vs. Proactive Upshift:** Our MP’s crisis downshift is unlike Sigfússon’s, who assumes a proactive leadership role during a crisis, with rich agency, and shifts his style upward until the crisis has passed. Importantly, the dominant lifespan effect on style is sometimes an idiosyncratic property of a personal history — reversible if the property is temporary. Most studies on lifespan change lack a fine-grained time axis to detect year-by-year fluctuations.

Note how an emphasis on comparing two points in time shapes the field: Sankoff (2019), working with 2 points in time, describes a typology of three patterns of lifespan change. This typology de-emphasizes personal histories that diverge from community behavior. We argue that especially during personal crises, individual-specific effects

can temporarily be the dominant predictor of style, outranking longer-term trends, and levels of agency can shape the trajectory of the change.

- Bell 1984. Language style as audience design.  
Bjarki Ármannsson 2015. Var Ási að fá sér? Flugfarþegi segir þingmann hafa ælt á sig eftir ofdrykkju.  
Jóhann Óli Eiddsson 2015. #Ásiaðfásér. Tíst um uppsölu þingmannsins.  
Labov 1972. Sociolinguistic patterns. • Labov. Principles of Linguistic Change. Oxford, UK: Blackwell, 1994.  
MacKenzie forthcoming. Tracking stylistic variation over a very long lifespan.  
Magnús Geir Ólafsson 2021. Ásmundur Einar rýkur upp ánægjulistann.  
Maling 1980. Inversion in embedded clauses in Modern Icelandic.  
Maskína 2022. Traust til allra leiðtoga ríkisstjórnarflokkanna dregst saman frá 2021 – enginn ráðherra nýtur meira trausts í ár en í fyrrahaust.  
Sankoff, D. & Laberge 1979. The linguistic market and the statistical explanation of variability.  
Sankoff, G. 2004. Adolescents, young adults and the critical period: Two case studies from ‘Seven Up’.  
Sankoff, G. 2019. Language change across the lifespan: Three trajectory types.  
Sankoff, G. & Wagner 2006. Age-grading in retrograde movement: The inflected future in Montréal French.  
Sankoff, G. & Wagner 2011. The Long Tail of Language Change. Québécois French Futures in Real Time.  
Sneller & Fisher 2015. When GET got noticed: The emerging salience of GET passives.  
Stefánsdóttir & Ingason 2018. A high definition study of syntactic lifespan change.  
Wood 2011. Stylistic fronting in spoken Icelandic relatives.  
Wagner 2012. Age grading in sociolinguistic theory.

---

## An apparent-time study of Daejeon Korean stop laryngeal contrasts

Yoonjung Kang<sup>1,2</sup>, Yiran Gao<sup>2</sup>, Suyeon Yun<sup>3</sup>, Na-Young Ryu<sup>4</sup>

<sup>1</sup>University of Toronto Scarborough, Toronto, Canada. <sup>2</sup>University of Toronto, Toronto, Canada. <sup>3</sup>Chungnam National University, Daejeon, Korea, Republic of. <sup>4</sup>The Pennsylvania State University, State College, USA

**Background:** Korean has three-way laryngeal contrasts of fortis (shortest VOT), lenis (intermediate VOT), and aspirated (longest VOT) stops. Recent studies found a pan-Korean change in progress towards the merger of VOT distinction between aspirated and lenis stops in phrase-initial position while F0 of the following vowel is taking over as the primary cue for aspirated vs. lenis contrast (Silva 2006; Kang et al. 2022 for a recent overview). This change, led by younger females, is most advanced in Seoul Korean. Despite being one of the major dialects, Chungnam Korean, which Daejeon belongs to, has received little attention in instrumental phonetic studies (Figure 1). The current study fills this gap. Ahn (2017), the only previous acoustic study on Daejeon Korean stops, found longer VOT for lenis than aspirated stops in ten speakers (5M, 5F) in their 20s, suggesting the change is similar to or even more advanced in this dialect than reported in Seoul Korean.

**Data and analysis:** We conducted an apparent-time study with 81 speakers of Daejeon Korean born between 1932 and 2003, stratified for age and gender (Table 1). Each participant produced two repetitions of 18 di- or tri-syllabic words with word-initial stops as part of a larger list. Two lexical items were chosen at each of the three places of articulation for each laryngeal category with a high/non-high vowel following the stop. A total of 2,906 tokens (81 speakers \* 3 laryngeal \* 3 places \* 2 vowels \* 2 reps – 10 errors) were analyzed for VOT and f0 (converted to semitone) of the following vowel midpoint. VOT and f0 were normalized using by-speaker z-score transformation to eliminate speaker-specific speech rate and pitch-level and range effects. Statistical analysis focused on the contrast between aspirated and lenis stops. Linear mixed-effects regression models were used to test the effects of age (old, young), gender (M, F), and laryngeal type (len, asp) on F0 and VOT. Random effects of speaker and word were included.

**Results:** Daejeon Korean is losing the VOT distinction between aspirated and lenis stops. Our results show that this change is more advanced in younger than older and in female than male speakers’ speech (Figure 2, Table 2a). Post-hoc tests (not shown) show the difference is statistically significant only for the older male speakers. On the other hand, f0 is a robust cue across all speaker groups and more so for younger and female speakers (Figure 3, Table 2b). The three-way interaction in the f0 model shows that the expansion of f0 distinction across age is slowing down in the female speech, likely due to this change reaching the near-end state. This pattern is consistent with the diachronic change in Seoul dialects. F0, which used to be a redundant cue, has been enhanced to a primary cue (40 Hz difference on average), and as the f0 distinction is being established, the original VOT cue is allowed to merge fully and demonstrates that similar to other non-pitch accent dialects, Daejeon Korean is undergoing a restructuring of stop laryngeal contrasts.

Ahn, Mee-Jin. 2017. Prosodic Effects on Acoustic Cues for the Korean Stop Contrast: Evidence from Daejeon Korean. 언어과학연구 82: 177-198.



VOT merger in progress and speech rate accommodation in perception:  
a case study of Daejeon Korean

Yoonjung Kang<sup>1,2</sup>, Suyeon Yun<sup>3</sup>, Na-Young Ryu<sup>4</sup>

<sup>1</sup>University of Toronto Scarborough, Toronto, Canada. <sup>2</sup>University of Toronto, Toronto, ON, Canada. <sup>3</sup>Chungnam National University, Daejeon, Korea, Republic of. <sup>4</sup>The Pennsylvania State University, State College, PA, USA

**Background:** Speech rate variation can obscure duration-based contrasts. Perception studies have found that listeners can modulate their perception, compensating for the contextual speech rate (Miller, 1987). For example, English listeners are more likely to perceive a VOT value as a long-lag stop in fast speech. In this study, we examine if and how listeners adapt to varying speech rates when a VOT merger is in progress. A case in point is the Korean laryngeal contrasts of fortis (shortest VOT), lenis (intermediate VOT), and aspirated (longest VOT) stops. In phrase-initial position, the VOT distinction between aspirated and lenis stops is merging, while the F0 of the following vowel is taking over as the primary cue (Silva, 2006; Kang et al., 2022 for a recent overview). We investigate how this sound change affects listeners' speech rate compensation in aspirated-lenis contrast perception in Daejeon Korean, undergoing a VOT merger (Figure 1).

**Data and analysis:** Eighty-one speakers of Daejeon Korean born between 1932 and 2003 (Table 1) participated in a perception study. A different set of four Daejeon Korean speakers varying in age and gender (YF, YM, OF, OM) produced the speech materials (Table 2). Participants identified stops after hearing CV stimuli manipulated to vary in VOT (7 steps, 70-130 ms) and F0 (3 steps of 1 semitone interval: Low, Mid, High, level determined based on a pilot study) embedded in a carrier sentence, "문장 맨 마지막 말은 \_\_다." The carrier sentence was manipulated to vary between slow (120% of the mean duration of all speakers) and fast (80%) conditions. A mixed-effects logistic regression model, trimmed via backward stepwise reduction, was used to determine how the cue use for lenis-aspirated perception and the rate effect are modulated by the age and gender of the listeners and the talkers.

**Results:** We found that listeners used both VOT and F0 cues significantly, but the weighting of the cues varied depending on the age and gender of the talkers and the listeners (Table 1a-b, Figure 2). Younger and female listeners were more sensitive to the F0 cue than older and male listeners, reflecting their own group-level production patterns. Listeners were also sensitive to the talkers' age and gender – they relied more on the F0 cue for the young female talker than the old male talker, reflecting the expected pattern for the talker's age and gender. The rate effect was only modulated by the talkers' age and gender (Table 1c, Figure 3). Post-hoc tests find the expected rate effect (more aspirated stop responses in fast speech) only for the older male talker, a group that retains the VOT contrast more robustly. In contrast, a significant rate effect in the opposite direction was found for the young female talker, who belongs to a group where the VOT distinction has merged and possibly reversed (cf. Ahn (2017) reported longer VOTs for lenis than aspirated stops.) The results are significant in showing the way listeners compensate for speech rate reflects the contrastive status of the duration cue undergoing a sound change.

Ahn, Mee-Jin. 2017. Prosodic Effects on Acoustic Cues for the Korean Stop Contrast: Evidence from Daejeon Korean. *언어과학연구* 82: 177-198.

Kang, Yoonjung, Jessamyn Schertz, and Sungwoo Han. 2022. Phonology and Phonetics. In *Korean: A Linguistic Introduction*, ed. J. Whitman & S. Cho, 63-95. Cambridge: Cambridge University Press.

Miller, J. L. 1987. Rate-dependent processing in speech perception. In Ellis, A. W. (ed.), *Progress in the psychology of language*. London: Erlbaum, 119-157.

Silva, David J. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23: 287-308.

---

"I could probably assume what you did": Characterological creation in improvised comedy

Nour Kayali, Ian Schneider  
University of Kentucky, Lexington, USA

In improvised comedy (improv), actors build identities and social worlds with only their words, bodies, and limited time, and without the luxuries of planning and revision (Goodman, 2008; Stewart, 2016). These elements combine to make improv a fruitful yet overlooked data source in sociolinguistic inquiry, particularly in relation to indexical fields (Eckert, 2008) and characterological enregisterment (Agha, 2007; Johnstone, 2017). To bridge this gap, this paper examines the role of indexical fields in improvised character creation.

Data for this study came from improv scenes performed by four participants. Participants had a range of improv and acting experience, with two noting no experience and two with formal and practical theatre training (e.g., acting in a play, working as a theatre director). In one subset of data, participants secretly drew occupations as character prompts (e.g., “Banker”, “Restaurant Server”) and performed solo scenes where the character returned home from work. After a scene, the audience (consisting of the other participants) guessed what career the actor presented; the audience guessed the character’s occupation with 100% accuracy. After performing all scenes, we conducted a 35-minute group interview (Schilling, 2014), discussing scene-building decisions and interpretations of other participants’ performances.

We analyzed each scene alongside group interview metacommentary, focusing on discussion of associations the actors applied in character creation and cues the audience noted. From this, we propose a model of how actors and audience develop mutual indexical frames through spectator mentalization (Brown, 2022) (Figure 1). Once an actor draws an occupation, they derive an indexical field of that occupation (Figure 2), from which they draw elements they view as legible to an audience, or what Halpern et al. (1994) call “shared consciousness.” For example, in Excerpt 1, an actor (Rosalina) attempts to embody these elements as a character (Restaurant Server) through lexical choices (bolded), situational-emotional characteristics (italicized), and constructed relationships (underlined).

The audience also interprets the actor’s embodied character and generates a corresponding indexical field. This process is twofold. The audience (Excerpt 3) interprets the actor’s character (Burt’s Banker, Excerpt 2) choices. With this data, the audience generates an indexical field of that character, drawing from their own prior knowledge and associations (Excerpt 3), such as a distant relationship to family (Excerpt 2, italicized and underlined). As the actor continues their embodied performance, the audience iteratively updates and aligns their indexical field closer to the actor’s characterological field through a process of spectator mentalization (Brown, 2022). The audience provides feedback in the form of laughter or crickets, which informs the actor of the resonance of their characterological decisions.

Improv actors employ associations and indexicalities on a blank stage over a period of seconds to immerse audiences in shared, believable characters and scene-worlds (Lubet & Hankinson, 2006). A sociolinguistic analysis of improv scenes could illuminate oft-observed yet ubiquitous processes of indexical agreement in our day-to-day lives. We find improv especially valuable in confirming the characterizations of personae that are central to work on indexicality and showing the variation inherent in indexical fields.

Agha, A., 2007. ‘Recombinant selves in mass mediated spacetime.’ *Language & Communication*, 27(3), pp.320-335.

Brown, S., 2022. ‘Partnering Up: The Social Cognition of Partnered Interaction in Life and Art.’ *Frontiers in Communication*, 7, 1-10.

Eckert, P., 2008. Variation and the indexical field 1. *Journal of sociolinguistics*, 12(4), pp.453-476.

Goodman, K., 2008. *Improvisation for the spirit: Live a more creative, spontaneous, and courageous life using the tools of improv comedy*. Naperville, IL: Sourcebooks, Inc..

Halpern, C., Close, D. and Johnson, K., 1994. *Truth in comedy: The manual of improvisation*. Englewood, CO: Meriwether Publishing.

Johnstone, B., 2017. ‘Characterological figures and expressive style in the enregisterment of linguistic variety.’ In Montgomery, C. & Moore, M. (eds.) *Language and a sense of place: Studies in language and region*, Cambridge: Cambridge University Press, pp.283-300.

Lubet, S. and Hankinson, T., 2006. ‘In facietis veritas: How improvisational comedy can help trial lawyers get some chops.’ *Texas Review of Entertainment & Sports Law*, 7, 1-13

Schilling, N., 2014. ‘Surveys and interviews.’ In Podesva, P & Sharma, D. (eds.) *Research methods in linguistics*. Cambridge: Cambridge University Press, pp.96-115.

Stewart, C.P., 2016. *Effects of improv comedy on college students*. Normal, IL: Illinois State University.

---

## Ne-omission in Ivorian French celebrity interviews on YouTube

Julianne Kapner

University of California, Berkeley, Berkeley, USA

There has been extensive sociolinguistic research on French, but almost none on African varieties, even as they comprise 44% of French speakers (BBC News 2019). Meanwhile, the worldwide proliferation of social media opens up new data sources for sociolinguistic investigation. The present study focuses on ne-omission, a variable which has been well-studied worldwide but never quantitatively examined in West Africa. Using interview data from YouTube, I aim (a) to address whether the constraints on morphosyntactic variation posited for other varieties of French also apply in Côte d’Ivoire; and (b) to demonstrate the utility of social media as a data source for investigations of less-studied varieties.

Prescriptive French has negative concord, with pre-verbal and post-verbal negative markers. Ne ‘NEG’ precedes the verb, while the marker after the verb is selected from several with varying semantics (Example 1). In informal contexts, ne is optional (Example 2). Ne-omission is relatively consistent across worldwide varieties, including in

France (e.g., Ashby 1981), Switzerland (Fonseca-Greber 2007), Quebec (e.g., Martineau & Mougeon 2003; Poplack & St-Amand 2007; Sankoff & Vincent 1977), and Martinique (Roberts 2014). Ne-omission is less common in formal contexts, as it likely indexes seriousness (Fonseca-Greber 2007; Grieve-Smith 2009; Roberts 2014; Sankoff & Vincent 1977). Higher-frequency (“fixed”) constructions predict more ne-omission (Grieve-Smith 2009; Roberts 2014). It is more common among men than women and may be advancing in apparent time in some communities (Fonseca-Greber 2007; Roberts 2014; cf. Grieve-Smith 2009). Ne-omission has been observed (although not quantitatively) in Ivorian French (Ahua 2009).

The present investigation used 57 celebrity interviews representing a semi-formal style, taken from Abidjanshow Vidéo, a YouTube channel based in Abidjan, Côte d’Ivoire. Twelve speakers were female and forty-five were male; birth years ranged from 1945-2004. Negative polarity utterances were coded for the retention or omission of ne; 850 eligible tokens were included in the analysis.

The overall rate of ne-omission was 56.7%. A fixed effects logistic regression model found decade of birth to be a significant predictor ( $p < 0.05$ ) of ne-omission, indicating that omission may be advancing in apparent time. Verb subject was also a significant predictor of ne-omission, with higher rates of omission with pronominal verb subjects, especially *je* and *ce*, than with other subjects. While the rate of ne-omission was higher in fixed constructions than elsewhere, this difference was not statistically significant, nor was the effect of the identity of the post-verbal negative marker. There was no significant patterning by gender. These patterns are generally consistent with those observed in other varieties of French.

This study finds that even as Ivorian French develops its own phonology and distinctive lexical items (Ahouzi 2014; Boutin 2002; 2018; N’Guessan 2007; 2008; Kouame 2012; Detey et al. 2016), it appears to be mirroring other varieties with regard to sociolinguistic variation in ne-omission. Ivorian French ne-omission therefore likely shares indexicality with the other varieties. I also demonstrate the promise of social media as a new data source to address under-studied language communities in variationist sociolinguistics.

Ahouzi, A. A. 2014. *Français parlé de Côte d’Ivoire*. Paris: L’Harmattan.

Ahua, B. M. 2009. “Mots, Phrases et Syntaxe du Nouchi.” *Le Français En Afrique* 23:16.

Ashby, W. J. 1981. “The loss of the negative particle *ne* in French: a syntactic change in Progress.” *Language* 57(3):674–87. <https://doi.org/10.2307/414345>.

BBC News. 2019. “Why the Future of French Is African,” Apr. 7, 2019. <https://www.bbc.com/news/world-africa-47790128>.

Boutin, B. A. 2002. “Description de la variation: Études transformationnelles des phrases du français de Côte d’Ivoire.” Ph.D. Diss., U. Stendhal – Grenoble III.

———. 2018. “Plurilinguisme et francophonie en Côte d’Ivoire.” In *Aspects linguistiques et sociolinguistiques des français africains*, O. Floque, ed. 101–19. Sapienza U. Editrice.

Detey, S., J. Durand, B. Laks, & C. Lyche, eds. 2016. *Varieties of Spoken French*. Oxford: Oxford U. Press. <https://doi.org/10.1093/acprof:oso/9780199573714.001.0001>.

Etienne, C., and K. Sax. 2009. “Stylistic variation in French: bridging the gap between research and textbooks.” *Mod. Lang. Journal* 93(4): 584–606.

Fonseca-Greber, B. 2007. “The emergence of emphatic ‘ne’ in conversational Swiss French.” *J. Fr. Lang. Studies* 17(November):249–275. <https://doi.org/10.1017/S0959269507002992>.

Grieve-Smith, A. B. 2009. “The Spread of Change in French Negation.” Ph.D. Diss., Albuquerque, N.M.: U. of New Mexico.

Kouame, K. J.-M. 2012. “La langue française dans tous les contours de la société ivoirienne.” Quebec: Observatoire démographique et statistique de l’espace francophone.

Martineau, F., & R. Mougeon. 2003. “A Sociolinguistic Study of the Origins of Ne Deletion in European and Quebec French.” *Language* 79(1):118–52.

N’Guessan, J. K. 2007. “Le français: langue coloniale ou langue ivoirienne ?” *Hérodote* 126(3):69. <https://doi.org/10.3917/her.126.0069>.

———. 2008. “Le français en Côte d’Ivoire: de l’imposition à l’appropriation décomplexée d’une langue exogène.” *Documents pour l’histoire du français langue étrangère ou seconde*, no. 40/41(January):179–197. <https://doi.org/10.4000/dhfiles.125>.

Poplack, S., & A. St-Amand. 2007. “A real-time window on 19th-Century Vernacular French: The ‘Recits du français québécois d’autrefois.’” *Lang. in Society* 36(5):707–34.

Roberts, N. S. 2014. “A Sociolinguistic Study of Grammatical Variation in Martinique French.” Ph.D. Diss., Newcastle U.

Sankoff, G., & D. Vincent. 1977. “L’emploi productif du *ne* dans le français parlé à Montréal.” *Le Français Moderne* 45(3):243–56.

---

## The last L1 Judeo-Spanish speakers of New York City

Daniel Kaufman<sup>1,2</sup>, Ross Perlin<sup>2,3</sup>, Ariel Newman,<sup>3</sup> Diego de Acosta<sup>4</sup>

<sup>1</sup>Queens College/CUNY, <sup>2</sup>Endangered Language Alliance, <sup>3</sup>Columbia U., <sup>4</sup>U. of Minnesota

We present an overview of the *Languages of New York City* digital map (Perlin et al. 2021) and take an in-depth look at the community of Judeo-Spanish speakers based on recently conducted interviews. While the historical development of Judeo-Spanish has been described in detail (Bunis 2015 and refs. therein), less attention has been

paid to the linguistic practices of contemporary L1 speakers in NYC under more recent layers of language contact (see Kraus 1951, Harris 2006, Varol 2006, Romero 2008 for situations outside of NYC and Agard 1950, Harris 1994, 2006 for an earlier look at New Yorkers). The goal here is two-fold: to better understand the distinguishing features of Judeo-Spanish as spoken by the last generation of local L1 speakers, and to understand those aspects of the community that are most problematic for purposes of mapping in the context of our digital map.

Four of the speakers forming our current corpus were born in Izmir/Smyrna (Turkey), Rhodes (Greece), and Tetouan (Morocco), while three were born in the Bronx, as first or second generation Americans. We focus on the speech of these seven speakers with occasional comparisons made to two other recent public audio-visual corpora.

Judeo-Spanish is characterized by the retention of several archaisms from Medieval Spanish (e.g. the \*b/\*v distinction) as well as unique innovations both in the historical phonology (e.g. *r*-metathesis, \*r/\*r merger, excrescent nasals, sporadic \*n>m) as well as the lexicon (e.g. *meldar* 'to read'). Regional dialects share most of these features but are distinguished by contact effects from a large number of co-territorial languages, as well as from more recent contact situations of the 20th century, including French, English, Argentinian Spanish and Caribbean Spanish (in New York). We show that English appears to exert a far stronger influence on NY-born speakers than other Spanish varieties, although there is wide variation by speaker in the extent of contact effects. Despite five centuries of intense multilingualism, we argue that it is the highly conservative nature of Judeo-Spanish, even in the face of language attrition, which remains its most surprising characteristic.

Our second aim focuses on questions of identity and language use. Earlier endonyms, most popularly *Spañol* or *Español*, did not distinguish Judeo-Spanish from other varieties. In New York, however, terms such as *Muestro Spañol* 'our Spanish' came into more popular use for just this purpose, as distinguishing features became newly salient in the local context. Interestingly, interviewees often identified Judeo-Spanish not as a (geographic) dialect of Spanish but rather as a "chronolect", ostensibly frozen in time and standing in contrast to *Español moderno*, a term meant to encompass all other varieties of Spanish. Several speakers furthermore view themselves as the last speakers of their *language* despite being esconced in Spanish-speaking communities, in distinction to younger language activists participating in online revitalization efforts (Brink-Danan 2011). We conclude with a reflection on the lessons of NYC Ladino and its community of speakers for the larger language mapping project.

Agard, Frederick B. 1950. "Present-Day Judaeo-Spanish in the United States". *Hispania* 33(3):203-210. Brink-Danan, Marcy.

2010. The meaning of Ladino: The semiotics of an online speech community. *Language & Communication* 31:107-118.

Bunis, David M. 2015. Judezmo (Ladino). In Kahn, Lily and Rubin, Aaron D. (eds.), *Handbook of Jewish languages*, 366–451. Leiden: Brill.

Harris, Tracy. 1994. *The Death of a Language: the History of Judeo-Spanish*. Newark: University of Delaware Press.

Harris, Tracy. 2006. "The sociolinguistic situation of Judeo-Spanish in the 20th century in the United States and Israel" *Revista Internacional de Lingüística Iberoamericana*, 4(2): 115-133

Kraus, Karl. 1951. Judeo-Spanish in Israel. *Hispania* 34(3):261-270.

Perlin, Ross, Daniel Kaufman, Jason Lampel, Maya Daurio, Mark Turin, Sienna Craig, eds., *Languages of New York City* (digital version), map. New York: Endangered Language Alliance. (Available online at <http://language-map.nyc>, Accessed on 2021-04-15.)

Varol, Marie-Christine. 2006. "The sociolinguistic situation of Judeo-Spanish in the 20th century in the United States and Israel" *Revista Internacional de Lingüística Iberoamericana*, 4(2): 99-114.

Romero, Reynaldo Jr. 2008. Structural Consequences of Language Shift: Judeo-Spanish in Istanbul. Ph.D. diss, Georgetown University.

---

## Analyzing the performance of male sexuality in Sevillian Spanish through remote speech data collection

Caleb Kelley

University of Alabama, Tuscaloosa, USA

Recent scholarship in linguistics, social psychology, and other related disciplines has demonstrated a significant interest in the topic of gay-sounding speech. Various work on sociophonetic variation has identified sibilants (i.e., /s/-like consonants), pitch, and vowel formants as being the most commonly correlated features with gay-sounding male speech (Duarte, 2022; Ezquerra, 2015; Gaudio, 1994; Kachel et al., 2018; Mack, 2010, 2011; Tracy et al., 2015; Zimman, 2017). However, a majority of this research has been conducted in English and there remains a significant lack of representation of LGBTQ+ voices within the field of Spanish sociophonetics. Some of the recent studies that focus on the social meaning of language have highlighted the importance of understanding sociolinguistic personae, which are holistic social types and styles linked with particular ways of being and speaking, in understanding how individuals convey and receive social meaning in interactional contexts (D'Onofrio, 2018, 2020). In light

of this, this study builds on the construct of “gay” and “straight” personae in sociolinguistic styles to investigate Spanish syllable-final /s/ variation from a third-wave (Eckert, 2016) perspective. In the present study a performance-based methodology was adopted, following Russell (2015, 2017). An online survey was conducted among 13 cisgender, college-educated male Spanish speakers (5 self-identified LGBTQ+ and 8 non-LGBTQ+) from Seville, Spain, of varying ages. The participants provided socio-demographic data, including age, origin, and sexual orientation, and then recorded themselves reading passages according to three personae tasks.

Reading tasks (i.e., short introductory, scientific, and narrative texts) each elicited speaker production of three distinct stylistic personae: “gay” and “straight” personae, as well as “neutral” (or baseline) speech. Speakers were initially asked to read the passages in their natural voice. Subsequently, they were instructed to read the passages as if they were attempting to convey both a gay and straight persona, akin to an actor preparing for a role. The recordings were analyzed acoustically, and tokens of syllable-final /s/ were coded for various social and linguistic factors. Measurements of coda /s/ duration and spectral center of gravity (COG) were taken in Praat using scripts.

Due to the COVID-19 pandemic, the data used in this study were collected remotely via an online Qualtrics survey with a recording plugin as in-person data collection was not feasible due to travel restrictions. Preliminary results show significant interactions between reading persona task and self-identified sexual orientation for coda /s/ duration and spectral center of gravity (COG), revealing potential connections to gay speech stereotypes in this Spanish-speaking region. The goal was to enhance our comprehension of how male sexuality is performed in Spanish by utilizing an innovative digital tool for corpus analysis, showcasing a new methodological design for conducting sociolinguistics research in the digital age.

D’Onofrio, A. (2018). Personae and phonetic detail in sociolinguistic signs. *Language in Society*, 47(4), 513–539. <http://dx.doi.org.libdata.lib.ua.edu/10.1017/S0047404518000581>

D’Onofrio, A. (2020). Personae in sociolinguistic variation. *WIREs Cognitive Science*, 11(6). <https://doi.org/10.1002/wcs.1543>

Duarte, J. (2022). Sociophonetic differences in queer speech of Spanish speakers [Undergraduate Honors Thesis, University of California, Berkeley]. <https://escholarship.org/uc/item/3009g0js>

Eckert, P. (2016). Third wave variationism. In *Oxford Handbook Topics in Linguistics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199935345.013.27>

Ezquerro, Á. O. (2015). The impact of sexual orientation on the pronunciation of stressed vowels in Peninsular Spanish: An acoustic analysis. *Sociolinguistic Studies*, 9(1), 137–150. <https://doi.org/10.1558/sols.v9i1.18395>

Gaudio, R. P. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech*, 69(1), 30. <https://doi.org/10.2307/455948>

Kachel, S., Radtke, A., Skuk, V. G., Zäske, R., Simpson, A. P., & Steffens, M. C. (2018). Investigating the common set of acoustic parameters in sexual orientation groups: A voice averaging approach. *PLOS ONE*, 13(12), e0208686. <https://doi.org/10.1371/journal.pone.0208686>

Mack, S. (2010). Perception and identity: Stereotypes of speech and sexual orientation in Puerto Rican Spanish. In C. Borgonovo, M. Español-Echevarría, & P. Prévost (Eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium* (pp. 136–147). Cascadilla Proceedings Project.

Mack, S. (2011). A sociophonetic analysis of /s/ variation in Puerto Rican Spanish. In L. Ortiz-López (Ed.), *Selected Proceedings of the 13th Hispanic Linguistics Symposium* (p. 14). Cascadilla Proceedings Project.

Tracy, E. C., Bainter, S. A., & Satariano, N. P. (2015). Judgments of self-identified gay and heterosexual male speakers: Which phonemes are most salient in determining sexual orientation? *Journal of Phonetics*, 52, 13–25. <https://doi.org/10.1016/j.wocn.2015.04.001>

Zimman, L. (2017). Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society*, 46(3), 339–370. <https://doi.org/10.1017/S0047404517000070>

---

## Inter- and intra-speaker /ai~ei/ variation in ‘ōlelo Hawai‘i: Acoustic and historical-comparative evidence

Thomas Kettig

York University, Toronto, Canada

‘Ōlelo Hawai‘i, the Austronesian language indigenous to the Hawaiian Islands, uses a transparent, highly regular orthography. However, a small number of words spelled with *ai* may be pronounced instead as *ei*. These exceptions have been impressionistically noted (Newbrand 1951; Kinney 1956; NeSmith 2005) without aid of acoustic analysis. Given a heavy reliance on orthography and few L1 models for learners acquiring this endangered language, identifying early 20th century phonetic norms and variation is of particular

pedagogical as well as linguistic interest. This study of 1970s conversational interviews with elders finds a 96% rate of /ei/ pronunciation for kinship terms spelled with (e.g. *kaikua‘ana* “older sibling”, *kaikunāne* “brother”; see Table 1). More complex patterns of inter- and intra-speaker variation are found for *maika‘i* (“good”), which exhibits some intermediary reflexes, and *laila* (“there”), which exhibits /i/ as a common variant. Gradient spectral reduction in high-frequency content and function words may be occurring, and cognates in related languages reveal that some of this variation may predate the 19th-century development of Hawaiian orthography.

This study considers interviews with eight native speaker elders recorded in 1972–1973 on the radio program *Ka Leo Hawai‘i* (KLH) (Kimura 1972-3): the sample comprises one male and one female speaker each from the four largest islands of *Kaua‘i*, *O‘ahu*, *Maui*, and *Hawai‘i*. F1/F2 measurements for all /ei/ and /ai/ tokens were extracted at nine time points using Fast Track (Barreda 2021). Unlike the purely auditory studies of previous literature, this research additionally considers visually plotted vowel trajectories to investigate how these words compare to speakers’ typical (orthographic) /ei/ and /ai/. 183 tokens of the target words were identified – a relatively substantial dataset considering the critically endangered nature of these dialects.

Table 1 summarizes the number of tokens coded as /ai/, /ei/, or other based initially on the author’s auditory impressions. All five speakers whose transcripts contained kinship terms consistently used /ei/, with just one token more like /ai/. *Laila* was pronounced fairly consistently as /ei/ by four speakers, while others alternated with /i/ and /ai/. Speakers also differed in their treatment of *maika‘i*, with only three exclusively using /ai/ and several instances of auditorily ambiguous or reduced trajectories (cf. Figure 1).

The /ei/ and /i/ tokens observed in *laila* might indicate severe undershoot of /ai/ typical of high-frequency words in Hawaiian (Kettig 2023). However, /ei/ is the majority form in *laila* cognates such as *reira* in nearly all closely related languages, including Māori, Tahitian, and Marquesan, reflecting Proto-Central-Eastern-Polynesian \**reira* (Greenhill & Clark 2011); the rarer but orthographically standard /ai/ is thus an innovative Hawaiian form. While the standard orthography of *maika‘i* accurately reflects Proto-Eastern-Polynesian \**maqitaki*, several languages have *meitaki*; this instability in vowel assignment may thus be widespread and several centuries old. Finally, *kai-* reflects Proto-Central-Eastern-Polynesian \**tai-*, with /ai/ reflexes reported in all related languages; the shift to /ei/ in the present Hawaiian sample therefore appears to be a more recent or geographically restricted innovation, possibly motivated by analogy with the high-frequency term *keiki* (“child”).

Barreda, S. (2021). Fast Track: Fast (nearly) automatic formant-tracking using Praat. *Linguistics Vanguard* 7(1).

Greenhill, S.J., & Clark, R. (2011). POLLEX-Online: The Polynesian Lexicon Project Online. *Oceanic Linguistics* 50(2), 551-559.

Kettig, T. (2023). Word class and frequency effects in Hawaiian stressed vowel clusters. *ICPhS 2023*.

Kimura, L. L. K. (1972–1973). (Producer). *Kani‘āina*, *ulukau.org* Digital Repository of *Ka Haka ‘Ula O Ke‘elikōlani* College of Hawaiian Language, University of Hawai‘i at Hilo, *Ka Leo Hawai‘i*. Retrieved from: *ulukau.org/kaniaina*.

Kinney, R. K. (1956). A non-purist view of morphomorphemic variations in Hawaiian speech. *Journal of the Polynesian Society* 65(3), 282-286.

NeSmith, R. K. (2005). *Tūtū’s* Hawaiian and the Emergence of a Neo Hawaiian Language. *‘Ōiwi Journal* 3–A Native Hawaiian Journal.

Newbrand, H. L. (1951). A phonemic analysis of Hawaiian. MA thesis, University Hawai‘i at Mānoa.

*Do You Have Any Pain?: Linguistic and Social Variables’ Impact on Doctor Perceptions and Treatments*

Raef Khan

Northwestern University, Evanston, USA

Perceptions of voices have been shown to correlate with attitudes about a speaker, including ratings of truthfulness, credibility, and intelligibility [e.g., 1-2]. Linguistic profiling has real-world effects on housing, employment, and judicial outcomes [e.g., 3-4]. However, it is unclear how a speaker’s perceived identity based on their voice, combined with their pragmatic choices, such as using modifiers (e.g., “extremely”), influences listener attitudes and affects real world outcomes. Similarly, evidence shows that pragmatic inferencing can be socially mediated [e.g., 5], but less is known about how social information, specifically race and gender, interacts with pragmatic inferences related to these choices. This research aims to understand how the interaction between social information and pragmatic choice affects listener perceptions and real-world outcomes.

I propose to study this interaction by probing doctor judgments and treatment options for patients presented with different voices and different uses of modifiers. Racial and gender biases also occur in the medical field, especially for women and Black patients, compared to men and White patients, including in attitudes and stereotypes [e.g., 7-

8], perceptions of pain [e.g., 9-10], treatment outcomes [e.g., 11], patient descriptions [e.g., 12-13], and doctor-patient interactions [e.g., 14]. It has been shown that these populations may need to express their symptoms more intensely (e.g., greater use of modifiers) in order to be heard and treated [e.g., 15-16]. However, little research has been conducted on the impact of linguistic variables on medical biases, particularly in relation to race and gender. This study therefore seeks to understand how doctor judgements and treatment decisions are affected by a patient's perceived social characteristics, cued by voice differences and their choice of using modifiers when describing symptoms.

In this proposed research, medical students in their third year will receive an audio recording of a patient describing their symptoms. Recordings will be manipulated in a 2x2x2 design for perceived voice gender (men and women), race (Black and White), and symptom descriptions (a bare statement and a positively modified one; e.g., “It's painful” vs. “It's extremely painful”), and presented using a matched guise task with a Latin square design. Participants will be asked to indicate their attitudes toward the patient on a 5-point Likert scale (1: negative judgment, 3: neutral judgment, 5: positive judgment), explain general reasons for these attitudes, rate participants' level of pain, and provide potential treatment options. While positive modifiers are theoretically meant to limit the scope of interpretation to strictly greater than the baseline [17], the degree to which doctors infer modifiers can vary based upon social and linguistic biases.

Though previous research would predict stereotyping and treatment differences between speakers of different perceived races and genders, it is not fully known how the speaker's usage of modifiers impacts these outcomes. This research aims to further improve our understanding of how pragmatic choice affects linguistic bias, how social information impacts inferences about statements, and to better inform factors of medical biases.

1. Chao, M., & Bursten, J. R. (2021). Girl talk: Understanding negative reactions to female vocal fry. *Hypatia*, 36(1), 42-59.
2. King, S., Vaughn, C., & Dunbar, A. (2022). Dialect on Trial: Raciolinguistic ideologies in perceptions of AAVE and MAE codeswitching. *University of Pennsylvania Working Papers in Linguistics*, 28(2), 7.
3. Baugh, J. (2016). Linguistic profiling and discrimination. *The Oxford handbook of language and society*, 349-368.
4. Rickford, J. R., & King, S. (2016). Language and linguistics on trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond. *Language*, 948-988.
5. Beltrama, A. (2020). Social meaning in semantics and pragmatics. *Language and Linguistics Compass*, 14(9), e12398.
6. Hunt, M., & Acton, E. K. (2022). “How's the wife?”: Pragmatic reasoning in spousal reference. *Journal of Pragmatics*, 188, 152-170.
7. Beach, M. C., Saha, S., Park, J., Taylor, J., Drew, P., Plank, E., ... & Chee, B. (2021). Testimonial injustice: linguistic bias in the medical records of black patients and women. *Journal of general internal medicine*, 36(6), 1708-1714.
8. Madeira, F., Costa-Lopes, R., Do Bú, E. A., & Tato Marinho, R. (2022). The role of stereotypical information on medical judgements for black and white patients. *Plos one*, 17(6), e0268888.
9. Samulowitz, A., Gremyr, I., Eriksson, E., & Hensing, G. (2018). “Brave men” and “emotional women”: A theory-guided literature review on gender bias in health care and gendered norms towards patients with chronic pain. *Pain Research and Management*, 2018.
10. Trawalter, S., Hoffman, K. M., & Waytz, A. (2012). Racial bias in perceptions of others' pain. *PloS one*, 7(11), e48546.
11. Schulman, K. A., Berlin, J. A., Harless, W., Kerner, J. F., Sistrunk, S., Gersh, B. J., ... & Escarce, J. J. (1999). The effect of race and sex on physicians' recommendations for cardiac catheterization. *New England Journal of Medicine*, 340(8), 618-626.
12. P Goddu, A., O'Conor, K. J., Lanzkron, S., Saheed, M. O., Saha, S., Peek, M. E., ... & Beach, M. C. (2018). Do words matter? Stigmatizing language and the transmission of bias in the medical record. *Journal of general internal medicine*, 33, 685-691.
13. Kelly, J. F., & Westerhoff, C. M. (2010). Does it matter how we refer to individuals with substance-related conditions? A randomized study of two commonly used terms. *International Journal of Drug Policy*, 21(3), 202-207.
14. Street Jr, R. L., Gordon, H., & Haidet, P. (2007). Physicians' communication and perceptions of patients: is it how they look, how they talk, or is it just the doctor?. *Social science & medicine*, 65(3), 586-598.
15. Bever, L. (2022, Dec. 3) From heart disease to IUDs: How doctors dismiss women's pain. *Washington Post*. <https://www.washingtonpost.com/wellness/interactive/2022/women-pain-gender-bias-doctors/>
16. Rao, V. (2020, July 27). “You are not listening to me”: Black women on pain and implicit bias in medicine. *Today*. <https://www.today.com/health/implicit-bias-medicine-how-it-hurts-black-women-t187866>
17. Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and philosophy*, 30, 1-45.

---

## Myriad orientations: Sexuality and locality in Manx English

Elijah King, Erez Levon

University of Bern, Bern, Switzerland

Regional dialect features do more than reflect a speaker's geographical background. Research has shown how they become indexically linked to local identities and the characterological traits associated with them (Britain 2009;

Labov 2012). By examining whether communities of speakers participate in regional patterns of variation and change, we can therefore track how individuals use language to construct local orientations and identities and gain insight into the social meanings of regional variables (Moore & Carter 2015; D’Onofrio & Benheim 2020).

In this paper, we draw on this understanding of regional variation to investigate speech among members of the queer community on the Isle of Man. A self-governing British Crown Dependency located off the northwest coast of England, the Isle of Man is a historically very conservative island. As a result, some Manx queer individuals distance themselves from identifying with the island and instead orient to queer-friendly urban centres on the English mainland (e.g., Manchester, Brighton), while others do not experience this kind of tension between their gender/sexuality and their local identity. We investigate how this situation affects the use of local variable patterns among queer speakers, focusing specifically on how local orientation constrains the appearance of Manx English features.

Data is drawn from 16 hours of sociolinguistic interviews with 14 members of the Manx queer community. Traditionally, Manx English shares many similarities with Northern English varieties, including a merger of the FOOT and STRUT lexical sets (Hamer 2007). Yet large-scale migration to the island in the twentieth century (the island’s population doubled between 1960-2000) brought large numbers of speakers of other English varieties, introducing the FOOT/STRUT split, among other variables, into the local feature pool. We investigate the extent to which the contrast between merged and unmerged FOOT/STRUT has taken on a local Manx indexical meaning and how queer speakers use this variable to present local vs. non-local selves. To test this, we extracted Lobanov normalized F1 and F2 values of 4,480 tokens of FOOT and STRUT from the interviews, divided between talk on “queer” topics (e.g., local queer life) and other topics. Analyses investigate FOOT/STRUT quality as a function of topic and local orientation, as well as standard internal factors (Jansen & Braber 2020; Turton & Baranowski 2021).

Initial results indicate an effect of local orientation on FOOT/STRUT realization. Speakers with stronger local orientations show less overlap between FOOT and STRUT (average Bhattacharyya’s Affinity = 0.829), than those with weaker orientations (average = 0.926). In addition, the divergence between FOOT and STRUT appears to be driven by the fronting of FOOT rather than the lowering of STRUT, a pattern distinct from other Northern English varieties (e.g., Strycharczuk et al. 2019) but that parallels a shift taking place across the English South and Midlands (Jansen & Braber 2020). We discuss the relevance of these patterns to the construction of local vs. non-local orientations among queer Manx speakers and what they can tell us about the status of the FOOT/STRUT merger more broadly (cf. Turton & Baranowski 2021).

- Britain, David (2009). Language and space: The variationist approach. In Peter Auer & Jürgen Schmidt (eds.), *Language and Space: An international handbook of linguistic variation*, 142–63. Berlin: de Gruyter.
- D’Onofrio, Annette & Benheim, Jaime (2020). Contextualizing reversal: Local dynamics of the Northern Cities Shift in a Chicago community. *Journal of Sociolinguistics* 24(4):469–91.
- Hamer, Andrew (2007). English on the Isle of Man. In David Britain (ed.), *Language in the British Isles*, 171–75. Cambridge: Cambridge University Press.
- Jansen, Sandra & Braber, Natalie (2021). FOOT-fronting and FOOT-STRUT splitting: Vowel variation in the East Midlands. *English Language and Linguistics* 25(4):767–97.
- Labov, William (2012). What is to be learned: The community as the focus of social cognition. *Review of Cognitive Linguistics* 10(2):265–93.
- Moore, Emma & Carter, Paul (2015). Dialect contact and distinctiveness: The social meaning of language variation in an island community. *Journal of Sociolinguistics* 19(1):3–36.
- Strycharczuk, Patrycja; Brown, Georgina; Leemann, Adrian; & Britain, David (2019). Investigating the FOOT-STRUT distinction in Northern Englishes using crowdsourced data. *Proceedings of the 19th International Congress of Phonetic Sciences*, 1337–41. Australasian Speech Science and Technology Association.
- Turton, Danielle & Baranowski, Maciej (2021). Not quite the same: The social stratification and phonetic conditioning of the FOOT-STRUT vowels in Manchester. *Journal of Linguistics* 57(1):163–201.

---

“We work hard and get nothen for it”[1]: Regional variation in pin/pen merger in a post-Reconstruction corpus [1] Reverend John Jacobs, Escambia County, Alabama, October 7, 1879  
Mary Kohn, Cassidy Hartig  
Kansas State University, Manhattan, USA

In 1879 African Americans from across the South began writing Governor John St. John in Kansas looking for refuge from waves of terrorism that followed Reconstruction (Painter 1977). Many of these authors had limited access to education, driven by extreme and at times life-threatening circumstances to communicate through writing. Historical data on African American Language (AAL) is scarce, but letters by semi-literate authors can provide insight into the history, regional variation, and development of regional southern American English and varieties associated with African American communities (Siebers 2019, Elspass 2012, Van Herk & Walker 2005). We analyze a unique data set from 1879 from the post-Reconstruction South, representing a period after the antebellum



documents analyzed by Montgomery and Eble (2004) and the Civil War Veteran letters explored by Brown (1990), but well before the WPA-era recordings of elderly formerly enslaved individuals (for a review, see Schneider 2015). Our work offers new insight into theories of southern English variants and regionalization in early varieties of AAL by filling in gaps in 19th century data.

The digitally available corpus includes eighteen folders of Governor St. John's received correspondence, comprising 2638 images of original letters and a 548-page pdf transcript. Each letter includes the county of origin, author's name, and date of authorship. Our initial analysis indicates that approximately two-thirds of the letters were written by African American authors. From this corpus, we selected 40 letters by semi-literate authors who self-identified as African American (ex: I am a coler man; H. Swift, Mobile County). We focus on the pin/pen merger as previous evidence suggests that this phonological feature may have emerged during the last quarter of the 19th century (Brown 1990). We selected forty letters, ten each from the states of Alabama, Louisiana, Mississippi, and Texas, resulting in approximately 11,000 words under analysis. From this selection, a total 346 potential environments for merger were identified, with 11% of the data showing evidence of merger through phonological or inverse spellings.

The majority of merged pin/pen tokens were produced by ten authors, but there is at least one example of the merger from each state. While merged tokens appeared in stressed environments for two authors out of the forty analyzed, the vast majority were present in unstressed environments. These patterns are surprisingly similar to findings in Brown (1990). This evidence suggests that the pin/pen merger was largely idiosyncratic, mostly confined to unstressed syllables, and regionally diffuse in our corpus of letters from the lower South. Such findings provide little evidence to support Montgomery and Eble's (2004) hypothesis that the pin/pen merger began phonologizing to stressed environments among African Americans in the Deep South in the later part of the 19th century; but, instead, confirm Brown's (1990) results suggesting that pin/pen mergers were rare and regionally diffuse at this time. While the pin/pen merger is widespread in contemporary varieties of AAL, we do not have convincing data that early AAL is the source of this variant.

Brown, Vivian. 1990. The social and linguistic history of a merger: /ɪ/ and /ɛ/ before nasals in southern American English. PhD Dissertation, Texas A&M University.

Elspace, Stephen. 2012. "The Use of Private Letters and Diaries in Sociolinguistic Investigation." The Handbook of Historical Sociolinguistics edited by Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, Wiley, pp. 156-169.

Painter, Nell Irvin. 1977. Exodusters: Black Migration to Kansas after Reconstruction. W. W. Norton & Company, Inc.

Montgomery, Michael, and Connie Eble. 2004. "Historical perspectives on the pen/pin merger in Southern American English." Studies in the History of the English Language II: Unfolding Conversations, edited by Anne Curzan and Kimberly Emmons, Mouton de Gruyter, pp. 415-434.

Siebers, Lucia. 2019. "[T]his is all answer soon": African American vernacular letters from the eighteenth and nineteenth centuries." Keeping in Touch: Emigrant letters across the English-speaking world Vol. 10, edited by Raymond Hickey, John Benjamins Publishing Company, pp. 213-237.

Schneider, E. W. 2015. Documenting the history of African American Vernacular English. The Oxford handbook of African American language, edited by Jennifer Bloomquist, Lisa Green, and Sonja Lanehart, Oxford, 125-139.

Van Herk, Gerard and James A. Walker. 2005. "S marks the spot? Regional variation and early African American correspondence." Language Variation and Change (17) 113-131.

## Double Negation in Russian: Variation in the functional structure of PPs

Alex Kohut<sup>1</sup>, Clara McMahon<sup>2</sup>

<sup>1</sup>Boston University, Boston, USA. <sup>2</sup>CUNY Graduate Center, New York, USA

As a strict negative concord (NC) language, Russian is predicted to not permit double negation (DN) readings within a single clause [4, 7]. However, certain freestanding Negative Concord Items (NCIs) allow for DN readings in small clause predicates and PP complements [3]. One analysis predicts that while certain NCIs are ambiguous between a DN and SN (single negation) reading in small clause predicates, PPs containing NCIs have a DN or SN reading, depending on word order [3]. However, recent research shows at least one NCI-containing PP (*ni za chto*) under sentential negation is ambiguous between a DN and an SN reading, as in (1) [6].

(1) On **ne** otdal svoju zhizn' **ni za chto**.

He **NEG** gave.perf his life **for nothing**.

DN/SN: 'He did not give his life for nothing/DN/anythingSN.'

The present study expands upon previous work by considering the PP *ni za chto* with a different set of VPs (*otdatj svoj post* 'give up her post', *pročitatj knigu* 'read the book', *podpisatj dogovor* 'sign the contract', *proigratj match* 'lose the match'), and a second NCI-containing PP with a fifth verb (*obraschatjsja ni k komu* 'address nobody'). We conducted a survey of native Russian speakers ( $N = 52$ ;  $Age = 31$ , range = 19-65). Bolded test sentences appeared in contexts compatible with one reading: DN or SN. Participants were asked to rate how well the

sentence fit the context (Contextual Felicity (CF) score) on a 6-point scale (1=strongly disagree, 6=strongly agree). Test sentences consisted of sentential negation, one of five VPs, and an NCI-containing PP. Test sentences varied by aspect (perfective/imperfective) and mood (indicative/subjunctive). Participants were additionally asked to rate four statements to probe their attitude on language variation and change (fig. 1). We also collected information on participants' age, gender, education, and linguistic background.

In contrast to previous work, we found variation with respect to the acceptance of DN readings in Russian [3, 6]. Out of 52 participants, 36 accepted a DN reading in at least one context (assigned a CF score of  $\geq 4$ ): 16 participants did not accept DN readings in any context. This suggests that different grammars exist in Russian: one that accepts DN readings and one that does not. Furthermore, participants seem to show variable sensitivity to the semantic status of the NCI-containing PP. A significantly higher percent of participants accept DN readings of the test item *ne obraschatjsja ni k komu* ('to not address nobody'), where the PP is semantically entailed, compared to each of the other VPs where the PP is not semantically entailed. The sensitivity to the semantic status of the PP contradicts previous claims that the availability of DN readings of PPs is not related to argument/adjunct distinctions [3].

We argue the argument/adjunct distinction is crucial in understanding the availability of DN readings. We propose the locus of variation is the presence or absence of a CP<sup>PP</sup> layer in the functional structure of PPs [2, 3, 5]. Russian NCIs cannot be licensed across CP boundaries [3]. When a CP<sup>PP</sup> layer is present, DN readings result due to CP-intervention effects. We suggest that speakers who accept DN readings prefer CP<sup>PP</sup> in argument positions. Speakers for whom NCI-containing PPs are ambiguous between SN and DN readings optionally select for CP<sup>PP</sup>. For speakers who do not allow any DN readings, CP<sup>PP</sup> is not present; CP-intervention effects do not occur and only SN readings are available.

Lastly, we find that more female than male participants accept DN readings (78% and 62%, respectively). However, male participants rate the linguistic attitude question 'Russian has changed for the better over the past 30 years' significantly higher ( $M = 3.48, SD = 1.05$ )

compared to female participants ( $M = 2.78, SD = 1.2, t(51) = -2.2, p < .05$ ), suggesting potentially more liberal attitudes towards language change (fig. 1). These findings seem to tentatively indicate that whereas variation with respect to negative concord is highly salient and stigmatized by English speakers in the U.S., its counterpart in Russian is not [1].

- [1] Blanchette, F. 2017. Micro-syntactic variation in American English Negative Concord.
- [2] Den Dikken, M. 2010. On the Functional Structure of Locative and Directional PPs. In *Mapping Spatial PPs: The Cartography of Syntactic Structures*.
- [3] Fitzgibbons, N. 2010. Freestanding n-words in Russian: a syntactic account. *Journal of Slavic Linguistics*.
- [4] Giannakidou, A. 2006. N-words and Negative Concord. In *The Blackwell Companion to Syntax*, Vol. 3
- [5] Koopman, H. 2010. Prepositions, Postpositions, Circumpositions, and Particles: The Structure of Dutch PPs. In *Mapping Spatial PPs: The Cartography of Syntactic Structures*.
- [6] McMahon, C. & K. Pember. Accepted. Rethinking Strict Negative Concord: An investigation into double negation readings in Russian.
- [7] Zeijlstra, H. 2008. Negative Concord is Syntactic Agreement.

---

## An analysis of Ukrainians' language attitudes and ideology: A conflict-catalyzed identity shift

Kateryna Kravchenko, Joseph A. Stanley  
Brigham Young University, Provo, USA

On February 24, 2022, the Russian Federation invaded Ukraine. According to Vladimir Putin's speech, the main goal of the war was to protect the Russian speaking population of the Donbas region in eastern Ukraine. Many Ukrainians indeed spoke Russian on a daily basis, largely due to the long and complicated history between the two countries. Recently, there has been an increase in pro-Ukrainian ideologies (Kulyk 2016), however, as the full-scale invasion began, people started advocating for Ukrainian usage in all public domains and many even vowed to never speak Russian again. Since then, we have seen a profound change in language attitudes and usage; for example, Russian-speaking television is prefaced with content warnings and is dubbed in Ukrainian and President Zelenskyy now asks for an interpreter when interacting with Russian speakers.

Following Droogsmas (2017), we examine this conflict-catalyzed change in language ideologies and attitudes in Ukraine and how it facilitates language and identity shift. Our aim is to answer the following research questions: What are the primary motivations for people who decide to switch from speaking Russian to Ukrainian? What is the process like? How do they manage the switch? What aspects/domains are most difficult?

Our results are based on social media posts and semi-structured interviews with Ukrainians who have decided to switch from speaking primarily Russian to using Ukrainian in their daily lives. All interviewees were born and still resided in Ukraine as of February 24, 2022, and varied in age, gender, and education. The Languages of Manukau questionnaire (Taumoeofolau et al, 2002) was used as a base for the sociolinguistic interviews. Specifically, The Demographics, Childhood and Family, Language Usage, and Language Attitudes sections were adapted to fit the linguistic situation in Ukraine. Additionally, we used questions from Droogsma (2017), who studied the effects of violent conflict on language attitudes. Thematic analysis was applied in order to find answers to the research questions.

These interviews illuminate nuance in Ukrainians' shift from Russian to Ukrainian. Pro-Ukrainian ideologies and attitudes appear to be people's motivation for switching. It appears that many people begin their switch with social media, often with a post explicitly saying so. They then begin new relationships with Ukrainian. The final stage of the shift is encouraging friends and family to use Ukrainian with them. Some people make concerted efforts to stick to their vow of only speaking Ukrainian; others make accidental slips, rare exceptions for certain friends or family, or only switch in public domains. It appears to be hardest to switch with people with whom they have spoken exclusively Russian. Interestingly, Surzhyk, a historically stigmatized Ukrainian-Russian mixed language (Bilaniuk 2004), has been reclaimed by Russian-dominant Ukrainians as a helpful bridge towards Ukrainian-dominance.

These results help to document the ongoing language shift in Ukraine, contributing to the existing body of sociolinguistic research about Eastern European languages (Gulida 2009), as well as deepening our understanding of how language attitudes and ideologies change in times of conflict and what implications this process has for speech communities.

Bilaniuk, Laada. 2004. A typology of Surzhyk: Mixed Ukrainian-Russian language. *International Journal of Bilingualism* 8(4). 409–425.

Droogsma, Marieke. (2017). The effect of the violent conflict in Ukraine on language use and language attitudes. MA Thesis, Leiden University.

Gulida, Victoria. 2009. Sociolinguistics in the Russian Federation, Ukraine, Belarus and Moldova. In Ball, Martin J. (ed.) *The Routledge Handbook of Sociolinguistics Around the World: A Handbook*. Taylor & Francis.

Kulyk, Volodymyr. 2016. National Identity in Ukraine: Impact of Euromaidan and the War. *Europe-Asia Studies*, 68(4) 588–608.

Taumoeofolau, Melanite, Donna Starks, Karen Davis, & Allen Bell (2002). Linguists and language maintenance: Pasifika languages in Manukau, New Zealand. *Oceanic Linguistics*, 41(1), 15-27.

---

## Longitudinal change in the adjective intensifier system of Hexagonal French

Scott Kunkel

Queen Mary University of London, London, United Kingdom

Intensifiers are known for their dynamic nature, due in part to the expressive function they serve in speech (Tagliamonte 2008). However, while the quantitative patterning of intensifiers in English has been studied extensively (e.g. Barnfield & Buchstaller 2010, Fuchs 2017, Ito & Tagliamonte 2003), the intensifier system of French has yet to be well documented from a variationist standpoint. This study therefore examines intensifier use from a sociolinguistic perspective in the *Enquêtes Sociolinguistiques à Orléans* (ESLO) corpus of spoken French in Orléans, France. The quantitative distributions of adjectival intensifiers are compared across two time points: the early 1970s (ESLO1) and the early 2010s (ESLO2). Spontaneous speech data from forty-eight total speakers are analyzed, with twenty-four speakers –stratified by age and gender– selected from each corpus.

All intensifiable adjectives were quantified ( $n=8,103$ ) to allow for analyses relating to the rate of intensification across time and social categories. Although a slight decrease in intensification rate was found over time (18.7% in the ESLO1 vs. 16.4% in the ESLO2), results from a mixed-effects logistic regression model show this difference to be non-significant ( $\beta = -.17$ ,  $SE = .11$ ,  $p = .09$ ). On the other hand, a significant difference is found for gender (Figure 1), with female speakers showing higher intensification rates than male speakers at both time points ( $\beta = -.63$ ,  $SE = .26$ ,  $p = .01$ ). This latter finding corroborates studies showing female speakers to use more intensification strategies than male speakers (e.g. Fuchs 2017, Stratton 2020).

Longitudinal analyses of intensifier tokens collected from the ESLO ( $n=1,455$ ) show three patterns of change over time: some intensifiers have decreased in use over time (e.g. *très*, *tellement*), others have increased in usage (e.g. *vraiment*, *tout(e)(s)*), and others show innovative usage between the two corpora (e.g. *super*, *hyper*). One of the most noteworthy findings shows that, whereas *très* dominates as the intensifier of choice among older generations, this intensifier holds less of a monopoly for younger speakers in the ESLO2 corpus (Figure 1). Changes are also found in the grammatical patterning of intensifiers in French. For instance, an increasing preference for intensifiers to occur with adjectives in predicative versus attributive position is found, a shift that has been previously noted as a

correlate of intensifiers becoming increasingly delexicalized (e.g. Barnfield & Buchstaller 2010). However, this trend is also found to vary by intensifier (Figure 2), thus complexifying claims that have previously been made about the implications of adjective position on changes in the intensification system.

Overall, this research provides a view of how the intensifier system of French has changed over the past half century, while also contributing to a more general knowledge of intensifiers as they relate to mechanisms of language change. It also advances a growing body of research that pushes the quantitative study of intensifiers beyond an Anglocentric purview (e.g. Alshaboul et al. 2022, Roels & Enghels 2020, Stratton 2020), thus leading to a larger, cross-linguistic understanding of intensifiers.

- Alshaboul, A., Jarrah, M., Alghazo, S., & Al-Shawashreh, E. (2022). A sociolinguistic analysis of intensifiers in Ammani Arabic. *Ampersand*, 9, 100086.
- Barnfield, K., & Buchstaller, I. (2010). Intensifiers on Tyneside: Longitudinal developments and new trends. *English World-Wide*, 31(3), 252-287.
- Fuchs, R. (2017). Do women (still) use more intensifiers than men? Recent change in the sociolinguistics of intensifiers in British English. *International Journal of Corpus Linguistics*, 22(3), 345-374.
- Ito, R., & Tagliamonte, S. (2003). Well weird, right dodgy, very strange, really cool: Layering and recycling in English intensifiers. *Language in society*, 32(2), 257-279.
- Roels, L., & Enghels, R. (2020). Age-based variation and patterns of recent language change: A case-study of morphological and lexical intensifiers in Spanish. *Journal of Pragmatics*, 170, 125-138.
- Stratton, J. M. (2020). Adjective intensifiers in German. *Journal of Germanic Linguistics*, 32(2), 183-215.
- Tagliamonte, S. A. (2008). So different and pretty cool! Recycling intensifiers in Toronto, Canada. *English Language & Linguistics*, 12(2), 361-394.

---

## I need to, for personal reasons: Shifting forms and functions in American English obligation modals

Dora LaCasse, Jenna McVey, Kevin Vrla  
University of Montana, Missoula, USA

Modal verbs encoding obligation are a longstanding site of variation, and differing patterns in modals' relative frequencies have been reported across dialects of English (e.g. Collins 2005, Smith 2003 among others). While dialects examined differ in frequencies of these modals, one commonality that has emerged is the reduction of modal *must* diachronically, in favor of different quasi-modals, especially *need to*. (Penry Williams & Korhonen 2020 and references therein).

The present study investigates distributions of modals (*have*) got to, *have to*, *must*, *need to*, and *should* diachronically in American English, utilizing a subsection of the Corpus of Contemporary American English (COCA). Data come from the 1990 and 2019 Television Comedy sections. The variable context for these modals is limited to affirmative, simple present declarative sentences (e.g. Tagliamonte & D'Arcy 2007a). Epistemic and formulaic uses of the modals were identified and set aside (N=298). After exclusions, the exhaustive search for these five modals yielded 2197 tokens. Coding was conducted as a class-wide project over several sessions, resulting in high reliability. The relative distributions of these modals, as well as their functions in discourse, are analyzed for both years to track changes in modal use and meaning.

In order to understand how obligation meaning may affect shifting distributions across time, a three-way distinction based on the source of obligation is implemented, distinguishing between Personal choice (1), General circumstances (2), and Hierarchical authority (3) (Travis & Torres Cacoullos 2022).

- (1) But I need to get my beauty sleep, especially if I'm getting my crown tomorrow.
- (2) Well, I have to fly through Denver. There's nothing direct.
- (3) Okay, we'll make lunch, but you kids have to set the table.

Changes in modal distributions across years is significant,  $X^2(4, N = 2197) = 242.8, p = .00001$  (Figure 1). Unsurprisingly, *must* findings follow previous studies, decreasing from 5% to 2%. A marked increase in *should* and *need to* is evident, accompanied by a corresponding decrease in (*have*) got to and *have to*. While a rise in *have to* has been reported for Canadian English (Tagliamonte & D'Arcy 2007a:72), the increase here in *need to* parallels recent findings in Australian (Penry Williams & Korhonen 2020) and UK English (Fehringer and Corrigan 2015:369). The more than four-fold increase in *need to* (6% to 26%) within 29 years may indicate communal change (cf. Labov 1994:84; Tagliamonte & D'Arcy 2007b:213 on *be like*).

These changing modal form distributions are accompanied by changes in modal function, i.e. sources of obligation  $X^2(2, N=2197) = 11.1, p = .0037$  (Figure 2). In particular, Personal obligation rises (74% to 80%). Both *need to* and *should* seem to disproportionately carry this function, with a higher rate of Personal uses for both (86% for each) relative to the overall proportion of Personal obligation in the sample (77%). However, the rise in Personal

obligation rates also holds for (have) got to. This across-the-board change in source of obligation would be consistent with a shift in modal function across the system (Myhill 1995; Travis & Torres Cacoullos 2022).

- Collins, P. C. (2005). The modals and quasi-modals of obligation and necessity in Australian English and other Englishes. *English World-Wide*, 26(2).
- Davies, M. (2008-) The Corpus of Contemporary American English (COCA). Available online at <https://www.english-corpora.org/coca/>.
- Fehring, C. & Corrigan, K. (2015). 'You've got to sort of eh hoy the Geordie out': Modals of obligation and necessity in fifty years of Tyneside English. *English Language and Linguistics*, 19(2): 355-381.
- Labov, W. (1994). *Principles of linguistic change: Internal factors*. Vol. 1, 3 vols. Blackwell.
- Myhill, J. (1995). Change and continuity in the functions of the American English modals. *Linguistics*, 33(2): 157-211.
- Penry Williams, C., & Korhonen, M. (2020). A sociolinguistic perspective on the (quasi-) modals of obligation and necessity in Australian English. *English World-Wide*, 41(3), 267-294.
- Smith, N. (2003). Changes in the modals and semi-modals of strong obligation and epistemic necessity in recent British English. *Topics in English linguistics*, 44, 241-266
- Tagliamonte, S. A., & D'Arcy, A. (2007a). The modals of obligation/necessity in Canadian perspective. *English World-Wide*, 28(1), 47-87.
- Tagliamonte, S., & D'Arcy, A. (2007b). Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change*, 19(2), 199-217.
- Travis, C.E. & Torres Cacoullos, R. (to appear). Form and function covariation: Obligation modals in Australian English. *Language Variation and Change*.

---

## Spanish Vowel and Consonant Contributions to Talker Identification and Lexical Contrast

Stephanie Lain

University of California Santa Cruz, USA

Acoustic properties of the input determine how speech sounds are processed, categorized, and encoded in memory. This information is used to identify words and convey information about the speaker. In perceptual studies conducted with English language prompts, the ability to accurately perform lexical decision or talker identification tasks has been shown to vary depending on whether the participants are attending to consonants or vowels [1-3]. However, given the substantial cross-linguistic variation in how consonants and vowels may function in a system of lexical contrast [4-6] or be used for social indexing [7-8], it is less clear how previous findings apply to languages other than English. For example, regional varieties of English differ substantially in their use of vowel space [9-10], while regional varieties of Spanish differ primarily with regard to stop consonant realization [11-12].

The series of experiments described in this study were undertaken with the goal of clarifying the roles vowels and consonants play in lexical decision making and talker identification in Spanish. Participants in the study were 98 listeners who self-identified as native speakers of Spanish. They were asked to complete a questionnaire to assess their language dominance (Spanish versus English) [13], then perform one of six same-different auditory discrimination tasks. The experiments varied according to task (lexical decision or talker identification) and condition (unaltered stimuli, vowels excised, consonants excised). Stimuli were a set of nonsense words fitting a CV<sub>r</sub>VCV pattern, with stress on the penultimate vowel. All consonants besides [r] were phonologically voiceless stops (/p/, /t/, /k/) with C1 and C2 identical. All vowels in Spanish (/a e i o u/) were represented, with V1, V2, and V3 identical.

Responses from each participant were used to calculate a D prime score (evaluating the participant's ability to discriminate between tokens [14]), as well as a language dominance score. A one-way ANOVA analysis revealed that the responses differed significantly by task and condition. A Tukey post hoc test was used to further break down the differences. The main findings revealed significant differences between the vowels excised condition and the other two conditions but no significant difference between the unaltered stimuli and consonants excised conditions. Results for the lexical decision and talker identification tasks differed significantly in the unaltered stimuli and consonants excised conditions, but not in the vowels excised condition. Discrimination was best in the lexical decision task across all conditions. Only one of the six experiments showed an effect for language dominance.

Findings from the current study largely confirm results from previous studies conducted in English which suggest a greater reliance on consonants when performing lexical decision tasks and vowels when performing talker identity tasks. Variation observed in response to the acoustic properties of vowels and consonants appears to be universal to linguistic processing and not a result of the interaction between speech sounds within a given language system. These results have implications for theories of speech perception, particularly with regard to the role of listener experience in the perception of phonemes and talker-specific acoustic properties.

- [1]Fogerty, D., and Humes, L. 2012. The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences. *Journal of the Acoustical Society of America*, 131(2), 1490-1501.
- [2]Fogerty, D., and Kewley-Port, D. 2009. Perceptual contributions of the consonant-vowel boundary to sentence intelligibility. *Journal of Acoustical Society of America*, 126(2), 847-857.
- [3]Owren, M., and Cardillo, G. 2006. The relative roles of vowels and consonants in discriminating talker identity versus word meaning. *Journal of the Acoustical Society of America*, 119(3), 1727-1739.
- [4]Amengual, M., and Chamorro, P. 2015. The effects of language dominance in the perception and production of the Galician mid vowel contrasts. *Phonetica*, 72, 207-236.
- [5]Cutler, A., Mehler, J., Norris, D., and Segui, J. 1983. A language-specific comprehension strategy. *Nature*, 304(5922), 159-160.
- [6]Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., and van Ooijen, B. 2000. Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory and Cognition*, 28(5), 746-755.
- [7]Floccia, C., Butler, J., Goslin, J., and Ellis, L. 2009. Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research*, 38, 379-412.
- [8]Floccia, C., Girard, F., Goslin, J., and Konopczynski, G. 2006. Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance*, 32(5), 1276-1293.
- [9]Hillenbrand, J., Getty, L., Clark, M., and Wheeler, K. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5), 3099-3111.
- [10]Hughes, A., Trudgill, P., and Watt, D. 2005. *English Accents and Dialects: An Introduction to the Social and Regional Varieties of English in the British Isles*. Hodder Arnold, 4th Ed.
- [11]Lipski, J. 1994. *Latin American Spanish*. New York, NY: Addison, Wesley, Longman, Inc.
- [12]Hualde, J.I. 2005. *The Sounds of Spanish*. New York, NY: Cambridge University Press.
- [13]Birdsong, D., Gertken, L.M., and Amengual, M. 2012. *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. Austin, TX: COERLL; University of Texas at Austin.
- [14]MacMillan, N.A., and Creelman, C.D. 2005. *Detection Theory: A user's guide*. Psychology Press.

---

## Constructions of whiteness in college students' conversations

Grace LeFevre

Northwestern University, Evanston, IL, USA

It is well-established that language and race are co-constructed: speakers use language to construct their racial identities, while their ideologies about race simultaneously influence their language use (Alim et al., 2016; Rosa and Flores, 2017). Consequently, studying linguistic variation within a group requires understanding the ideologies that people in the group hold about their racial identities (among other identities), as well as how they construct those identities linguistically (Eckert, 2008; King, 2020). Many linguistics studies have relied predominantly on white speakers—most often monoracial white undergraduates at R1 institutions in the US—without any explicit discussion of these participants' race. This trend reflects the prevailing ideology that treats whiteness as the unracialized default, associating it with the absence of defining characteristics (Trechter and Bucholtz, 2001) and with standard English varieties (Fought, 2006). Over the past 20 years, relatively little work has investigated white speakers' ideologies about language and whiteness (c.f. Bucholtz, 2001; Kiesling, 2001; Walton and Jaffe, 2011), leaving it less clear how white speakers' racial identities influence their language use and obscuring that whiteness itself is a racialization process constituted through social and language variation (Trechter and Bucholtz, 2001). The proposed study will examine how monoracial white speakers from a Chicago-area university use linguistic resources to ideologize and construct whiteness in conversations with other white speakers.

20 monoracial white, American, English-speaking undergraduate students in the 18-24 age range will interact in 30-minute dyadic conversations in a private laboratory space. To structure their conversations, participants will be instructed to use provided prompts, which will be developed based on a focus group to elicit discussion about participants' constructions of their own racial identities and their ideologies about language and whiteness. After each conversation, both participants will complete a questionnaire: 1) rating their conversation using a series of Likert scales, including conversational quality, connection to partner, and similarity to partner; and 2)

providing qualitative descriptions of their own identities and any speech patterns they associate with those identities.

Using audio recordings and transcripts of these conversations, I will perform qualitative discourse analysis to uncover how the participants linguistically indexed whiteness and/or their racial identities throughout their interactions. Participants could do this in many ways, including using explicit identity labels (Bucholtz and Hall, 2005), situating themselves as part of a white group through references to non-white groups (Kiesling, 2001), and aligning themselves with a higher-order liberal whiteness by expressing awareness of their white privilege (Walton and Jaffe, 2011). I will also analyze quantitative linguistic measurements previously shown to indicate social connection or cooperation, including speech rate convergence (Manson et al., 2013) and response time speed (Templeton et al., 2022). Along with participants' ratings of their conversations, these measurements will enable analysis of how participants' identity-indexing linguistic acts are received by their conversation partner, both initially and over time.

Through facilitating these conversations, this study will contribute to our understanding of white speakers' racial identity construction and the ways that speakers constitute and interpret their own and others' identities in interactions.

- Alim, H. Samy, Rickford, John, & Ball, Arnetha, editors. 2016. *Raciolinguistics: How language shapes our ideas about race*. Oxford University Press.
- Bucholtz, Mary. 2001. The whiteness of nerds: superstandard English and racial markedness. *Journal of Linguistic Anthropology* 11(1):84-100.
- Bucholtz, Mary and Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse studies* 7(4-5):585–614.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4):453-476.
- Fought, Carmen. 2006. Are white people ethnic? Whiteness, dominance, and ethnicity. In *Language and Ethnicity: Key Topics in Sociolinguistics*, 112-132. Cambridge University Press.
- Kiesling, Scott. 2001. Stances of whiteness and hegemony in fraternity men's discourse. *Journal of Linguistic Anthropology* 11(1):101-115.
- King, Sharese. 2020. From African-American Vernacular English to African American Language: rethinking the study of race and language in African Americans' speech. *Annual Review of Linguistics* 6:285-300.
- Manson, Joseph, Gregory Bryant, Matthew Gervais, and Michelle Kline. 2013. Convergence of speech rate in conversation predicts cooperation. *Evolution and Human Behavior*, 34(6):419–426.
- Rosa, Jonathan and Nelson Flores. 2017. Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society* 46(5):621-647.
- Templeton, Emma, Luke Chang, Elizabeth Reynolds, Marie Cone LeBeaumont, and Thalia Wheatley. 2022. Fast response times signal social connection in conversation. *Proceedings of the National Academy of Sciences*, 119(4).
- Trechter, Sara and Mary Bucholtz. 2001. White noise: bringing language into whiteness studies. *Journal of Linguistic Anthropology* 11(1):3-21.
- Walton, Shana, & Alexandra Jaffe. 2011. "Stuff White People Like": Stance, Class, Race, and Internet Commentary. In Crispin Thurlow and Kristine Mroczek, editors, *Digital Discourse*, 287–301. Oxford University Press.

---

## Lexical effects in the change of Cantonese neutral question forms

Margaret Lei, Ivy Kwan

The Chinese University of Hong Kong, Hong Kong

Neutral questions in contemporary Cantonese (a Southern Chinese variety) rely primarily on the use of the [V-neg-VP] form, which involves the juxtaposition of two reduplicated verb phrases and the insertion of a negator in between (Cheung 2007). This interrogative form emerged in the early 20th century, evolving from earlier forms of [VP-neg-V], [VP-neg] and [VP-neg-VP]/[V-neg-V] (Yue-Hashimoto 1993, 2006; Cheung 2001). The adoption of the new form has been observed to vary among different verbs under the effect of lexical diffusion (Yue-Hashimoto 1993).

The present study investigates the morphosyntactic change in the Cantonese neutral question forms comprising of the high-frequency copula verb 'hai6' ("be"). Based on language data of early Cantonese (1828-1931) retrieved from nine Cantonese textbooks in three online databases (Cheung 2012; Yiu 2012, 2019) and contemporary Cantonese (1939-1999) retrieved from 69 films in two online databases (Chin 2012, 2020), the following research questions are addressed:

- (i) What are the neutral question forms adopted by the copula verb 'hai6' ("be") in early and contemporary Cantonese?
- (ii) What are the similarities and differences in the morphosyntactic changes involved in the neutral question forms of 'hai6'?

(iii) How can the variation with ‘hai6’ be accounted for?

Our findings identify a couple of distinct morphosyntactic developments involved in the neutral question forms of ‘hai6’. Unlike other verbs, the most dominant form adopted by ‘hai6’ was [VP-neg-VP] rather than [VP-neg-V] in early Cantonese, constituting 60% (30/50 tokens) of the data in the first period (1828-1902) and 85% (11/13 tokens) in the second period (1906-1931). [VP-neg-V] was the second dominant form in these two early periods (1st period: 40%; 2nd period: 15%), but became non-existent in the next two periods (1939-1964 and 1965-1999). Only two variants of neutral question forms are found in contemporary Cantonese, with [V-neg-VP] being the dominant one (3rd period: 58%; 4th period 60%) followed by [VP-neg-VP] (3rd period: 42%; 4th period: 40%). Among the [VP-neg-VP] tokens, in addition to main verb and tag usages similarly seen in other verbs, two extra variants are found: ‘hai6’ were used in (i) Pseudo-tag questions with the copula verb ‘hai6’ occurring in the main clause and the ‘hai6-neg-hai6’ form in the tag clause; and (ii) Main-plus-tag questions with the ‘hai6-neg-hai6’ form repeatedly occurring in both the main clause and the tag clause.

Taken together, the old [VP-neg-V] form died out earlier in ‘hai6’ than in other verbs, which became totally obsolete in the 1940s with the former but still constituted 26% of the neutral question forms with other verbs (cf. Cheung 2001). Additionally, the switch to the new [V-neg-VP] took place earlier in ‘hai6’ than in other verbs, with the new form dominating the ‘hai6’ tokens since 1940s but constituting only 40% of the data in other verbs. We argue that the high-frequency usage of the copula verb ‘hai6’ in the various tag and tag-like forms, which partially overlap with the structure of the new [V-neg-VP] form, facilitates the process of the morphosyntactic change and lends further support to the role of lexical effects.

- Cheung, Hung-nin. (2001). The interrogative construction: (re)constructing early Cantonese grammar. In H. Chappell (Ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives* (pp.191-231). Oxford: Oxford University Press.
- Cheung, Hung-nin. (2007). *A grammar of Cantonese as spoken in Hong Kong*. Hong Kong: The Chinese University of Hong Kong.
- Cheung, Hung-nin. (2012). *Early Cantonese Colloquial Texts: A Database*.
- Chin, Andy Chi-on. (2012). *A Linguistic Corpus of Mid-20th Century Hong Kong Cantonese (Phase 1)*.
- Chin, Andy Chi-on. (2020). *A Linguistic Corpus of Mid-20th Century Hong Kong Cantonese (Phase 2)*.
- Yiu, Yuk Man Carine. (2012). *Early Cantonese Tagged Database*.
- Yiu, Yuk Man Carine. (2019). *Database of Early Chinese Dialects*.
- Yue-Hashimoto, A (1993). The lexicon in syntactic change: Lexical diffusion in Chinese syntax. *Journal of Chinese Linguistics*, 21(2), 213-254.
- Yue-Hashimoto, A (2006). Syntactic typology in Chinese (part I): the neutral question forms—v-not-v. *Bulletin of Chinese Linguistics*, 1(1), 201-253.

---

## Setting {straight} the record {straight}: Acceptability of alternative word orders in resultatives by heritage Cantonese speakers

Justin Leung

University of Toronto, Toronto, Canada

Heritage speakers are often similar to their homeland counterparts in their syntactic grammar. However, one domain where this generalization does not hold is word order, vulnerable to change by two main forces: transfer from the dominant language and amplification/overgeneralization of existing patterns (Polinsky, 2018). Much work in syntactic transfer/change in heritage languages focuses on situations with variable word orders in the languages in contact; in these cases, structures not present in the dominant language and/or non-canonical structures in the non-dominant language are often reduced in usage among heritage speakers (Lohndal, 2021). Little work has been done on phenomena in which there is little variability and/or no shared word order options between the languages in contact. We conduct a case study on such a phenomenon in heritage Cantonese speakers in Toronto, Canada—word order in the resultative construction—in order to shed light on how the two forces of change operate in these situations.

Resultatives represent a “conflict site” (Poplack & Meechan, 1998) between Cantonese and the majority language of English: while English has adjectival resultatives (1) in which the result (R) adjective (bolded) appears after a theme object (O), Hong Kong Cantonese (2a) puts the result component (bolded) directly after the means (M) verb (underlined), before the theme object (and aspectual marking, A, if present). It also presents a possible case of overgeneralization: many disyllabic verbal constructions allow aspectual markers such as perfective *zo2* to



optionally intervene between the two syllables, as shown in (3), but this is generally resisted for resultatives in the homeland (Chan & Cheung, 2020).

Experiment. We designed an auditory yes/no acceptability judgement task to test Cantonese speakers' judgements of resultative constructions with different word orders. We test the acceptability of four word-order variants: the canonical M-R-A-O (2a), the English order M-A-O-R (2b), the morphological variant M-A-R-O (2c), and the ungrammatical baseline M-A-O-Adverbial-R (2d). We compare judgements from two groups: heritage speakers from Toronto (HER) and homeland speakers from Hong Kong (HOM). The experiment is run online on Gorilla ([www.gorilla.sc](http://www.gorilla.sc)). Sentences are presented auditorily, since (i) heritage speakers tend to have low literacy and (ii) judgements of written stimuli may be affected by the Mandarin-based literary register (cf. Sedarous & Namboodiripad, 2020). Stimuli were synthesized with the online voice generator by Narakeet ([www.narakeet.com](http://www.narakeet.com)).

Data collection is ongoing. The canonical M-R-A-O order (2a) should be grammatical in both heritage and homeland speakers, while the intervening adverbial order (2d) is expected to be ungrammatical in both groups. We expect two possible differences between HER and HOM. First, HER may transfer the English word order into their grammar, in which case the M-A-O-R order (2b) would be deemed more acceptable than the baseline condition (2d). Similar outcomes have been found in other contact situations between Chinese varieties and M-O-R languages (Kwok, 2010; Lee, 2016). Second, HER may overgeneralize *zo2* intervention (3) to resultatives and would thus show higher acceptability of the M-A-R-O order (2c) than the baseline. In both these cases, HOM should instead show low acceptability.

- Chan, Sheila S.L., & Lawrence Y.L. Cheung. 2020. Morpho-syntax of non-VO separable compound verbs in Cantonese. *Studies in Chinese Linguistics* 41(2): 185–206.
- Chan, Sheila S.L., Tommy T.M. Lee & Ka-Fai Yip. 2022. Discontinuous predicates as partial deletion in Cantonese. *University of Pennsylvania Working Papers in Linguistics* 28(1): 5.
- Kwok, Bit-Chee. 2010. Grammatical change in language contact: On the origin of the “verb + object + complement” construction in Nanning Yue. *Journal of Chinese Linguistics Monograph Series* 24: 201–216.
- Lee, Hui-chi. 2016. Hainan Min resultative structures. *Bulletin of Chinese Linguistics* 9: 67–82.
- Lohndal, Terje. 2021. Syntax of heritage languages. In Silvina Montrul & Maria Polinsky (eds.), *The Cambridge handbook of heritage languages and linguistics*, 644–667. Cambridge: Cambridge University Press.
- Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge: Cambridge University Press.
- Poplack, Shana, & Marjory Meechan. 1998. Introduction: How languages fit together in codemixing. *International Journal of Bilingualism* 2(2): 127–138.
- Sedarous, Yourdanis, & Savithry Namboodiripad. 2020. Using audio stimuli in acceptability judgment experiments. *Language and Linguistics Compass* 14(8): e12377.

---

### Place, ideology, and identity:

#### The construction of African American linguistic identity in Jackson, MS

Tom Lewis, Ane' Scott

Tougaloo College, Jackson, MS, USA

This paper presents a description of African American Language (AAL) in Jackson, MS, focusing on key factors shaping speakers' use of vowel productions associated with the African American Vowel Shift (AAVS) (Thomas, 2007; King, 2016; Jones, 2020) and on how interspeaker variation points to the ways in which speakers manipulate linguistic resources in the construction of ethnic identity. Recent research has indicated substantial regional and interspeaker variation in AAL (Kohn & Farrington, 2015; King, 2016, 2018; Kendall, 2018; Jones, 2020). Jones (2020) proposes an AAL dialect region extending along the Mississippi River towards the Great Lakes, delineated in part by participation in the AAVS. Our data indicates that while many Jackson residents participate in at least certain features of the AAVS, demonstrating tendencies towards reversing the nuclei of /e/ and /ɛ/ and towards the fronting and raising of /æ/, there is substantial and significant interspeaker variation. Inferential analysis identifies that several sociolinguistic variables significantly shape this variation, including gender, social network density, and participant generation. The overall picture that emerges is that of a community in flux, with younger generations being the most likely to adopt the vowel productions predicted by the AAVS. These patterns may indicate a contextually conditioned shift in progress towards AAVS shaped realizations among Jackson residents. However, this data also clearly points to the fluid and dynamic nature of linguistic variation, as individual speakers vary significantly in terms of production of the vowels in question, indicating that the linguistic construction of African American identity is a dynamic, individualized process.

The analysis is based on a series of sociolinguistic interviews (n=30) conducted with African Americans in Jackson. The interviews collected demographic information and included a reading passage (Jones, 2020), which

was transcribed and then analyzed in Praat. The DARLA interface was used for forced alignment and formant extraction. Vowel plots were created using NORM for the community and for individual speakers. Inferential statistics were modeled using R. This data is part of an ongoing project to examine AAL in Jackson, MS. This project contributes the exploration of variation in AAL in two ways: first, by describing the vowel systems of African American speakers in Jackson, MS, and second, by considering the social factors that shape interspeaker variation among African Americans in the city in relation to the use of linguistic resources in the construction of layered, multiplex African American identities. Viewed holistically, our work further advocates for continued attention to variation within African American Language and additional exploration of the ways individual speakers construct an African American identity using linguistic resources.

- Benor, S. B. (2010). Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics*, 14(2), 159-183.
- Durian, D., Dodsworth, R., & Schumacher, J. (2010). Convergence in blue-collar Columbus, Ohio African American and White vowel systems? *Publication of the American Dialect Society*, 94(1), 161-190.
- Farrington, C. (2019). *Language variation and the Great Migration: Regionality and African American Language*. (Doctoral Dissertation). University of Oregon.
- Farrington, C., King, S., & Kohn, M. (2021). Sources of variation in the speech of African Americans. *WIREs Cognitive Science*, 12.
- Fasold, R. W. (1972). *Tense Marking in Black English: A Linguistic and Social Analysis*. Arlington, VA: Center for Applied Linguistics.
- Forest, J. & Dodsworth, R. (2016). Towards a sociologically grounded view of occupation in sociolinguistics. *University of Pennsylvania Working Papers in Linguistics* 22(2), 31-40.
- Fridland, V. (2003a). Network strength and the realization of the southern vowel shift among African Americans in Memphis, Tennessee. *American Speech* 78(1), 3-30.
- Fridland, V. (2003b). "Tie, tied, and tight": The expansion of /ai/ monophthongization in African-American and European-American speech in Memphis, Tennessee. *Journal of Sociolinguistics*, 7(3), 279-298.
- Green, L. J. (2002). *African American English: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Jones, T. (2020). *Variation in African American English: The Great Migration and Regional Differentiation*. (Doctoral Dissertation). The University of Pennsylvania.
- Kendall, T. & Wolfram, W. (2009). Local and external language standards in African American English. *Journal of English Linguistics* 37, 5-30.
- King, S. (2016). On negotiating regional and racial identities: Vocalic variation among African Americans in Bakersfield, California. *University of Pennsylvania Working Papers in Linguistics*, 22(2).
- King, S. (2018). *Exploring social and linguistic diversity across African Americans from Rochester*. (Doctoral Dissertation). Stanford University.
- King, S. (2021). From African American Vernacular English to African American Language: Rethinking the study of race and language in African American's speech. *Annual Review of Linguistics*, 6(1), 285-300.
- Kohn, M. E. (2014). "The way I communicate changes but the way I speak don't": A longitudinal perspective on adolescent language variation and change. Durham, NC: Duke University Press.
- Kohn, M. E. & Farrington, C. (2013). A tale of two cities: Community density and African American English vowels. *University of Pennsylvania Working Papers in Linguistics*, 19(2).
- Labov, W. (1972). *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
- Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English*. Berlin & New York: Mouton de Gruyter.
- Labov, W., Cohen, P., Robins, C., Lewis, J. (1968). *A Study of the Non-Standard English of Puerto Rican and Negro Speakers in New York City*. Washington, D.C.: United States Office of Education Final Report, Research Project 3288.
- Labov, W., Rosenfelder, I., & Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, 89, 30-65.
- Legum S. E., Pfaff, C., Tinny, G., Nicholas, M. (1971). *The Speech of Young Black Children in Los Angeles*. Los Angeles: Southwest Regional Laboratory for Research and Development, Technical Report 33, 01.
- Mizumoto, A. (2022). Calculating the relative importance of multiple regression predictor variables using dominance analysis and random forests. <https://doi.org/10.31219/osf.io/w8nb3>.
- Podesva, R. J., D'Onofrio, A., Van Hofwegen, J., & Kim, S. K. (2015). Country ideology and the California vowel shift. *Language Variation and Change*, 27, 157-186.
- Rickford, J. R. & Price, M. (2013). Girlz II women: Age-grading, language change, and stylistic variation. *Journal of Sociolinguistics*, 17(2), 143-179.
- Shands, H. A. (1893). *Some Peculiarities of Mississippi Speech*. (Doctoral Dissertation). University of Mississippi.
- Tagliamonte, S. & Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, 24(2), 135-178.
- Thomas, E. R. (2007). Phonological and phonetic characteristics of African American Vernacular English. *Language and Linguistics Compass*, 1(5), 450-475.
- Thomas, E. R. (2011). *Sociophonetics: An Introduction*. London: Palgrave MacMillan.
- Weldon, T. L. (2021). *Middle Class African American English*. Cambridge: Cambridge University Press.

- Wilkerson, R. (2015). African American English in the Mississippi Delta: A case study of copula absence and r-lessness in the speech of African-American women in Coahoma County. In S. Lanehart (Ed.), *The Handbook of African American Language*, 338-352. New York: Oxford University Press, 201-218.
- Wolfram, W. (1969). A sociolinguistic description of Detroit Negro Speech. Washington, D.C.: Center for Applied Linguistics.
- Wolfram, W. (2007). Sociolinguistic folklore in the study of African American English. *Language and Linguistic Compass*, 1(4), 292-313.
- Wolfram, W. (2015). The sociolinguistic construction of African American Language. In S. Lanehart (Ed.), *The Handbook of African American Language*, 338-352. New York: Oxford University Press, 338-352.
- Wolfram, W. & Kohn, M. (2015). Regionality in the development of African American English. In S. Lanehart (Ed.), *The Handbook of African American Language*, 338-352. New York: Oxford University Press, 140-159.

## The effects of stylistic covariation in phonological variant choices

Aini Li

University of Pennsylvania, Philadelphia, USA

Language variation is stylistically conditioned, often with standard variants being favored in formal speech styles and nonstandard features being favored in casual ones [1, 2]. However, it remains unclear what kind of knowledge individuals may have about the stylistic conditioning on variation, and whether they utilize this knowledge when they process variation in real time [3, 4]. Using the sociolinguistic variable (ING), this study probes the tacit knowledge listeners may have about the stylistic conditioning on variation through their behaviors of uncertainty resolution in speech perception.

A forced-choice variant identification experiment was conducted. Listeners (N=40) were asked to choose out of two written word options the one they thought they heard (e.g., *sleepin'* vs. *sleeping*). 24 critical and 72 filler sentences were constructed and produced by a phonetically-trained female American English speaker from Ohio under both casual (i.e., conversational) and formal (i.e., public speech) styles in a professional sound-attenuated booth. To implement greater experimental control, the speaker targeted features like vowel reduction and nasal flapping in casual guises and t-release and unreduced vowels in formal ones. Moreover, sentences were intentionally created to vary the number (ranges from 2 to 6) of discrete stylistic cues to formality (e.g., t/d-deletion, glottalization) before and after ING-suffixed target words. The two guises of each sentence were similar in duration. For critical items, the ING portion in all targets (e.g., *carvING*) was zeroed out and replaced with white noise. White noise bursts in both guises were of the same length (i.e., the average of the original ING duration of the two guises of that item) and amplitude, to prevent listeners from relying on variant relative durational cues in their variant identification. To not draw attention to the critical ING bursts, white noise was added to the sentences throughout, and additional bursts were added in noncritical spots. For filler items, similar manipulation was performed except that no ING-suffixed words were included and white noise bursts were randomly interspersed throughout the recordings. Critical sentences were normed for their perceived formality.

The experiment employed a within-subjects design where each listener heard both styles, but only one guise for each sentence. Mixed-effects regression showed listeners were significantly less likely to perceive the uncertain white noise as -ing when the utterance was produced casually ( $\beta = -0.63$ ,  $p = .02$ ) (Figure 1). Interestingly, the number of discrete stylistic cues after the ING-suffixed target turned out to be statistically significant ( $\beta = -0.31$ ,  $p = .04$ ), suggesting that -ing response rates decreased significantly when there existed more covarying stylistic cues to informality before variant identification decisions. The results show that listeners make use of co-varying stylistic cues to infer their variant choices under uncertainty during speech perception. It is possible that the preceding stylistic cues did not directly influence listeners' inferences. Rather, these stylistic cues following critical targets impacted listeners' memory of what they heard, further influencing their variant choices. Taken together, individuals are indeed sensitive to stylistic conditioning on variation and can utilize this knowledge in real-time instances of language use.

- [1]. Vaughn, C., & Kendall, T. (2018). Listener sensitivity to probabilistic conditioning of sociolinguistic variables: The case of (ING). *Journal of Memory and Language*, 103, 58-73.
- [2]. Campbell-Kibler, K. (2010). The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change*, 22(3), 423-441.
- [3]. Tamminga, M., MacKenzie, L., & Embick, D. (2016). The dynamics of variation in individuals. *Linguistic Variation*, 16(2), 300-336.
- [4]. Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change*, 21(1), 135-156.

---

# The Taishanese Belted L: Acoustic Features, Lateralization from Cantonese /s/, and Heritage vs. Native Speakers

Vicki Li

CUNY - Graduate Center, New York, USA

Taishanese is a Siyi dialect of Cantonese that existed as the lingua franca of Chinese immigrants into the U.S. before being replaced by Cantonese (Lee 2017; Leung 2012). Though there are similarities, some phonological and phonetic features of Taishanese do not exist in Cantonese nor Mandarin, resulting in a lack of complete mutual intelligibility (Yue-Hashimoto 1971; Cheng 1973). For example, the voiceless lateral fricative [ɬ] or the “belted l” is such a feature, previously studied through its role in internet discourse about Taishanese language attitudes and maintenance. This sound also contributes to perceptions of the variant as “sounding harsh” (Leung 2010, 41). Considering its prominence in indexing Taishanese identity and the lack of research done on the sound in general, study of the belted l serves as an untapped source of information.

Currently, the belted l exists as a counterpart to the Cantonese /s/, which is theorized as part of a larger push chain phenomenon of fricatives changing in places of articulation: merging and fronting of the retroflex ([ʂ]) and palatals ([j], [ç]) to a post alveolar [ʃ], then to an alveolar ([s]), which was lateralized to a ([ɬ]) (Cheng 1973). This last lateralization is rare, especially as it has only occurred in Taishanese and not Cantonese. Also, the change is incomplete as variation between Taishanese and Cantonese on this phenomenon still exists. Partially due to Taishanese-Cantonese bilingualism, the lateralized fricative is a variant of the Cantonese alveolar fricative. Rather than through places of articulation as past studies have done, I investigate this variation and the nature of the change through an acoustic phonetic lens.

This study addresses the following regarding the Taishanese belted l:

1. What is the Taishanese belted l? (Acoustic Phonetic documentation) 1. Is there a possibility of an acoustic cline from the Cantonese /s/ when compared to Taishanese heritage and native [ɬ]?
  2. Is there a difference (acoustics-wise) between native and heritage speakers?
  3. Is there a difference in usage of the belted l between native and heritage speakers? 1. Does age influence usage and acoustics of the belted l?
2. How do results speak to past sociolinguistic studies that report either assimilation to Cantonese or resistance to said assimilation in younger groups?

Following a demographic interview (birth year, place of birth/childhood, languages spoken & self-reported fluency from 1-5), 15 Taishanese speakers (1 elderly, 5 middle-aged, 9 young/ 6 native, 9 heritage) were given 40 flashcards with a photo, the word in English, and in simplified Chinese. Each word was used in the following sentence in Taishanese (This is the \_\_\_\_ card.) All IPA readings, tones for words, and word bank taken from Taishanese and Cantonese dictionaries. The spectral moments (center of gravity, standard deviation, skewness, kurtosis) were recorded through a Praat script (DiCanio 2013). Comparison of this data is currently underway. However, analysis of speakers’ productions will contribute to a deeper understanding of phonological change with evidence from an under-examined language while also providing a sociophonetic perspective on these changes.

Cheng, Teresa M. 1973. THE PHONOLOGY OF TAISHAN. *Journal of Chinese Linguistics* 1, no. 2, [Chinese University Press, Project on Linguistic Analysis]: pp. 256–322, <http://www.jstor.org/stable/23749797>.

DiCanio, Christian. 2013. Spectral moments of fricative spectra script in Praat. Haskins Laboratories & SUNY Buffalo. Version 4.0 revised in 2021.

Leung, Genevieve. Y. 2010. Learning about Your Culture and Language Online: Sifting Language Ideologies of Hoisan-wa on the Internet (“I especially liked how you used the thl-sound.”) *Working Papers in Educational Linguistics* 25: no. 1. Retrieved from <https://repository.upenn.edu/wpel/vol25/iss1/3>.

\_\_\_\_\_. 2012. Contemporary Hoisan-wa Language Maintenance in Northern California: Evidence From Fourteen Frog Story Narratives. *International Multilingual Research Journal* 6, no. 2: 104–120. <https://doi.org/10.1080/19313152.2011.651427>

Yue-Hashimoto, Anne. 1971. *A Guide to the Tai-shan Dialect*. Chinese Linguistics Project, Princeton Univ., NJ.

---

## French clitics: A paradox from variation

Yiming Liang, Caterina Donati, Heather Burnett

Laboratoire de linguistique formelle (Université Paris Cité/CNRS), Paris, France

**1. Introduction.** In the research on the syntax of French clitics in Spoken French, there have been many arguments that subject clitics are agreement markers (Auger 1995, Culbertson 2010, Authors 2023) and, when there is a lexical doubled DP, it is a Topic (Authors 2023). In this paper, we raise a challenge for this conclusion involving the possible intervention of object clitics in between the subject clitic and the inflected verb, as in Marie<sub>i</sub> elle<sub>i</sub> le mange ‘Mary<sub>i</sub> she<sub>i</sub> eats it’. Given the lexical integrity hypothesis (DiSciullo and Williams 1987) (or more syntactic

re-formulations where words are phrases (Embick and Noyer 2007)), this co-occurrence can mean only one of two things: 1) either object clitics are also agreement markers, and their compatibility is thus not an issue; or 2) object clitics are *not* agreement markers, but a syntactically merged head (Nevins 2011), and a derivational paradox arises, with a morphological affix merged to a syntactically complex head. Based on new evidence from a quantitative study of sociolinguistic variation, we argue that object clitics are very different from subject clitics, and do not behave like agreement markers, leaving us with a puzzle.

**2. Subject doubling in the MPF.** Subject doubling, where a nominal subject and a coreferential subject clitic co-occur (*Marie<sub>i</sub> elle<sub>i</sub> mange* ‘Mary<sub>i</sub> she<sub>i</sub> eats’) is a common feature of spoken French. We tracked the contours of this phenomenon in the *Multicultural Parisian French* corpus (Gadet & Guerin 2016). We extracted all occurrences of preverbal nominal subjects (e.g., *mon pe`re* ‘my father’, *Marie*, etc.) and preverbal quantified subjects (*certain* ‘certain people’, *personne* ‘nobody’) (N=3620, strong pronouns excluded) from the corpus, and coded the following factors: social factors (age, gender, education, profession), the polarity of the clause (i.e., affirmative, negative with ‘ne’, negative without ‘ne’), verb frequency, DP subject head frequency, the distance (in words) between subject and verb, subject type and clause type. Speaker and verb lemma were considered as random intercepts. Mixed effects logistic regression analysis of our data reveals significant effects that provide support for a) clitics being agreement markers, and b) the lexical DP being dislocated, as in (2).

(2) Our analysis: [<sub>TOP</sub> Marie [<sub>TP</sub> *pro* [<sub>T</sub> elle-mange ]]

**Supporting (a):** we find a very high rate of subject doubling (77% for preverbal and postverbal DP subjects) and an extremely low rate of *ne* with doubled subjects (0%). Furthermore, within the 23% of occurrences without subject clitics, most of them co-occur with ‘ne’ or are about education and religion, suggesting they belong to the formal Standard French register. We also find that verb frequency is positively correlated with subject doubling ( $p < 0.01$ ), which, we argue, provides indirect evidence in favour of (a), since high frequency is well known to preserve complex morphosyntactic forms (Bybee, 2003). **Supporting (b):** we find that doubling is governed by subject type in terms of topicality hierarchy: definite DPs (rate of doubling: 82%) > indefinite DPs (39%) > universal QPs (7%). Crucially, we also find that root clauses have the highest rate of subject doubling; whereas subordinate clauses disfavor it. Among them, relative clauses disfavor doubling most, while other subordinate clauses are in-between (root (79%) > other subordinates (62%) > relatives (31%); differences between two adjacent categories being significant  $p < 0.001$  for both subject type and clause type). This asymmetry is expected if the doubled DP is some kind of topic, since it is well known that there are topics (eg. Cinque 1977’s “hanging topic”) that are only available in root clauses (cf. Albrecht et al. 2012, a.o.).

**3. Object doubling in the MPF.** What about object doubling, where a nominal object co-occurs with a coreferential object clitic (e.g., *Le film<sub>i</sub>, je l<sub>i</sub>’ai vu* ‘The film I saw it.’ or *Je l<sub>i</sub>’ai vu, le film<sub>i</sub>* ‘I saw it the film’)? If object clitics are agreement markers, we would expect the same kind of facts supporting (a). If doubled NPs are topics, we would expect similar patterns supporting (b). In order to test these predictions, we extracted from the MPF corpus all occurrences of nominal direct objects and coded whether the object is doubled by a direct object clitic “le”, “la”, “les” (N=11 273). Object doubling cases involve both preverbal and postverbal direct DP objects. Strong pronouns were excluded. We annotated the following factors: social factors (age, sex, education, profession), the polarity of the clause, verb frequency, DP object head frequency, object type and clause type, and added a new factor, DP object animacy. Two random intercepts (speaker and verb lemma) were included in the mixed effects logistic regression model. Given the limited number of doubling cases (N=254), we had to analyze social factors and linguistic factors in separate mixed models.

Statistical analyses revealed linguistic effects supporting (b) but against (a). Against (a): the rate of object doubling is extremely low (2.3%, preverbal and postverbal DP object included). We also observe a negative correlation between verb frequency and object doubling ( $p < 0.05$ ), which contrasts with the directionality observed in subject doubling. This fact suggests that object clitic is **not** an agreement marker, otherwise highly frequent verb would have favored the more complex morphosyntactic form as object doubling. Supporting (b): we find that object doubling is also favoured with definite DPs (rate of doubling: 4.4%) compared with indefinite DPs (0.2%) and universal QPs (0.4%) (definite vs. indefinite DPs:  $p < 0.001$ , no difference between indefinite and universal DPs:  $p = 0.57$ ). Furthermore, root clauses (2.5%) favour object doubling compared with subordinate clauses (1.5%,  $p < 0.01$ ), and no difference in object doubling was revealed between relative clauses (1.5%) and other subordinate clauses (1.5%). DP object animacy is also an important predictor of doubling, with animate objects associated with a higher rate of doubling compared with inanimate objects ( $p < 0.001$ ). These facts strongly suggest that object doubling is very different from subject doubling, and that object clitics are not agreement markers, but rather (resumptive) pronouns associated with topicalization.

**4. The paradox.** Our two corpus studies in the MPF reveal a remarkable difference in frequency and distribution between subject and object doubling, and suggest that subject clitics are agreement markers, while direct object clitics are resumptive pronouns. If this is the case, the object clitic should block merge of the subject affix to the

verb head under the Lexical Integrity Hypothesis. We tested this prediction and found the presence of object clitics indeed disfavors subject doubling ( $p < 0.001$ ). This being said, when an object clitic intervenes, subject doubling is still used at a rather high frequency (55%). By contrast, we saw that the negative “ne” clitic rarely co-occurs with subject doubling in the MPF corpus. The following puzzle thus arises: if subject clitics are affixes, why can object clitics intervene between the affix and the hosted verb, while “ne” cannot? One possibility is that Lexical Integrity is not an absolute condition, as discussed in Bosque (2019), but rather a reflex of locality conditions (phase domains) on lexicosyntactic derivations. In a derivational perspective, object clitics, which are merged to the V lower than T, can be included in a complex derived form with subject markers within the same phase, while negation, which is merged higher, cannot. A radical alternative would be to argue that the incompatibility of ‘ne’ with subject clitics is not syntactic after all, but rather due to simple socio-stylistic reasons: since ‘ne’ is a marker of formal register, it is stylistically incompatible with subject doubling. This conclusion appears to be reinforced by the existence of Northern Italian dialects where weak negative elements can regularly occur between the subject clitic and the verb (Manzini and Savoia 2005: 2.5). But if weak elements such as negation and object clitics can both intervene between an inflectional prefix and its host, the divide between syntax and morphology underlying the distinction between the notions of pronoun and inflectional marker gets even further blurred, and we are left with no real explanation for the differences we observed between subject and object clitics.

Aelbrecht, L., Haegeman, L., & Nye, R. (2012). *Main Clause Phenomena: New Horizons*. John Benjamins.

Alexiadou, A., & Anagnostopoulou, E. (1998). Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language & Linguistic Theory*, 16(3), 491–539.

Auger, J. (1995). Les clitiques pronominaux en français parlé informel: Une approche morphologique. *Revue québécoise de linguistique*, 24(1), 21–60.

Culbertson, J. (2010). Convergent evidence for categorial change in French: From subject clitic to agreement marker. *Language*, 86(1), 85–132.

---

## You are not a native speaker, but I want to sound like you: Unveiling alternative VOT convergence patterns in L2 English speech

Ke Lin

Georgetown University, Washington, DC, USA

A robust body of research has examined convergence at various phonetic levels in both L1 and L2 contexts; however, much less work in the field has investigated phonetic convergence by L2 speakers with attention to consonant features. To address this gap, the present study uses Voice Onset Time (VOT), a feature of stop consonants, to investigate phonetic convergence in English spoken by native speakers of Mandarin. In particular, we explore the effect of participants’ beliefs about the L1 status of an interlocutor on the degree of VOT convergence for voiceless aspirated stops /p, t, k/. To this end, we designed a shadow task contextualized as an escape room activity, in which we constructed impressions of a model talker’s L1 status via a customized online avatar and its given name. We randomly assigned participants to L1 or L2 interlocutor condition, and presented them with the same audio stimuli produced by a male native speaker of American English. In the L1 interlocutor condition, they shadowed an avatar of a white student with an Anglo-Saxon name. In the L2 interlocutor condition, participants repeated words after an avatar of an Asian student with a Chinese name. Our results showed that all participants shortened their VOT from baseline production to shadow task. This effect was strongest for /p/ and weakest for /t/. Notably, the convergence magnitude was on average larger for participants exposed to the L2 interlocutor condition than those exposed to the L1 condition. This outcome lends support to the Interactive Alignment Model (Pickering & Garrod, 2004), which posits that speech alignment primarily serves the purpose of enhancing mutual comprehension. We suggest that our participants, when communicating in their L2 with a presumed Mandarin-speaking model talker, adapted their speech to improve communication by mimicking the model talker’s speech patterns.

1. Chao, K. Y., & Chen, L. M. (2008, June). A cross-linguistic study of voice onset time in stop consonant productions. In *International Journal of Computational Linguistics & Chinese Language Processing*, Volume 13, Number 2, June 2008 (pp. 215-232)
2. Jiang, F., & Kennison, S. (2022). The Impact of L2 English Learners’ Belief about an Interlocutor’s English Proficiency on L2 Phonetic Accommodation. *Journal of Psycholinguistic Research*, 1-18.
3. Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39 (2), 132-142.
4. Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of experimental psychology: General*, 115(1), 39.
5. Pickering, M., & Garrod, S. (2004). The interactive-alignment model: Developments and refinements. *Behavioral and Brain Sciences*, 27(2), 212-225.

---

## Variable adaptation of English /ə/ in Quebec French

Massimo Lipari

McGill University, Montreal, Canada

Introduction: Languages use a variety of strategies to parse new loanwords containing a sound absent from the native phonology (Kang 2011). One option involves relying on the orthography, substituting the expected sound given the usual grapheme-to-phoneme mapping in the recipient language. Alternatively, the offending phoneme may be approximated with a phonetically or featurally similar one from the existing segment inventory. A third possibility is to import the missing sound wholesale, creating new (marginal) phonemic contrasts. Speakers of Quebec French have been claimed to employ all three of these strategies to integrate /ə/ from English borrowings, like in the word soccer (/səkə/), producing the alternation [səkɛɾ]~[səkœɾ]~[səkə] (Côté 2021). No study, however, has examined the acoustics of this variable loanword vowel (henceforth labelled ‘X’). As a result, its relation to the two native phonemes perceptually closest to the donor language sound, /œ/ and /ø/, is largely a mystery. Moreover, it is not known how the evolving social dynamics between French and English in Quebec have influenced its realization. Indeed, the increase in individual bilingualism from the late 20th century onwards may be expected to have favoured the maximally faithful realization (i.e., [ə]), leading to sound change.

Research questions: How does the acoustic realization of English loanword /ə/ in Quebec French compare to the similar vowels /œ, ø/? How has this realization evolved over time?

Data: Data are from an expanded version of Milne’s (2014) AssNat corpus. The dataset for this pilot includes 317 tokens of X from the loanword Orford (a placename) across 26 speakers—9F, 17M, born 1941 – 1991—along with 797 tokens of (final syllable, pre-rhotic) /œ/ and 1157 tokens of (final, open syllable) /ø/. Formant measures were normalized using Nearey2 (Nearey 1978).

Methods: The overlap between empirical by-speaker F1 × F2 × F3 distributions for pairs X – /œ/ and X – /ø/ is measured using Pillai scores (cf. Hay et al. 2006, Nycz & Hall-Lew 2014). In the future, empirical distributions will be replaced with the posteriors of a Bayesian regression model.

Preliminary results: F2 of vowel X shows the clearest evidence of change, especially amongst women (Figure 1): this fronting is evidence of the regression of the orthography-based variant [ə]. Almost across the board, F3 of X is lower than that of /ø, œ/, suggesting some degree of successful imitation and importation of rhoticity from English /ə/. F3 of X begins at a similar level for men and women, but declines over time for the latter group only (consistent with change towards increased use of [ə]-like variants). The evidence from Pillai scores (Figure 2) is less clear, but does suggest that X is more similar to /ø/ than to /œ/, contrary to the impressionistic descriptions. While X is fairly distinct from both /ø/ and /œ/ at the outset, its distribution increasingly approximates that of /ø/ (and, for men, also that of /œ/, but to a lesser extent). Confidence intervals here are wide, however: more data is necessary to ascertain the nature of these trends.

Côté, M.-H. (2021). L’adaptation des emprunts de l’anglais en français: Variation dialectale, phonologie, lexicographie. In *L’adaptació de manlleus en català i en altres llengües romàniques* (pp. 43–47). Publicacions i Edicions de la Universitat de Barcelona.

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), 458–484.

Kang, Y. (2011). *Loanword Phonology*. In *Blackwell Companion to Phonology* (sec. 95). Wiley.

Milne, P. (2014). The variable pronunciations of word-final consonant clusters in a force aligned corpus of spoken French.

Nearey, T. M. (1978). *Phonetic feature systems for vowels*. Indiana University Linguistics Club.

Nycz, J., & Hall-Lew, L. (2014). Best practices in measuring vowel merger. *Proceedings of Meetings on Acoustics*, 20(1), 1–19.

---

## “Read a news report and make a blessing”: accessing the social meaning of the pharyngeals variants in Hebrew

Marcelo Melo, Gisele Costa

Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

This study presents an analysis of the behaviour of native Hebrew speakers based on two sociolinguistic variables: (a) *het*, produced as a voiceless pharyngeal fricative [ħ] or a voiceless velar fricative [x]; (b) *ayn*, produced as a voiceless pharyngeal fricative [ʕ] and a glottal plosive voiceless [ʔ]. According to Gafter (2016), the realization of these two variables is conditioned by the ethnic group to which speakers belong, as well as by the context of interaction in which speakers are. The pharyngeal variants of these two variables, according to Gafter, are stigmatized by Hebrew speakers, especially by speakers who belong to socially prestigious ethnic groups. It happens because different usages and behaviours of Hebrew speakers are related to different Jewish origins: Jews from Europe (Ashkenazi) constitute the

prestigious social group, which means that linguistic variants associated with this group are the prestigious variants as well; on the other hand, Jews from the Middle East (Mizrahi) constitute the less privileged group of Israeli society, and their speech is therefore less prestigious. If pharyngeal variants of *het* and *ayn* are associated with Mizrahis speaking, these variants tend to be stigmatized and consequently avoided in broader social contexts that are not related to religion. In religious contexts, despite being stigmatized, Mizrahi way of speaking is associated with a purer and more conservative way of speaking once it is associated with the historical origins of Hebrew. In this sense, pharyngeal variants sound to be more appropriate for religious contexts. In order to observe the behaviour of different groups of Hebrew speakers, participants from three ethnic groups (12 Ashkenazi, 10 non-Yemeni Mizrahi, and 12 Yemeni Mizrahi) read two texts: one non-religious (news report) and one religious (blessing). The aim was to verify (1) if there would be different behaviours of the speakers from the different ethnic groups concerning the two variables and (2) if there would be an influence of the context (religious and non-religious) for the realization of the variables. The results point to a difference in the behaviour of the groups based on ethnic origin: Ashkenazi and non-Yemeni Mizrahis speakers did not produce any pharyngeal variants in either of the two reading contexts. Regarding the Yemeni Mizrahi speakers, pharyngeal variants of *het* and *ayn* were produced in non-religious context (13.23% and 9.38%, respectively) and even more produced in the reading of the religious text (61.11% and 83.85%, respectively). Differences in percentages of the realization of the pharyngeals for the two variables may indicate that both variables are socially stratified, but *ayn* seems to be more sensitive to stylistic variation, since there is a more expressive difference in the realization of *ayn* variants depending on the context (religious or non-religious). Thus, the results suggest that different structural and social patterns of variation can be developed from values and social meanings built around the linguistic forms that are adopted by speakers of different social groups and in different contexts of interaction.

GAFTER, R. J.. What's a stigmatized variant doing in the word list? Authenticity in reading styles and the Hebrew pharyngeals. *Journal of Sociolinguistics* 20, 2016: 31-58.

---

Language-specific prosody, convergence, and hyperarticulation in statements of  
 Palenquero/Spanish bilinguals  
 Wilmar Lopez-Barrios  
 University of Massachusetts Amherst, Amherst, USA

This study explores the extent to which Palenquero/Spanish bilinguals, a community that appears to have a residual high tone of African origin[5], keep their two languages prosodically distinct across statements. It has been shown that bilinguals' languages may be coactivated[4], showing cross-linguistic interactions such as transfer or interference. If such interactions are maintained over time, language-specific prosody could eventually converge. This phonological process is undoubtedly more accelerated when bilingual's non-dominant language corresponds to the minority language in the community. Prosody is more susceptible to the adverse effects of majority languages than segmental information[1]. Nevertheless, some creole languages that developed from the contact of African and European languages may exhibit hybrid prosodic systems with tones from substrate languages, and stress from the majority language. As prosodic convergence seems to be inescapable under these circumstances, which factors are supporting language-specific prosody in these populations?

Palenquero is a creole language, and its most salient prosody was described as a type of phrase-final "cadence"[3], correlated with a falling pitch and a lengthening effect, stretched from the penultimate syllable until the end of the statement[5]. Correa[2] suggested that although Palenquero/Spanish bilinguals produce more penultimate lengthening in Palenquero than Spanish, the prosody of the two languages did not hide great differences. The speakers use the same inventory of pitch accents and nuclear configurations in the two languages[2, p.264]. However, Lipski[6] noticed that, when these bilinguals listened to Palenquero intonation in Spanish utterances, they biased language identification in the direction of Palenquero (p.53). To better understand language-specific prosody and convergence in Palenquero/Spanish bilinguals, this study aims to answer whether the speakers use phrase-final lengthening language-specifically, and if they maintain the two languages intonationally distinct across statements.

234 five-syllable statements were elicited through a discourse completion task[9], with the participation of ten Palenquero/Spanish bilinguals, in two unilingual sessions. Phrase-final lengthening was assessed with linear mixed-effects models predicting vowel length in final trochees and iambs, from both adult and elderly bilinguals. Language-specific intonation was explored through F0 contours which were dimensionally reduced using Functional



Principal Component Analysis[7]. Final stress, language, and generational effects were tested on each F0 dimension, using linear mixed-effects models. Results indicate that penultimate lengthening is more pronounced in Palenquero than Spanish across elderly bilinguals, while adults do not show this language-specific distinction, having a shorter lengthening. Both generations kept their two languages intonationally distinct using plateau-shaped contours in Palenquero, and initial rises followed by steeper declinations in Spanish (see Fig.1). However, the elderly show a wider distance between the two languages. While adults exhibit more convergence, there does seem to be an underlying process (perhaps a substrate effect) driving plateau-shaped intonation in Palenquero, regardless of generation. Nevertheless, Spanish intonation of elderly speakers was significantly different from adults, suggesting that the former not only hyperarticulate Spanish prosody—likely as a result of imperfect group learning[8]—but also, and as a consequence, enhance language-specific prosody incidentally, with no apparent cross-linguistic interaction.

- [1] Bullock, B. E. (2009). Prosody in contact in French: A case study from a heritage variety in the USA. *International Journal of Bilingualism*, 13(2), 165–194.
- [2] Correa, J. A. (2017). Intonation in Palenquero creole and Palenquero Spanish, Colombia. *Orality, Identity, and Resistance in Palenque, Colombia: An interdisciplinary approach*, 54, 245.
- [3] Friedemann, N., & Patiño-Roselli, C. (1983). *Lengua y sociedad en el Palenque de San Basilio*. Instituto Caro y Cuervo.
- [4] Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.
- [5] Hualde, J. I., & Schwegler, A. (2008). Intonation in palenquero. *Journal of Pidgin and Creole Languages*, 23(1), 1–31.
- [6] Lipski, J. M. (2016). Palenquero and Spanish. *Journal of Pidgin & Creole Languages*, 31(1), 42–81.  
<http://silk.library.umass.edu/login?url=http://search.ebscohost.com/login.aspxdirect=true&db=ufh&AN=114849234&site=ehostlive&scope=site>
- [7] Ramsay, J., Hooker, G., & Graves, S. (2009). *Functional data analysis with R and Matlab* (1st ed.). Springer.  
<https://doi.org/https://doi.org/10.1007/978-0-387-98185-7>
- [8] Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
- [9] Vanrell, M. d. M., Feldhausen, I., & Astruc, L. (2018). The discourse completion task in romance prosody research: Status quo and outlook. In I. Feldhausen, J. Fliessbach, & M. d. M. Vanrell (Eds.), *Methods in prosody: A romance language perspective* (pp. 191–227). Language Science Press.

---

## Age effects and language differentiation for yes/no question intonation in two groups of Palenquero/Spanish bilinguals

Wilmar Lopez-Barrios, Meghan Armstrong  
 University of Massachusetts Amherst, Amherst, USA

Palenquero is an Afro-Hispanic creole language which is being revitalized since the early 1990s. Fluent Palenquero speakers avoided its use, and did not transmit it to their children, due to the stigma associated with the creole[6]. Those former children are now adults, and are also the architects of an extraordinary revival program for the creole. They are heritage speakers of Palenquero, and have additionally learned the creole in school[5]. Thus, Palenquero/Spanish bilinguals speak both a European majority language, and an Afro-Hispanic creole which still preserves some of the phonotactic restrictions of their African ancestors (e.g., Kikongo speakers[4]). With respect to prosody, previous studies have suggested that these bilinguals use penultimate lengthening language-specifically, as it seems to be less frequently used in Spanish statements[3]. The present study aims to determine to what extent two groups of bilinguals, adult and elderly speakers, keep their Palenquero and Spanish prosodically distinct for a previously unstudied sentence type: Yes-No questions.

Ten Palenquero/Spanish bilinguals participated in two unilingual discourse completion tasks[10], eliciting 230 yes/no questions containing segmental strings of five syllables. Penultimate lengthening was tested with linear mixed-effects models predicting vowel length in final trochees and iambs, from both adult ( $\mu_{age}=39.5; SD=1.97$ ) and elderly ( $\mu_{age}=64.3; SD=5.44$ ) bilinguals. Language-specific intonation was explored on dimensionally-reduced F0 contours, using Functional Principal Component Analysis[2,8], while final stress, language, and generational effects were assessed on each F0 dimension, using linear mixed-effects models. Results demonstrate that adult bilinguals adapted boundary tones to the stress pattern of the final word. They produced L% for trochees, but tended to truncate this boundary tone for questions ending in iambs (see Fig.1). On the other hand, the elderly tended to produce boundary tones for yes-no questions in a language specific way: L% in Palenquero and H% in Spanish, regardless of the stress pattern of the final word.

It has been documented in the past that varieties of Spanish such as Canary Islands Spanish, Puerto Rican Spanish and Buenos Aires Spanish in Argentina exhibit L% for yes/no questions, while truncation of the L% can result when utterances end in iambs[7]. Adults appear to behave in a way similar to other Spanish varieties in the Caribbean. The elderly, by contrast, have consistently used L% in Palenquero, and this does not seem to be sensitive to the stress pattern of the final word in the yes-no question. Conversely, Spanish yes/no questions tend to end in a high, plateau-like tone. One explanation for this is that the elderly speakers have overgeneralized the truncation rule followed by the adults (and shown in other Spanish varieties) so that all yes-no questions end in a “truncated” way. This phenomenon has been shown for toddlers acquiring Puerto Rican Spanish, which has variable truncation for their  $\uparrow\text{H}^*\text{L}\%$  nuclear configuration, and the L% tone is often (but not always) truncated.[1]. In language contact, overgeneralization might be connected to the imperfect group learning that non-native speakers may reveal during the process of language shift[9].

- [1] Armstrong, M. E. (2012). The development of yes-no question intonation in Puerto Rican Spanish (Doctoral dissertation). The Ohio State University.
- [2] Asano, Y., Gubian, M., & Sacha, D. (2016). Cutting down on manual pitch contour annotation using data modeling. *Speech Prosody 2016*, 282–286.
- [3] Correa, J. A. (2017). Intonation in Palenquero creole and Palenquero Spanish, Colombia. *Orality, Identity, and Resistance in Palenque, Colombia: An interdisciplinary approach*, 54, 245.
- [4] Moñino, Y. (2017). Orality, identity, and resistance in Palenque, Colombia. In A. Schwegler, B. Kirschen, & G. Maglia (Eds.). John Benjamins Publishing Company.
- [5] Morton, T. (2005). Sociolinguistic variation and language change in el Palenque de San Basilio, Colombia (Doctoral dissertation). University of Pennsylvania. Pennsylvania, USA.
- [6] Patiño-Roselli, C. (1983). *Lengua y sociedad en el Palenque de San Basilio*. In N. S. de Friedemann & C. Patiño-Roselli (Eds.). Instituto Caro y Cuervo.
- [7] Prieto, P., & Roseano, P. (2010). Transcription of intonation of the Spanish language. In P. Prieto & P. Roseano (Eds.). Lincom Europa.
- [8] Ramsay, J., Hooker, G., & Graves, S. (2009). *Functional data analysis with R and MATLAB* (1st ed.). Springer. <https://doi.org/https://doi.org/https://doi.org/10.1007/978-0-387-98185-7>
- [9] Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
- [10] Vanrell, M. d. M., Feldhausen, I., & Astruc, L. (2018). The discourse completion task in Romance prosody research: Status quo and outlook. In I. Feldhausen, J. Fließbach, & M. d. M. Vanrell (Eds.), *Methods in prosody: A Romance language perspective* (pp. 191–227). Language Science Press.

---

## Pronominal Expression in Child Heritage Georgian

Cass Lowry

The Graduate Center, City University of New York, New York, USA

The expression of personal pronouns in languages which allow both overt and null pronouns—so-called pro-drop or zero anaphora languages—is a well-studied topic in the variationist tradition. Given that sentences with either an overt or a null personal pronoun are grammatical in these languages, the factors that condition which pronoun type a speaker chooses are probabilistic in nature. Subject personal pronouns in Spanish varieties in particular have a rich literature in which numerous grammatical, discourse, and genre variables have been found to condition the choice of subject pronoun (Bayley, 2013; Flores-Ferrán, 2007). More recent work has employed variationist methods to investigate the same phenomena in non-Indo-European languages as well, including Mandarin (Li & Bayley, 2018) and several Oceanic languages (Meyerhoff, 2000, 2009; Schnell & Barth, 2020).

This study investigates the use of phonetically overt and phonetically null (or zero) personal pronouns in monolingual and Heritage Georgian. Contrary to null subject languages of the Romance variety, such as Italian and Spanish, Georgian allows pro-dropping for all syntactic arguments, including objects as well as subjects. This study presents three studies investigating the linguistic constraints that determine the choice between overt and null pronouns in Georgian. First, a hand-coded analysis of a Georgian short story finds that the morphosyntactic features of person-number and the verb's tense-aspect-mood constrain the realization of pronouns. Study 2 replicates the results of the first experiment with a larger corpus analysis of 1,652 Georgian texts, and also finds syntactic position to affect pronoun realization, with the subject position favoring overt pronouns. Study 3 tests for these constraints in narratives produced by child Georgian heritage speakers ( $n = 26$ ) and age-matched baseline monolingual peers ( $n=30$ ). Contrary to what has often been found in pronoun production studies of bilingual speakers of other languages, Georgian heritage speakers produced more null subjects and fewer overt subjects than the baseline group. For pronouns in object position, the two groups did not differ. Taken together, the results suggest that overt

pronouns have a limited distribution and discursive use in Georgian, and that this general reliance on pro-drop is magnified in heritage grammars.

- Bayley, R. (2013). Variationist Sociolinguistics. In R. Bayley, R. Cameron, & C. Lucas (Eds.), *The Oxford Handbook of Sociolinguistics*. Oxford University Press. 10.1093/oxfordhb/9780199744084.013.0001
- Flores-Ferrán, N. (2007). A bend in the road: Subject personal pronoun expression in Spanish after 30 years of sociolinguistic research. *Language and Linguistics Compass*, 1(6), 624–652.
- Harris, A. C. (1981). *Georgian syntax: A study in relational grammar*. Cambridge University Press.
- Li, X., & Bayley, R. (2018). Lexical frequency and syntactic variation: Subject pronoun use in Mandarin Chinese. *Asia-Pacific Language Variation*, 4(2), 135–160.
- Meurer, P. (2021). The Georgian National Corpus. <http://gnc.gov.ge>
- Meyerhoff, M. (2000). The emergence of creole subject–verb agreement and the licensing of null subjects. *Language Variation and Change*, 12(2), 203–230.
- Meyerhoff, M. (2009). Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change*, 21, 297–317.
- Schnell, S., & Barth, D. (2020). Expression of anaphoric subjects in Vera'a: Functional and structural factors in the choice between pronoun and zero. *Language Variation and Change*, 32, 267–291.

---

## A dynamic neural model of the interaction between social and lexical influences on speech production: the case of retroflex sibilants in Taiwan Mandarin.

Noah Macey<sup>1,2</sup>, Michael Stern<sup>1</sup>, Sang-Im Lee-Kim<sup>3</sup>, Jason Shaw<sup>1</sup>

<sup>1</sup>Yale University, New Haven, CT, USA. <sup>2</sup>UC Berkeley, Berkeley, CA, USA. <sup>3</sup>Hanyang University, Seoul, Korea, Republic of

Speakers of Taiwan Mandarin navigate two types of stigmas associated with retroflex sibilants. Extreme retroflexion, which facilitates contrast with alveolar sibilants, is stigmatized as a marker of mainland China. Absence of retroflexion leads to merger with alveolar sibilants, an historically stigmatized pattern. Speakers who merge sibilants in some situations have been shown to de-merge when interacting with a non-merged experimenter (Lee-Kim & Chou 2022), a result we interpret as an effect of social expectation priming (e.g., Wade 2022). Our aim in this study is to model the cognitive mechanism underlying the interaction between contrast maintenance and social context.

Our approach builds on Stern & Shaw (2022/2023), who use a Dynamic Neural Field (DNF) (e.g., Schöner & Spencer 2016) to model contrastive hyperarticulation as a consequence of neural inhibition—a minimal pair competitor inhibits a portion of the DNF, effectively repelling speech planning away from the competitor. We incorporate social context by modeling expectation priming as an excitatory input to a DNF (see also Yi et al. 2023), driving neural activation towards social expectations.

We ran 500 simulations of retroflex productions for each of the three interactive conditions reported in Lee-Kim & Chou (2022) (see Table 1) as well as in corresponding non-interactive (no expectation priming) situations. We assume the relevant dimension of variation for retroflex sibilants is the length of the vocal tract cavity in front of the source of turbulence, i.e., front cavity length. We derived measures of front cavity length from the maximum amplitude frequency of retroflex fricative spectra in Lee-Kim & Chou's data. These measurements define Gaussian-shaped inputs to the DNF centered on 25 mm for the retroflex target; 13 mm for the alveolar competitor; and 40 mm (extreme retroflexion) for the social priming input. In each simulation, neural activation evolves over time under the influence of these inputs until an activation peak forms.

Figure 1 shows field evolution for retroflexes without minimal pair competitors (Fig. 1a), with minimal pair competitors (Fig. 1b), and with minimal pair competitors salient in the interactive context (Fig. 1c). Figure 2 shows box plots aggregating across 500 simulations of each condition. With no interlocutor (Fig. 2: left), there is a small difference between conditions, consistent with past work on contrastive hyperarticulation (Baese-Berk & Goldrick, 2009; Wedel et al. 2018). In the interactive condition, neural activation peaks at a significantly more retroflexed location due to social priming, but only in the salient competitor condition. The interaction arises because strong inhibition of the alveolar (in the salient competitor condition) suppresses activation of the speaker's typical retroflex, allowing the field to evolve towards the social expectation input. This pattern of results matches the trends reported Lee-Kim & Chou (2022), who found significantly more retroflexed sounds only when there was a contextually salient minimal pair competitor.

With increasing attention to the cognitive mechanisms through which sociolinguistic variation is processed (e.g., Chevrot, Drager, Foulkes 2018), DNF models show promise for integrating linguistic and social influences on variation using cognitively plausible mental representations of both.

- Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, 24(4), 527-554.
- Chevrot, J. P., Drager, K., & Foulkes, P. (2018). Editors' introduction and review: Sociolinguistic variation and cognitive science. *Topics in Cognitive Science*, 10(4), 679-695.
- Lee-Kim, S. I., & Chou, Y. C. (2022). Unmerging the sibilant merger among speakers of Taiwan Mandarin. *Laboratory Phonology*, 13(1).
- Schöner, G., & Spencer, J. P. (2016). *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press.
- Stern, M. C., & Shaw, J. A. (2022/2023). Neural inhibition during speech planning contributes to contrastive hyperarticulation. arXiv preprint arXiv:2209.12278 (accepted in *Journal of Memory and Language*).
- Yi, I., Bowern, C., & Shaw J.A. (2023). A dynamic field theory of (socio)phonetic convergence and divergence in sound change. *Linguistic Society of America Annual Meeting*, Boulder, Colorado.
- Wade, L. (2022). Experimental evidence for expectation-driven linguistic convergence. *Language*, 98(1).
- Wedel, A., Nelson, N., & Sharp, R. (2018). The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language*, 100, 61-88.

---

## Real-time interpretation of morphosyntactic variation across varieties and listeners

Zachary Maher<sup>1,2</sup>, Jan Edwards<sup>1</sup>, Jared Novick<sup>1</sup>

<sup>1</sup>University of Maryland, College Park, MD, USA. <sup>2</sup>Pennsylvania State University, University Park, PA, USA

Listeners are sensitive to the grammatical conditions of sociolinguistic variables[1], and there is evidence that phonological variation influences lexical processing[2]. But to what extent do listeners use their knowledge of sociolinguistic variation to guide sentence comprehension in real time? In this study, we examine how listeners use morphosyntactic variation to guide interpretation, depending on whether (1) the speaker uses African American Language (AAL) or Mainstream American English (MAE); and (2) the listener is a monodialectal speaker of MAE or a bidialectal speaker of AAL and MAE.

Participants were college students who spoke English as a first language and identified as either Black or white. Participants heard naturalistic audio examples of AAL and MAE and rated how frequently they use each variety. Black participants were included in the bidialectal group (n=36) if they responded in the upper half of the AAL usage scale, and white participants were included in the monodialectal group (n=31) if they responded at the bottom two levels of the AAL usage scale and top two levels of the MAE usage scale. The bidialectal group also accurately distinguished between correct and incorrect uses of habitual be[3], while the monodialectal group did not.

**Method:** Participants listened to sentences spoken by two speakers of AAL and two speakers of MAE (guise manipulation). Filler items used negative concord and null copula (or MAE equivalent sentences) to clearly associate each voice with AAL (or MAE) grammar. In critical items, a word segmentation ambiguity forced listeners to rely on the agreement morphology of the verb to determine if the subject of the sentence was singular or plural (e.g., The duck(s) swim in the pond). Crucially, in AAL, a singular interpretation is available, while in MAE, only a plural interpretation is available. Participants' eye movements were tracked as they listened to each sentence, and they selected the image that best matched their interpretation. Four images were on the screen: a singular image (one duck), plural image (two ducks), and two distractors.

**Results:** Picture selections were analyzed with logistic mixed-effects models and gaze patterns with cluster-based permutation analysis. Participants most frequently selected the plural image, which aligns with the grammar of both MAE and AAL. However, across groups, participants were more likely to look at (Fig. 1) and select (Fig. 2) the singular image in the AAL guise, compared to the MAE guise (ps<0.001), reflecting the increased availability of a singular interpretation in AAL. Additionally, bidialectal participants were overall more likely to look at and select the singular image, compared to monodialectal participants (ps<0.001). There was no interaction between guise and group.

**Discussion:** These results show that listeners can use their knowledge of variation to guide interpretation. Remarkably, this is true even for "monodialectal" MAE speakers, who are unlikely to produce regularized variants. This suggests that sociolinguistic repertoire extends beyond what might be observed in production[4], including to morphosyntactic variables, inviting further study on the ways in which listeners adapt to different sources of sociolinguistic variation in real-time processing.

1. Vaughn, C., & Kendall, T. (2018). Listener sensitivity to probabilistic conditioning of sociolinguistic variables: The case of (ING). *Journal of Memory and Language*, 103, 58-73.

2. King, S., & Sumner, M. (2014). Voices and variants: Effects of voice on the form-based processing of words with different phonological variants. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, No. 36).
3. Wolfram, W. (1982). Language knowledge and other dialects. *American Speech*, 57(1), 3-18.
4. Labov, W. (1973). Where do grammars stop. In *Report of the twenty-third annual round table meeting on linguistics and language studies* (pp. 43-88). Washington DC: Georgetown University Press.

---

## Race in the performance of a Baltimore persona

Aidan Malanoski

The Graduate Center, City University of New York, New York, USA

Dialect performance reflects indexical relationships between linguistic features and place (Johnstone et al. 2006). I discuss the role of language in the performance of the Hon, a persona based on white working-class women from 1950s Baltimore. I consider two questions: which linguistic features are involved in the performance of the Hon, and how does this performance differ between Black and white speakers? The present work constitutes one of the first studies of Baltimore English, and adopts a third-wave approach (Eckert 2012) to the growing literature on regional African American Englishes and their relationship to white varieties.

The Hon is the focus of HonFest, a Baltimore festival where many attendees dress up as Hons and adopt features of “Bawlmerese”, white working-class Baltimore English. HonFest’s main event is the Best Hon contest, where contestants compete for the title of Baltimore’s Best Hon. While attendees are mostly white, the organizers have pushed to increase diversity at the event, and several Black women participated in last year’s Best Hon Contest.

This study’s participants are seven Baltimore-area women (six identified as white/“Caucasian”, one as Black) who attended HonFest dressed as a Hon. Most were previous winners of the Best Hon contest. I conducted sociolinguistic interviews with two reading passages and two wordlists. Participants were asked to read the first passage and wordlist “normally” and the second passage and wordlist “as a Hon.” Presentation order was counterbalanced across participants.

Table 1 gives the target variables. I assume variants found in dialect respellings are stereotypes. Other variants shown by prior research (Hisley 1964, Labov et al. 2005) to be associated with Baltimore are assumed to be markers or indicators. Montreal Forced Aligner (McAuliffe et al. 2017) was used to align transcripts, and FAVE (Rosenfelder et al. 2022) was used to extract formant measurements.

Figure 1 compares the difference between the normal and Hon guise for the Black speaker, Ms. Annett, to the difference between guises for the white speakers. In the Hon guise, both Black and white speakers produce frontier GOOSE and MOUTH, stereotypes of Bawlmerese, but white speakers otherwise use a wider range of Bawlmerese stereotypes (fronted GOAT, monophthongal PRICE/TILE, tensed BATH, monophthongal TOIL). On the other hand, Ms. Annett dramatically raises THOUGHT in the Hon guise, suggesting that raised THOUGHT may be a stereotype of local speech among Black Baltimoreans. Ms. Annett also raises TRAP in the Hon guise, erasing the Bawlmerese short-a split. These data suggest that Ms. Annett’s performance does not target the same norms as the white speakers’. Indeed, before reading the passage and wordlist in the Hon guise, she said that she would read them in “[her] African American style.” Although we cannot know without comparing Ms. Annett to other Black speakers whether her performance is idiosyncratic or normative, it is likely that her performance targets a Black counterpart of Bawlmerese. Even if such a register is not yet established, performances by Ms. Annett and others like her will bring it into being.

Hisley, A.M. 1964. An historical analysis of the development of the Baltimore dialect. University of Maryland MA thesis.

Johnstone, Barbara, J., J. Andrus & A.E. Danielson. 2006. Mobility, indexicality, and the enregisterment of “Pittsburghese”. *Journal of English Linguistics* 34(2). 77–104.

Eckert, P. 2012. Three waves of variation study: The emergence of meaning in the study of variation. *Annual Review of Anthropology* 41. 87–100.

Labov, W., S. Ash & C. Boberg. 2005. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin/New York: Mouton de Gruyter.

McAuliffe, M., M. Socolof, S. Mihuc, M. Wagner & M. Sonderegger. 2017. Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. In *Proceedings of the 18th Conference of the International Speech Communication Association*, 498–502.

Rosenfelder, I., J. Fruehwald, C. Brickhouse, K. Evanini, S. Seyfarth, K. Gorman, H. Prichard & J. Yuan. 2022. FAVE (Forced Alignment and Vowel Extraction) program suite v2.0.2. Thomas, E.R. 2007. Phonological and phonetic characteristics of African American Vernacular English. <https://doi.org/10.5281/zenodo.7143517>.

---

Community embeddedness, social practice and cognition:  
Explaining was/were variation in an economically-deprived community

Katie Mansfield

University of Sheffield, Sheffield, United Kingdom

Variation in the past tense forms of the verb BE (e.g., I were, you was, they was) is found across the entire English-speaking world. Within Britain, was/were variation behaves differently across dialects, with differential levelling of the paradigm towards was and were (e.g., Britain, 2002; Durham, 2013; Rautionaho and Kaunisto, 2022). Previous studies have found that gender (e.g., Tagliamonte, 1998), ethnicity and social class (e.g., Cheshire and Fox, 2009) and community of practice engagement (Moore, 2011) affect was/were variation. Furthermore, Moore (forthcoming) proposes that the sociolinguistic vitality of levelled were in Bolton, Greater Manchester, is a consequence of indexical associations which span class, local personas, and relaxed, playful interactional moves.

Viewing linguistic variation as social practice has enabled variationists to better understand the social meaning of items like levelled was and were. However, questions remain about the constraints that cognitive factors play on our ability to vary morphosyntactic variation (Meyerhoff and Walker, 2013; Levon and Buchstaller, 2015). Work on phonological variation suggests that increased cognitive load correlates with speakers defaulting to their first acquired 'lect' (Sharma & McCarthy, 2018), but it remains unclear how cognitive constraints interact with identity presentation - especially in the domain of morphosyntax. To what extent does early acquisition of features like levelled was and were inhibit the extent to which we can vary their use?

This paper reports on the first study of was/were variation to examine the interaction between cognitive factors (shifting ability and working memory capacity), linguistic factors (such as polarity and grammatical person/number), and a range of social factors (such as social class, community embeddedness and community of practice). The data come from a school ethnography of 10-11 year olds in a socio-economically deprived area in Leicester (an understudied English Midlands area; Braber and Robinson, 2018). Children were recorded during semi-structured interviews, in peer/play settings, and during a dual-task speech-based experiment which measures their language output during varying levels of cognitive load. The children's working memory was also tested via a novel interactive superhero game (Mansfield, 2022).

Analysis of 1258 tokens of was and were shows a dominant pattern of levelled was in Leicester. Regression analysis reveals correlations with the expected social and linguistic factors (subject type and polarity affect frequencies of levelling and lower class speakers use more levelled was). However, anomalies in the social class results reflect an interaction between class and community engagement, suggesting that the strongest correlation is between levelled was and community embeddedness. This cognitive data allows me to nuance this finding in an unprecedented way: preliminary analyses suggest that, in conditions of high cognitive load, children who have poor working memory and high community embeddedness are less able to switch from nonstandard was to standard were than those children with good working memory and high community embeddedness. I consider this tentative finding in the context of classroom interaction where young people are expected to switch between home dialects and standard English on demand.

- Braber, N., & Robinson, J. (2018) *East Midlands English (Dialects of English)*, Vol. 15. Berlin: De Gruyter Mouton.
- Britain, D. (2002) 'Diffusion, Levelling, Simplification and Reallocation in Past Tense BE in the English Fens', *Journal of Sociolinguistics*, 6(1): 16-43.
- Cheshire, J., & Fox, S. (2009) 'Was/Were Variation: A Perspective from London', *Language Variation and Change*, 21(1): 1-38.
- Durham, M. (2013) 'Was/Were Alternation in Shetland English', *World Englishes*, 32(1): 108-128.
- Levon, E., & Buchstaller, I. (2015) 'Perception, Cognition, and Linguistic Structure: The Effect of Linguistic Modularity and Cognitive Style on Sociolinguistic Processing', *Language Variation and Change*, 27(3): 319-348.
- Mansfield, K. (2022) *Calling all Superheroes*, Gorilla Experiment Builder (Gorilla Game Builder).
- Meyerhoff, M., & Walker, J.A. (2013) 'An Existential Problem: The Sociolinguistic Monitor and Variation in Existential Constructions on Bequia (St. Vincent and the Grenadines)', *Language in Society*, 42: 407-428.
- Moore, E. (2011) 'Interaction Between Social Category and Social Practice: Explaining Was/Were Variation', *Language Variation and Change*, 22: 347-371.
- Moore, E. (forthcoming) *Socio-Syntax: Exploring the Social Life of Grammar*. Cambridge: Cambridge University Press.
- Rautionaho, P., & Kaunisto, M. (2022) 'Was/Were Variation with Subject Pronouns We, You and They in Recent British English: Towards Standard Uses?', in *Social and Regional Variation in World Englishes*, eds. P. Rautionaho, H. Parviainen, M. Kaunisto and A. Nurmi. London: Routledge, 43-65.
- Sharma, D., & McCarthy, K. (2018) 'Attentional Load and Style Control', *University of Pennsylvania Working Papers in Linguistics*, 24(2): Article 15.
- Tagliamonte, S. (1998) 'Was/Were Variation Across the Generations: View From The City of York', *Language Variation and Change*, 10: 153-191.

---

Discourse/Pragmatic Variation in the Tag System of Southern Arizona Spanish  
Brandon Martínez

University of Arizona, Tucson, USA

Tags are discourse markers used in speech to accomplish a variety of pragmatic functions (Gómez González, 2014, 2012), whether eliciting information from an interlocutor (addressee-oriented), highlighting elements of the speaker's experience (speaker-oriented), or managing structural aspects of the discourse (exchange-oriented). Structurally, tag questions consist of an anchor clause with an appended tag which can be variably positioned (Allerton, 2009); Furthermore, tags exhibit either a positive or negative polarity (Allerton, 2009; Gómez González, 2014; Pichler, 2013), display intonational differences (Allerton, 2009; Gazdik, 2022; Kimps et al, 2014), and vary in terms of the mood of the anchor clause (Gómez González, 2014, 2012; Kimps et al, 2014).

The current project is an exploratory study investigating tags in the bilingual Spanish of Southern Arizona through a variationist lens (Tagliamonte, 2012, 2006). Due to its proximity to the border, this variety is similar linguistically to Sonoran Spanish (Bessett, 2018, 2015). On the other hand, as with other Borderlands varieties, the influence of American English on Southern Arizona Spanish have been shown in some studies to have an interlingual effect on morphosyntactic variation (Limerick, 2018; Llompart, 2016), while others have shown no cross-lingual influence (Bessett 2018, 2015). That being said, interlingual effects on discourse/pragmatic variation have to my knowledge not been previously investigated in this variety, and warrant further study. Utilizing data taken from the Corpus del Español en el Sur de Arizona (Carvalho, 2012-), a collection of sociolinguistic interviews from Southern Arizona, the study's participant population includes 30 individuals who are natives or long-term residents of the city of Tucson, primarily of Mexican origin. Tags are split into two categories, following Pichler's (2013) hypothesis that the discourse functions of tags will vary based on their polarity: The negative polarity tags "no" and "qué no," compared to the positive polarity tags "sabes" and "verdad." Furthermore, this study investigates the incoming positive variant "you know," a loanword from English, in order to establish the role which it plays (or will come to play) in the pragmatic system of this bilingual variety as it grammaticalizes.

Pilot data taken from the first eight interviews (7 female, 1 male; total tokens n=158) demonstrate that both Spanish "no" (n=84) and English "you know" (n=58) dominate the system numerically, although "you know" is restricted in scope to only a few speakers. In terms of discourse function, speakers of Southern Arizona Spanish seem to primarily utilize tags to convey speaker- or exchange-oriented pragmatic functions, both of which are non-conducive to an interlocutor response, and indicative of grammaticalization occurring in the tag question system (cf. Pichler, 2021). What's more, as further evidence of this phenomenon, tags in this variety tend to occur in a medial position (n=98) instead of the canonical final position (n=35), and with more tokens displaying falling (n=92) rather than rising (n=44) intonation. The above data is taken as preliminary evidence of grammaticalization occurring in this variety, with implications for future developments for the tag system of Southern Arizona Spanish.

Allerton, D. (2009). Tag questions. In G. Rohdenburg & J. Schlüter (eds.) *One Language, Two Grammars: Differences between British and American English* (pp. 306-323). Cambridge: Cambridge University Press.

Bessett, R. (2018). Testing english influence on first person singular "yo" subject pronoun expression in sonoran spanish. In J. MacDonald (ed.) *Contemporary Trends in Hispanic and Lusophone Linguistics: Selected papers from the Hispanic Linguistic Symposium 2015* (pp. 355-372). Amsterdam: John Benjamins.

Bessett, R. (2015). The extension of *estar* across the mexico-us border: evidence against contact induced acceleration. *Sociolinguistic Studies* 9 (4), pp. 421-443.

Carvalho, A. (2012-). *Corpus del Español en el Sur de Arizona (CESA)*. University of Arizona. Available from <https://cesa.arizona.edu>.

Gazdik, A. (2022). On the prosodic realization of spanish ¿no? – tags from a pragmatic perspective. *Isogloss* 8 (2), pp. 1-19.

Gómez González, M. (2014). Canonical tag questions in english, spanish and portuguese. *Languages in Contrast* 14 (1), pp. 93-126.

Gómez González, M. (2012). The question of tag questions in english and spanish. In I. Moskowich & B. Crespo (eds.) *Encoding the Past, Decoding the Future: Corpora in the 21st Century* (pp. 59-95). Cambridge: Cambridge Scholars Publisher.

Kimps, D. et al. (2014). The speech functions of tag questions and their properties: a comparison of their distribution in colt and llc. In K. Davids et al (eds.) *Corpus Interrogation and Grammatical Patterns* (pp. 321-350). Amsterdam: John Benjamins.

- Limerick, P. (2018). Variable clitic placement in u.s. spanish. In J. MacDonald (ed.) *Contemporary Trends in Hispanic and Lusophone Linguistics: Selected papers from the Hispanic Linguistic Symposium 2015* (pp. 49-70). Amsterdam: John Benjamins.
- Llompart, M. (2016). Subject-verb order variation with unaccusative verbs of change of location in Mexico and southern Arizona. *Studies in Hispanic and Lusophone Linguistics* 9 (1), pp. 143-173.
- Pichler, H. (2021). Grammaticalization and language contact in a discourse-pragmatic change in progress: the spread of *innit* in London English. *Language in Society* 50 (5), pp. 723-761.
- Pichler, H. (2013). Negative polarity question tags (neg-tags). In *The Structure of Discourse-Pragmatic Variation* (pp. 169-212). Amsterdam: John Benjamins.
- Tagliamonte, S. (2012). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley Blackwell.
- Tagliamonte, S. (2006). *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

## Social perceptions of Traditional New Mexican Spanish paragoge across communities

Aurora Martinez Kane

University of California, Berkeley, Berkeley, USA

Traditional New Mexican Spanish (TNMS), spoken by descendants of 16th century Spanish settlers in New Mexico and southern Colorado known as *Hispanos*, is the oldest surviving variety of US Spanish, though it is often described as moribund (Bills, 1997). However, TNMS is only one of many Spanishes now found in its region; as a border state, New Mexico also has a growing population of Mexican immigrants and their descendants that now comprises its largest Latinx group (Ennis et al., 2011). Though Mexicans share a common language with *Hispanos*, there are intergroup linguistic differences and sometimes cultural tensions (Bills & Vigil, 2008), but no formal linguistic attitude studies have ever been done on this topic. This project aims to examine covert language attitudes toward a TNMS feature on the parts of both TNMS speakers and Mexican Spanish participants using the matched guise technique (cf. Lambert et al. 1960).

One of TNMS's most unique features is *paragoge*, in which a vowel is sometimes added prepausally to words with a stressed final syllable ending in /l/, /t/, /n/, or /s/ (i.e. *mal* 'bad' → *male/mali*). Given descriptions in past literature of *paragoge* as a unique and recognizably TNMS feature (i.e. Wilson, 2015), this variable seems ideal for this first-ever TNMS matched guise test. Stimuli for the test will be pulled from spontaneous speech from the NMCOS (Bills & Vigil, 2007), one of the only available corpora of TNMS data. Two short sections of spontaneous speech on a neutral topic from the same speaker, one with use of *paragoge* and the other without, will be presented to roughly thirty Latin American Spanish-speaking participants and thirty TNMS-speaking participants, with filler speech from other Spanish speakers presented in between.

Participants will be asked to evaluate the voices they hear for qualities such as accent, status, solidarity, and rurality, and this data will be analyzed quantitatively for any possible trends. For example, in some past matched guise work on stigmatized variants and varieties, out-group participants have rated the variant or varieties in question lowly for accent, status, and solidarity, while in-group participants have rated them similarly but with a significantly higher rating for solidarity, especially for male speakers (i.e. Chappell, 2016; Luhman, 1990). It remains to be seen if the evaluations in this project will pattern similarly to this in any way, or indeed if any significant trends at all will be found. Regardless, any findings will offer insight into how speakers of local Spanishes view TNMS features.

TNMS is a highly endangered variety of Spanish, described as being destined to die out within the next few generations (i.e. Arnold & Martínez-García, 2019); furthermore, very little sociolinguistic analysis has been published on TNMS, with nearly nothing having been published on *paragoge*. Therefore, this study would represent the first attitudinal study of a unique feature of this endangered variety of Spanish. Taken together, these attitudinal findings can point toward a better understanding of the patterns of intersection amongst dialect contact, language attitudes, and social factors.

- Arnold, P., & Martínez-García, M. T. (2019). Traditional New Mexican Spanish: The past, present, and future. In G. L. Thompson & S. M. Alvord (Eds.), *Contact, community, and connections: Current approaches to Spanish in multilingual populations* (pp. 175-202). Vernon Press.
- Bills, G. (1997). New Mexican Spanish: Demise of the earliest European variety in the United States. *American Speech*, 72(2), 154-171. Retrieved from <https://www.jstor.org/stable/455787>
- Bills, G., & Vigil, N. (2007). *New Mexico and Colorado Spanish Survey [Data set]*. University of New Mexico. <http://hdl.handle.net/1928/20424>
- Bills, G., & Vigil, N. (2008). *The Spanish language of New Mexico and Southern Colorado: A linguistic atlas*. Albuquerque, NM: University of New Mexico Press.
- Chappell, W. (2016). On the social perception of intervocalic /s/ voicing in Costa Rican Spanish. *Language Variation and Change*, 28(3), 357-378. doi:10.1017/S0954394516000107
- Ennis, S.R., Rios-Vargas, M., & Albert, N.G. (2011). *The Hispanic population: 2010, 2010 census briefs*. United States Census Bureau. Retrieved from <https://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf>



- Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology*, 60(1), 44-51.
- Luhman, R. (1990). Appalachian English stereotypes: Language attitudes in Kentucky. *Language in Society*, 19(3), 331-348.
- Wilson, D. V. (2015). Panorama del español tradicional de Nuevo México. *Informes del Observatorio / Observatorio Reports*. <https://doi.org/10.15427/OR012-06/2015SP>

---

## Sociolinguistic study of partial question structures in Hexagonal French: conclusions from experimental data

Corentin Mazet  
Indiana University, Bloomington, USA

French allows for a number of different question structures, which are semantically equivalent (Kayne, 1972; Danjou-Flaux and Dessaux, 1976; Ashby, 1977). However, these questions are not in free-variation and are considered to vary in pragmatic value (Coveney, 1995, 1996, 1997; Boeckx 1999; Quillard, 2001; Beyssade, 2006). Specifically, they seem to vary in register (Coveney, 2011; Guryev, 2017), and degrees of formality seem to correlate with specific structures. Most partial questions fall into one of four types: questions with inversion between subject and verb, *est-ce que* particle questions, *in situ*/intonation questions, and fronting without inversion (Coveney, 2011; Guryev, 2017). These question structures exhibit a variety of degrees of syntactic movement, from maximal movement (inversion questions) to minimal movement (*in situ*/intonation questions).

The pragmatic variation between the question structures is well studied but other factors might play a role in this variation. Quillard (2001) found that younger speakers are more likely to use *in situ* structures, whereas older speakers are more likely to use inversion structures, which was then confirmed by Guryev (2017). Looking at gender, Guryev (2017) also found a slight preference for men to use inversion and women to use fronting, although not statistically significant. Finally, other studies have found diachronic variation over the 20th century (Farmer, 2015), hinting at a change in progress in the language.

In this study, I look at various factors that may influence the choice of partial question structure through experimental data from the production of 31 Hexagonal French speakers. By way of an elicitation task with pragmatic priming, speakers created partial questions, exhibiting all four structures mentioned above. Their productions were then coded for question structure, speaker, verb, WH-word, register, speaker age, speaker gender, speaker profession, and speaker region of origin. Multiple regressions were performed on the data, probing the influence of the pragmatic and sociolinguistic factors that might correlate with specific question structures. As expected from previous literature on the topic, register played a major role in the results found in this study. However, age, gender, and region of origin were also found to be statistically significant predictors for specific question structures, while other factors did not reach significance threshold. Overall, younger speakers seemed to favor structures showing less movement for all registers, which supports previous research. Looking at gender, women seem to favor structures with less movement, whereas men seem to favor them, across registers. Finally, speakers from different regions of France seem to exhibit different patterns with regards to the relationship between register and question structure.

These results confirm previous findings and add to them. Statistical significance was reached for three sociolinguistic factors, including one that had not been studied in the previous literature. In addition, general tendencies across results seem to indicate that different social groups use structures distinctively, and that they pair each structure with each register in a different manner.

- Ashby, W. (1977). Interrogative forms in Parisian French. *Semantica*, 4, 35-52.
- Beyssade, C. (2006). La structure de l'information dans les questions: quelques remarques sur la diversité des formes interrogatives en français. *LINX*, 55, 173-193.
- Boeckx, C. (1999). Decomposing French questions, in J. Alexander, N-R Han & M. Minnick Fox (eds), *Proceedings of the 23 rd Annual Penn Linguistics Colloquium*, Philadelphia, PA: University of Pennsylvania Working Papers in Linguistics, pp. 69-80.
- Coveney, A. (1995). The use of QU- final interrogative structure in spoken French. *Journal of French Language Studies* 5, 143-171.
- Coveney, A. (1996). *Variability in Spoken French: interrogation and negation*. Bristol: Intellect Books.
- Coveney, A. (1997). L'approche variationniste et la description de la grammaire du français : le cas des interrogatives. *La variation en syntaxe*. *Langue française* 115, 88-100.
- Coveney, A. (2011). L'interrogation directe. *Travaux de linguistique*, 63, 112-145.
- Danjou-Flaux, N. and Dessaux, A-M. (1976). L'interrogation en français: Données linguistiques et traitements transformationnels. *Grammaire transformationnelle: syntaxe et lexique*. Études réunies par Jean-Claude Chevalier. Publications de l'Université de Lille, 139-231.
- Farmer, K. L. (2015). *Sociopragmatic variation in yes/no and wh-interrogatives in hexagonal French: A real-time study of French films from 1930 to 2009*. Dissertation. Indiana University, USA.

- Guryev, A. (2017). La forme des interrogatives dans le Corpus suisse de SMS en français. Étude multidimensionnelle. Thèse de doctorat. Université de Neuchâtel, Suisse, et Université de Paris 3 Sorbonne Nouvelle, France.
- Kayne, R.S. (1972). Subject Inversion in French Interrogatives. In: J. Casagrande and B. Saciuk (eds.) *Generative Studies in Romance Languages*. Newbury House, Rowley, Mass., 70-126.
- Quillard, V. (2001). La diversité des formes interrogatives : comment l'interpréter ? *Langage et société*. 95, 57-72.

---

Pedro Castillo, la sierra, and la costa:

Intra-Speaker variation and dynamic bricolage in Peruvian vowel spaces

Alden McCollum

New York University, USA

Popular Peruvian discourse imagines Peru as divided into three regions: la sierra (the Andes), la selva (the Amazon), and la costa (the coast). In this project, I understand these regions as imagined communities (Anderson, 1983/2016), and consider how vowel spaces are recruited in stylistic performances that are in dialogue with a constructed opposition between la sierra and la costa, wherein la sierra is associated with indigeneity and Quechua, and la costa with whiteness and 'standard' Spanish.

Prior work on language contact between Spanish and trivocalic Quechuas finds that some Spanish speakers in Andean contexts exhibit partial or complete mergers of /i/ with /e/ and /o/ with /u/ (Lipski, 2015). Popular narratives tie such speakers to la sierra and view them through a deficit lens of "confusión vocálica" (de los Heros, 2016), wherein sierra vowel spaces are portrayed as the product of an 'inability' to produce vocalic distinctions. Such narratives depart from observed realities, erasing (Irvine & Gal, 2000) both inter- and intra-speaker variation. In the current project, I reject this deficit-based framing and ask instead: What might we find if we view these vowel spaces as a stylistic resource that is flexibly recruited in the service of identity projects?

In approaching this question, I use quantitative sociophonetic data to analyze the dynamic vowel space of former Peruvian president Pedro Castillo (president from 7/28/2021 through 12/7/2022), focusing on how his vowel space varies across and within three discourse events: a May 2021 debate with Keiko Fujimori; a September 2021 mensaje a la nación; and Castillo's December 2022 attempted dissolution of congress.

I operationalize merger via Pillai scores for the front and back vowel spaces, and run a series of Generalized Additive Mixed Models (GAMMs) to model how Pillai scores shift over time within individual discourse events. I consider Castillo's stylistic performance in relation to that of his opponent in Peru's 2021 election cycle, Keiko Fujimori, and ultimately identify two bricolage-like patterns in the vowel space, which I propose to label contrastive mirroring and additive mirroring. I discuss potential explanations for why it is that Castillo exhibits more contrastive mirroring (Figure 1) while Keiko[1] exhibits more additive mirroring (Figure 2).

In addition to the dynamic stylistic work observed in the vowel spaces here, another finding of note is that /o/ and /u/ appear to be underlyingly merged (Pillai score below .03 at the level of each discourse event) for both Keiko and Castillo. This finding is surprising in Keiko's case, and raises interesting questions about the role of racialization in the (il)legibility of 'accented' speech, given that Castillo is routinely described as exhibiting "confusion vocálica" while Keiko is not, despite the fact that both speakers exhibit this merger in the back vowel space.

[1] I refer to her as "Keiko" rather than "Fujimori" because this is the name she goes by politically and also because "Fujimori" with no first name attached to it is more often used to refer to Keiko's father, Alberto Fujimori.

Anderson, B. (2016). *Imagined communities: Reflections on the origin and spread of nationalism* (Revised edition). Verso. (Original work published in 1983.)

de los Heros, S. (2016). Humor étnico y discriminación en La paisana Jacinta. *Pragmática Sociocultural*, 4(1), 74–107.

Irvine, J. T., & Gal, S. (2000). Language Ideology and Linguistic Differentiation. In P. V. Kroskrity (Ed.), *Regimes of Language: Ideologies, Politics, and Identities*. School of American Research Press.

Lipski, J. M. (2015). Colliding Vowel Systems in Andean Spanish. *Linguistic Approaches to Bilingualism*, 5(1), 91–121. <https://doi.org/10.1075/lab.5.1.04lip>

---

Age and gender-mediated patterns in sociolinguistic perception – Comparing US and UK

listeners' evaluative reactions to Tyneside voices

Johanna Mechler

University of Duisburg-Essen, Germany

Research on the sociolinguistic monitor (Labov et al. 2011) suggests that evaluative reactions to voices regarding perceived professionalism are influenced by a variable's social salience and (social) characteristics of the listener (Levon and Buchstaller 2015). While we know that the socio-indexical meaning of (ing) differs in the US (Campbell-Kibler 2007) and the UK (Levon and Fox 2014), we lack comparative data from studies with the same

methodological setup across these locales (but see Schlee et al. 2017). My study aims to address this gap by exploring how the perception of (ing) (Hazen 2006) is mediated by age and gender, when US and UK listeners rate British speakers of English.

This paper reports on listener responses to 16 naturally occurring samples drawn from eight speakers from the North-East of England, two women and two men in their 20s and two women and two men in their 40s. Short guises were manipulated acoustically to create two stimuli per speaker: one with only alveolar and one with only velar variants of (ing). Careful sampling of a large corpus of sociolinguistic interviews (Corrigan et al. 2012) ensured that all stimuli-rich guises were stringently controlled for topic (hobbies), sound quality, and f0. Guises were presented in the context of a job application as a local radio host. Two participant pools (US and UK listeners) were asked to rate speakers on discrete 6-point Likert scales for attributes such as “friendly” and “professional” and to estimate speakers’ ages and places of residence (city/state/country).

Data was collected using PsychoPy with similar setups for both listener groups; trials were complemented with locally targeted post-experimental questionnaires. In-person data was collected in Columbus, Ohio (n=24), and Newcastle-upon-Tyne, England (n=21). All participants (aged 18-71) were native speakers of English and lived in the US or UK. Additional data are currently collected online in both locations.

Linear and ordinal regression models on age estimates (n=665) and professionalism ratings (n=622) reveal four main findings:

1. Listener evaluations of speaker professionalism do not vary by geographic region (US/UK).
2. Listeners in both locales react to real-life age differences in speakers’ voices, as reflected in higher age estimates for older speakers (p<0.001).
3. Perceptions of professionalism are highly contingent on both speaker age and gender: speakers in their 40s receive higher professionalism ratings (p<0.001), and older male speakers are perceived as more professional (p<0.05), even when they consistently produce vernacular [n] (cf. Conn 2005).
4. Listener age and self-reported gender play a mediating role on social evaluations (cf. Levon et al. 2021) with older male respondents rating speakers as less professional than younger female listeners (p<0.01).

Overall, the study reveals the intersectional nature of age and gender in sociolinguistic perception, providing a more finely differentiated picture of the ways in which age- and gender-related stereotypes impact speaker evaluation. These findings further provide a perceptual lens that allows insights into how indexicalities of stable variables may motivate differences in individual linguistic choices across the life-span.

- Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American Speech* 82(1). 32-64.
- Conn, Jeffrey Christopher. 2005. Of “moice” and men: The evolution of a male-led sound change. University of Pennsylvania ProQuest Dissertations Publishing. IRCS Technical Reports Series 5.
- Corrigan, Karen P., Isabelle Buchstaller, Adam Mearns, and Hermann Moisl. 2012. The Diachronic Electronic Corpus of Tyneside English. Newcastle University. <https://research.ncl.ac.uk/dectec>
- Hazen, Kirk. 2006. IN/ING Variable. In *Encyclopedia of Language and Linguistics*, 2nd ed., Vol. 5, Keith Brown (ed). Oxford: Elsevier. 581-584.
- Labov, William, Sharon Ash, Maya Ravindranath, Tracey Weldon, Maciej Baranowski, and Naomi Nagy. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15(4). 431-463.
- Levon, Erez, Devyani Sharma, Dominic J. L. Watt, Amanda Cardoso, and Yang Ye. 2021. Accent Bias and Perceptions of Professional Competence in England. *Journal of English Linguistics* 49(4). 355-388.
- Levon, Erez, and Isabelle Buchstaller. 2015. Perception, cognition, and linguistic structure: The effect of linguistic modularity and cognitive style on sociolinguistic processing. *Language Variation and Change* 27(3). 319-348.
- Levon, Erez, and Sue Fox. 2014. Social salience and the sociolinguistic monitor: A case study of ING and TH-fronting in Britain. *Journal of English Linguistics* 42(3). 185-217.
- Schlee, Erik, Nicholas Flynn, and William Barras. 2017. Regional diversity in social perceptions of (ing). *Language Variation and Change* 29(1). 29-56.

---

## Voicing Expertise: Exploring Strategic Use of Vowel Variants in the English Pronunciation of Chinese Language Instructors

Yunbo Mei

National University of Singapore, Singapore, Singapore

Building upon previous studies discussing the intertwined correlations between speakers' accent and identity (Bolton & Kwok, 1980; Cutler, 2014; Jenkins, 2003; Lippi-Green, 2012), this study employs a variationist approach to examine the agency of Chinese English language teachers (CELTs) in negotiating their professional identities. Instead of a loose association between identity and accent, detailed sociophonetic analyses reveal that these teachers construct their desired self-representations through the strategic use of linguistic resources.

15 CELTs were recruited for this study, and sociolinguistic interviews were conducted to collect 1253 vowel tokens, including TRAP, DRESS, MOUTH, PRICE, FACE, and GOAT. The selection of these vowels is based on

their relevance to Mandarin-accented English, particularly in terms of the smaller distinction between TRAP and DRESS, and the reduced tongue movements in diphthong pronunciation which is rather evident in those diphthongs with counterparts in Chinese. Given Chinese Second language learners' general difficulties with these vowels, the TRAP-DRESS distinction and the tongue movements of MOUTH, PRICE, FACE, and GOAT serve as potential variants that may index professionalism among CELTs. The F1 and F2 values of all tokens were extracted at 25%, 50%, and 75% of the vowel's duration using Praat (Boersma & Weenink 2009), followed by normalization using the Lobanov Metric (Lobanov, 1971) in NORM (Thomas & Kendall, 2015). To quantify the distinction variability among participants, a Pillai score was computed for each individual using the MANOVA test in R (R Core Team, 2013). Additionally, the Euclidean distance between the nucleus and off-glide over the duration of a diphthong was calculated in R to represent diphthongal movements. Mixed-effects linear regression modeling in R using lmerTest was employed to identify potential social predictors of the F1 and F2 values of these vowels.

The sociophonetic analyses reveal that participants' perceptions of the relationship between teacher qualifications and native-like/first-language-influenced English accent can predict their own pronunciation patterns. Specifically, participants who view Mandarin-accented pronunciation as undesirable for an English teacher demonstrated a more distinct TRAP and DRESS nuclei (Figure 1). Moreover, participants who consider native-like pronunciation as a benchmark for English teachers' expertise exhibited larger tongue movements in the pronunciation of PRICE (Figure 2) and MOUTH (Figure 3). Both groups employed standard variants to distance themselves from the perception of Mandarin-accented English and displayed a stronger ideological association between nativeness and teacher qualifications. Therefore, the use of robust DRESS-TRAP nuclei and larger tongue movements in MOUTH and PRICE can be interpreted as strategies employed by these teachers to distance themselves from a "non-native" identity, which is often stigmatized within the English language teaching community. In summary, the utilization of DRESS-TRAP nuclei, MOUTH, and PRICE as stylistic resources allows participants to construct a "native English speaker" persona and signify expertise in language teaching. These findings contribute to a deeper understanding of the social meanings conveyed by vowels among CELTs and add to the emerging body of literature on variationist studies of pronunciation among non-native English speaker teachers.

Bolton, K., & Kwok, H. (1990). The dynamics of the Hong Kong accent: Social identity and sociolinguistic description. *Journal of Asian Pacific Communication*, 1(1), 147-172.

Cutler, C. (2014). Accentedness, "passing", and crossing. In J. Levis & A. Moyer (Eds.), *Social Dynamics in Second Language Accent*. 145-167. Boston, MA: DeGruyter. <http://dx.doi.org/10.1515/9781614511762.145>

Jenkins, J. (2003). *World Englishes: A Resource Book for Students*. London: Routledge.

Lippi-Green, R. (2012). *English with an Accent: Language, Ideology, and Discrimination in the United States* (2nd ed.). Routledge.

## Linguistic Pride or Assimilation? An Examination of Lexical Change, Linguistic Practices, and Construction of Ethnic Identity among Chaoui Friendship Networks in Batna City, Algeria

Haroun Melgani

Oum El Bouaghi University, Oum El Bouaghi City, Algeria

Algeria is considered to be one of the most culturally and linguistically diverse countries in North Africa. It is a multilingual nation that is inhabited by communities of speakers who identify with different ethnic groups, and speak different linguistic varieties, such as Berber, Arabic, French, and korandjé. The long term contact between Berbers and other groups-eg., Arabs, Turkish, French, Romans-moulded the sociolinguistic situation of Algeria for centuries and paved the way for the emergence of new Koinés and establishment of new ethnolects and ethnolinguistic practices. By the same token, the massive migration movements into Batna City (Algeria) in the last three decades, coupled with extensive interethnic contact between Chaoui (Berber) speakers and their Arabic counterparts changed the socioethnic profile of the city, both culturally and linguistically, and gave rise to novel sociolinguistic outcomes across various social categories. Daily interethnic communication between migrant adolescents of Chaoui descent and their sedentary Arabic counterparts engendered cases of ethnic tension and yielded new speech practices and linguistic developments in the city. Grounded in variationist sociolinguistics, and theoretically informed by social constructivist paradigm, this quantitatively and ethnographically orientated research project seeks to examine dynamics of lexical variation and change and ethnically based salient linguistic practices across diametrically friendship networks in local areas in Batna city. By the same token, this research seeks to investigate the extent to which the linguistic variation and practices adopted by the two friendship networks co-vary with several, interrelated, sociocultural, educational, and ideological forces. Intensive ethnographic observations were carried out to explore the sociocultural composition of two ethnically based friendship networks, which vary on several sociocultural and linguistic grounds. A variety of research tools, including social network questionnaires, participant-observations, note taking and ethnographic interviews were incorporated to solicit socioethnic and sociolinguistic data about the two ethnic friendship networks under investigation. A set of Network-based analyses

and field-observations were carried out to gauge both networks' ethnic cohesiveness, dynamics of lexical change and linguistic practices. The results showed that ethnic strength and cohesiveness correlate, significantly, with dynamics of lexical change in Chaouia dialect. Chaoui adolescents with strong Chaoui ties prefer to adopt Chaoui words, whilst respondents with weak ties are less immune to Arabic influence and, hence, tend to be more amenable to use Arabic loanwords. Also, the research findings indicated that ethnically dense, many-stranded friendship networks favor dialect maintenance, whereas weak and multiethnic friendship networks favor speech innovation and diffusion. In urban interethnic settings, migrant Chaoui youngsters use various Chaoui linguistic practices as a socio-indexical function of membership, ethnic loyalty, cousinhood and regional identity, whereas urban Chaoui youngsters code-switch between local Arabic forms and Chaoui dialectal features to signal biethnic identity and urbanity. Non-Chaoui urban youngsters adopt salient Chaoui speech features to index complex personal stances and social affiliations, such as masculinity, assertiveness, and social distance from Chaouia culture. This study, also, discusses the role of speakers' sociopsychological perspectives in maintaining, loosening, reshaping, or subverting boundaries between ethnic groups in interethnic contact situations in Batna City.

---

## Learning to be (un)hip in panel data - Exploring quotative be like across the adult lifespan

Anne-Marie Moelders, Isabelle Buchstaller

Sociolinguistics Lab, University of Duisburg-Essen, Essen, Germany

Much of linguistic research relies on the critical period hypothesis, the assumption that “speech patterns are largely fixed by early adulthood” (Wagner 2012a:372). If speakers retain their early adult grammars, age-related variation in the community can be interpreted as a diagnostic for diachronic change (Sankoff 2005; Wagner 2012a,b). Only recently has the critical period hypothesis been subject to rigorous empirical testing on the basis of panel data. The relatively young field of lifespan research has revealed that changes do occur in language use across the lifespan, including into the older age ranges (Beaman 2022, Mechler et al. 2022, Buchstaller et al. 2017, Harrington et al. 2000). Sociolinguists have interpreted these findings as evidence that our linguistic habits are more flexible than a strict interpretation of the critical period would suggest (Wagner and Sankoff 2011; Sankoff and Blondeau 2007; *inter alia*).

To date, we lack consolidated empirical data about the age-related patterns and the conditioning factors of linguistic malleability across the adult lifespan. This is because the vast majority of panel-based research relies on relatively narrow windows of the human lifespan. The present paper aims to address this issue. We report on a panel dataset of 22 speakers from the Northeast of England who were grouped into five age cohorts and re-recorded an average of 7 years later. These age bands, when considered consecutively, cover the adult lifespan as a whole, which allows testing for age effects in linguistic lability during ongoing language change. The variable we investigate is the quotative system, the site of far-reaching changes due to the widely-reported newcomer *be like* (Buchstaller 2011; Tagliamonte & D’Arcy 2007). While younger age brackets have been reported to increment the variant to skyrocketing rates (Tagliamonte & D’Arcy 2007), the extent to which *be like* propagates into the linguistically sensitive middle age brackets is an unresolved question. Singler (2004) and Buchstaller (2013) provide apparent time data that suggests retrenchment amongst speakers in the mid-30s towards more conservative forms, whereas Wagner and Tagliamonte (2017) show that the majority of their speakers (ages 16-28) maintain high *be like* rates. Our analysis adds a lifespan perspective to the discussion. Our analysis of 1,351 tokens demonstrates that, rather than taking part in this community-wide change, many speakers – in particular those in their 30s – are moving away from innovative *be like* (as illustrated in the Figure overleaf). Our panel analysis further reveals that while frequencies of *be like* are malleable into the middle age ranges, the constraints that condition the variable remain intact, resulting in a shift in variable frequencies in otherwise stable speaker grammars (see also Tagliamonte and D’Arcy 2007). The analysis suggests that the retrenchment in middle age (Downes 1984), mainly postulated for stable variables, is also observable in changes in progress. We hypothesize that such “correction” away from incoming forms is particularly acute for variables that are associated with youthful styles, or considered inappropriate in more monitored styles (Buchstaller 2015, see also Conn 2005), making them socially costly for middle-aged speakers.

Beaman, K. V. (2022). (In)coherence across the linguistic architecture: Change in Swabian across the lifespan. *Methods in Dialectology (Methods XVII)*, Johannes Gutenberg-Universität Mainz.

Buchstaller, I. (2013). *Quotatives: New trends and sociolinguistic implications*. John Wiley & Sons.

Buchstaller, I. (2011). Quotations across the generations: A multivariate analysis of speech and thought introducers across 5 decades of Tyneside speech. *Corpus Linguistics and Linguistic Theory*, 7(1), 59–92.

Buchstaller, I. 2015. “Exploring linguistic malleability across the life span: Age-specific patterns in quotative use.” *Language in Society* 44(4):457–496.

Buchstaller, I., Krause, A., Auer, A., & Otte, S. (2017). Levelling across the life-span?: Tracing the face vowel in panel data from the North East of England. *Journal of Sociolinguistics*, 21(1), 3–33.

- Conn, J. C. (2005). Of “moice” and men: The evolution of a male-led sound change. University of Pennsylvania.
- Downes, W. (1984). *Language and Society* (2nd ed.). Cambridge University Press.
- Harrington, J., Palethorpe, S., & Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen’s Christmas broadcasts. *Journal of the International Phonetic Association*, 30(1/2), 63–78.
- Mechler, J., Grama, J., Bauernfeind, L., Eiswirth, M., & Buchstaller, I. (2022). Towards an Empirically-based Model of Age-graded Behaviour: Trac(ing) linguistic malleability across the entire adult life-span. *University of Pennsylvania Working Papers in Linguistics*, 28(2).
- Sankoff, G. (2005). Cross-Sectional and Longitudinal Studies. In U. Ammon, H. E. Wiegand, G. Ungeheuer, H. Steger, & K. Brinker (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft* (2., completely rev. and extended ed). de Gruyter.
- Sankoff, G., & Blondeau, H. (2007). Language Change across the Lifespan: /R/ in Montreal French. *Language*, 83(3), 560–588.
- Singler, J. (2004). Learning to be unhip, and other life lessons. Paper presented at Sociolinguistic Perspectives on Age. New York University, New York, NY.
- Tagliamonte, S. A., & D’Arcy, A. (2007). Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change*, 19(2), 199–217.
- Wagner, S. E. (2012a). Age Grading in Sociolinguistic Theory. *Language and Linguistics Compass*, 6(6), 371–382.
- Wagner, S. E. (2012b). Real-time evidence for age grad(ing) in late adolescence. *Language Variation and Change*, 24(2), 179–202.
- Wagner, S. E., & Tagliamonte, S. A. (2017). What Makes a Panel Study Work?: Researcher and Participant in Real Time. In *Panel Studies of Variation and Change*. Routledge.
- Wagner, S. E., & Sankoff, G. (2011). Age grading in the Montréal French inflected future. *Language Variation and Change*, 23(3), 275–313.

---

## Exploring variation in heritage Tamil retroflex perception and production.

Kamala Muthukumarasamy, Chandan Narayan  
York University, Toronto, Canada

This presentation investigates how sociolinguistic factors impact the psycholinguistics of heritage Tamil liquid retroflex perception and production. Heritage Tamil is an understudied language, with minimal literature on how speakers acquire, retain, and maintain their language. Since heritage language acquisition relies heavily on language input, familial usage, consistent exposure, individual attitudes, aligning beliefs, and personal motivation, this results in incredible variation among heritage language speakers, their proficiencies, and language abilities. In our presentation, we explore how social factors affect the perception and production of the alveolar-retroflex liquid contrast ([l]-[ɭ]) in heritage Tamil speakers in the greater Toronto area, a region with a high concentration of multi-generational Tamil speakers from India and Sri Lanka. Since heritage input differs greatly from majority or dominant language input, the relationship between these socio and psycholinguistic factors provides new insight into heritage phonological contrast retention and maintenance. In particular, we examine the connection between language and identity in the maintenance of this acoustically fragile contrast with similar spectral characteristics, shown through low perceptual salience. Heritage Tamil speakers (n=18) were recruited for participation in production and perception tasks, as well as a detailed language use questionnaire about language use and exposure, and follow-up sociolinguistic interviews about identity and motivation.

First, an AX discrimination task with non-word Tamil VCVs was administered. D-prime, a bias free measure of perceptual distance, was computed from the discrimination data. Then, participants provided minimal pairs with the target consonants in an elicited production task. An F3-F2 (Hz) score was taken as a measure of productive salience, with alveolars having a larger difference than retroflexes. Results showed a high degree of variation in productive salience, i.e., some speakers with clear differentiation between alveolar and retroflex, and others very little. Likewise, perceptual distance was also variable, with some participants clearly showing categorical discrimination while others have very perceptual space for both liquids. Interestingly, a simple language exposure metric did not predict productive or perceptual scores. Follow-up interviews, however, revealed that personal motivations, a concrete embedding of the heritage language within their individual culture, as well as a strong linguistic identity served as a strong indicator of accuracy in perception and production. Since diverse variation is so prevalent in heritage speaker populations and their language abilities, it is important to highlight both quantitative and qualitative approaches to data analysis to broaden our understanding of these speakers. Our research addresses whether acoustically fragile contrasts are realized in an understudied heritage language, and importantly, how, and why cultural identity and motivation serve to maintain them. Implications of this research include contributing to the literature on variation in understudied languages, the sociolinguistic and psycholinguistic patterning of heritage speaker populations, as well as providing baseline literature for further research on heritage Tamil and other heritage Dravidian languages.

Alarcon, I. (2010). Advanced Heritage Learners of Spanish: A Sociolinguistic Profile for Pedagogical Purposes. *Foreign Language Annals*, 43(2), 269–288. <https://doi.org/10.1111/j.19449720.2010.01078.x>

- Amengual, M. (2019). Type of early bilingualism and its effect on the acoustic realization of allophonic variants: Early sequential and simultaneous bilinguals. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior*, 23(5), 954–970. <https://doi.org/10.1177/1367006917741364>
- Brown, J. R., & Bousquette, J. (2018). Heritage Languages in North America: Sociolinguistic Approaches. *Journal of Language Contact*, 11(2), 201–207. <https://doi.org/10.1163/19552629-01102002>
- Canagarajah, S. (2019). Changing Orientations to Heritage Language: The Practice-Based Ideology of Sri Lankan Tamil Diaspora Families. *International Journal of the Sociology of Language*, 255, 9–44.
- Chang, C. (2016). Bilingual perceptual benefits of experience with a heritage language. *Bilingualism: Language and Cognition*, 19(4), 791–809. doi:10.1017/S1366728914000261
- Chang, C., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustical Society of America*, 129(6), 3964–3980. <https://doi.org/10.1121/1.3569736>
- Cognola, F., Baronchelli, I., & Molinari, E. (2019) Inter- vs. Intra-Speaker Variation in Mixed Heritage Syntax: A Statistical Analysis. *Frontiers in Psychology*, 10(1528). doi: 10.3389/fpsyg.2019.01528
- Comanaru, R. & Noels, K. A. (2009). Self-Determination, Motivation, and the Learning of Chinese as a Heritage Language. *Canadian Modern Language Review*, 66(1), 131–158. <https://doi.org/10.3138/cmlr.66.1.131>
- Flege, J. & Bohn, O. The Revised Speech Learning Model (SLM-r). (2021). In *Second Language Speech Learning* (pp. 3–83). <https://doi.org/10.1017/9781108886901.002>
- Gharibi, K. & Boers, F. (2017). Influential factors in incomplete acquisition and attrition of heritage speakers' vocabulary knowledge. *Language Acquisition*, 24(1), 52–69. <https://doi.org/10.1080/10489223.2016.1187613>
- Icardo Isasa, A. (2022). Preliminary Results on the Development of the Perception of Spanish /e/ and /ei/ by Heritage Learners vs. L2 Learners of Spanish in the Classroom. *Languages (Basel)*, 7(1), 7, 20–21. <https://doi.org/10.3390/languages7010007>
- Isakson, S. (2018). The Case for Heritage ASL Instruction for Hearing Heritage Signers. *Sign Language Studies*, 18(3), 385–411. <https://doi.org/10.1353/sls.2018.0008>
- Kan, R., & Schmid, M. (2019). Development of tonal discrimination in young heritage speakers of Cantonese, *Journal of Phonetics*, 73, 40–54. <https://doi.org/10.1016/j.wocn.2018.12.004>
- Kasstan, J. R., Auer, A., & Salmons, J. (2018). Heritage-language speakers: Theoretical and empirical challenges on sociolinguistic attitudes and prestige. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior*, 22(4), 387–394. <https://doi.org/10.1177/1367006918762151>
- Kim, JY (2020). Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress. *Bilingualism: Language and Cognition* 23, 233–250. <https://doi.org/10.1017/S1366728918001220>
- Kim, J. & Repiso Puigdellíura, G. (2020). Deconstructing Heritage Language Dominance: Effects of Proficiency, Use, and Input on Heritage Speakers' Production of the Spanish Alveolar Tap. *Phonetica*, 77(1), 55–80. <https://doi.org/10.1159/000501188>
- Lee-Ellis, S. (2012). Looking into bilingualism through the heritage speaker's mind. ProQuest Dissertations Publishing.
- Flege, J. & Bohn, O. The Revised Speech Learning Model (SLM-r). (2021). In *Second Language Speech Learning* (pp. 3–83). <https://doi.org/10.1017/9781108886901.002>
- Montrul, S. (2016). *The Acquisition of Heritage Languages*. Cambridge University Press.
- Narayan, C. (2019). An acoustic perspective on 45 years of infant speech perception, Part 1: Consonants. *Language and Linguistics Compass*, 13(10). <https://doi.org/10.1111/lnc3.12352>
- Narayanan, S., Kaun, A., Byra, D., Ladefoged, P., & Alwan, A. (1996). Liquids in Tamil. *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*, 2, 797–800 vol.2. <https://doi.org/10.1109/ICSLP.1996.607721>
- Narayanan, S., Byrd, D., & Kaun, A. (1999). Geometry, Kinematics and Acoustics of Tamil Liquid Consonants. *The Journal of the Acoustical Society of America*, 106(4), 1993–2007. <https://doi.org/10.1121/1.427946>
- Park, S. M. & Sarkar, M. (2007). Parents' Attitudes Toward Heritage Language Maintenance for Their Children and Their Efforts to Help Their Children Maintain the Heritage Language: A Case Study of Korean-Canadian Immigrants. *Language, Culture and Curriculum*, 20(3), 223–235. <https://doi.org/10.2167/lcc337.0>
- Scontras, G. & Putnam, M. T. (2020). Lesser-studied Heritage Languages: An Appeal to the Dyad. *Heritage Language Journal*, 17(2), 152–155. <https://doi.org/10.46538/hlj.17.2.2>
- Seo, Y., Dmitrieva, O., & Cuza, A. (2022). Crosslinguistic Influence in the Discrimination of Korean Stop Contrast by Heritage Speakers and Second Language Learners. *Languages*, 7(1):6. <https://doi.org/10.3390/languages7010006>
- Shea. (2019). Dominance, Proficiency, and Spanish Heritage Speakers' Production of English and Spanish Vowels. *Studies in Second Language Acquisition*, 41(1), 123–149. <https://doi.org/10.1017/S0272263117000328>
- Vijayakumar, K., Gunalan, S., & Rajeshwaran, R. (2021). Development of Minimal Pair Test in Tamil (MPT-T). *Journal of Clinical and Diagnostic Research*. 15(9), 1–4. 10.7860/JCDR/2021/46807.15357
- Wen, X. (2011). Chinese Language Learning Motivation: A Comparative Study of Heritage and Non-heritage Learners. *Heritage Language Journal*, 8(3), 333–358. <https://doi.org/10.46538/hlj.8.3.3>

---

## “My safe word will be [ʍ]iskey!” – An acoustic approach to the whine-wine split

Miriam Neuhausen

University of Freiburg, Freiburg, Germany

The voiceless fricative /ʍ/ is realized in words spelled with <wh->, such as whiskey and whine, while words spelled with <w->, such as wine, are produced with the voiced approximant /w/. For the past three centuries, /w/ has been replacing /ʍ/ in most varieties of English (Chambers 1998: 26). The few varieties of English outside Scotland that still feature /ʍ/ are places with geographically or historically close ties to Scotland, such as Ireland and Northumberland (Chambers 2013: 357), southern Ontario (Boberg 2010: 82), southeastern U.S. (Milroy 2004: 51; Chambers 2013: 357), and southern New Zealand (Schreier et al. 2003: 264). By now, the whine-wine split has attracted the attention of non-linguists and is sometimes even used for comical effects, such as in the movie *Hot Rod* (2007).

The vast majority of studies on whine-wine has so far been carried out based on auditory coding (Schreier et al. 2003; Schützler 2010; Brato 2014). Gut and Li (2022) are the first to suggest an innovative acoustic measurement: the median of harmonicity, which expresses the acoustic relation between voicing and friction of a sound. This talk raises the question whether the median of harmonicity should be normalized and proposes a normalization method.

Even though in Ontario at large the language change to /w/ has been completed since the 1960s (Chambers 1998: 26), /ʍ/ has been maintained in the speech of some Mennonites in southern Ontario, where it now serves as a community-specific marker in Labov’s (1972: 178) sense. Drawing on 38 sociolinguistic interviews with 51 speakers which I conducted in 2018-2019, I base my findings on four Mennonite communities in Ontario ranging from conservative to liberal.

My database includes 18,489 <w-> and 3,716 <wh-> words. I confirm the median of harmonicity as a measure for /ʍ/ in the current data set; the median of harmonicity significantly correlates with my auditorily hand-coded <wh-> words in the speech of 27 speakers (n = 1,610). In a following step, I address the fact that the median of harmonicity also captures vocal fry, which occurs particularly in the speech of elderly women (Ferrand 2002: 484; Lortie et al. 2015: 11). I discuss what gendered age-related vocal fry means for normalization and introduce a normalization method developed by Beke Hansen and me. The normalization is based on the individual speaker mean of [w]; the further away the <wh-> variant from the respective speaker mean for <w-> words, the closer it is to [ʍ]. In a final step, I compare two mixed-effects models with the raw and the normalized data.

Acoustic approaches to the whine-wine split are time- and energy-saving for researchers and provide objective measurements on a continuum (rather than subjective thresholds for cut-off points). Furthermore, they allow for the observation of stronger/weaker realizations within each category and shed light on potential “cross-over” effects (e.g. Labov 1966: 152; Trudgill 1997: 98), i.e., some groups produce stronger realisations than those who are typically associated with the conservative variant.

- Boberg, Charles. 2010. *The English Language in Canada: Status, History and Comparative Analysis* (Studies in English Language). Cambridge: Cambridge University.
- Brato, Thorsten. 2014. Accent Variation and Change in North-East Scotland: The Case of (HW) in Aberdeen. In Robert Lawson (ed.), *Sociolinguistics in Scotland*, 32–52. London: Palgrave Macmillan. <https://doi.org/10.1057/9781137034717>.
- Chambers, Jack K. 1998. Social Embedding of Changes in Progress. *Journal of English Linguistics* 26(1). 5–36. <https://doi.org/10.1177/007542429802600102>.
- Chambers, Jack K. 2013. Patterns of Variation including Change. In Jack K. Chambers & Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, 297–323. 3rd edn. Malden, MA: Blackwell.
- Ferrand, Carole T. 2002. Harmonics-to-Noise Ratio: An Index of Vocal Aging. *Journal of Voice* 16(4). 480–487.
- Hot Rod. 2007. Akiva Schaffer (dir.), Paramount Pictures. <https://www.youtube.com/watch?v=gbu88CKkEzI>.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center of Applied Linguistics.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.
- Li, Zeyu & Ulrike Gut. 2022. The Distribution of /w/ and /ʍ/ in Scottish Standard English. *Corpus Linguistics and Linguistic Theory*. <https://doi.org/10.1515/cilt-2021-0052>.
- Lortie, Catherine L., Mélanie Thibeault, Matthieu J. Guitton & Pascale Tremblay. 2015. Effects of Age on the Amplitude, Frequency and Perceived Quality of Voice. *AGE* 37(6). 117. <https://doi.org/10.1007/s11357-015-9854-1>.
- Milroy, Lesley. 2004. An Essay in Historical Sociolinguistics? On Donka Minkova’s “Philology, Linguistics, and the History of [hw]~[w].” In Anne Curzan & Kimberly Emmons (eds.), *Studies in the History of the English Language II: Unfolding Conversations*, 47–53. De Gruyter Mouton. <https://doi.org/10.1515/9783110897661>.
- Schreier, Daniel, Elizabeth Gordon, Jennifer Hay & Margaret Maclagan. 2003. The Regional and Sociolinguistic Dimension of /hw/ Maintenance and Loss in Early 20th Century New Zealand English. *English World-Wide* 24(2). 245–269. <https://doi.org/10.1075/eww.24.2.06sch>.
- Schützler, Ole. 2010. Variable Scottish English Consonants: The Cases of /ʍ/ and Non-Prevocalic /r/. *Research in Language* 8. 5–21. <https://doi.org/10.2478/v10015-010-0010-9>.
- Trudgill, Peter. 1997. The Social Differentiation of English in Norwich. In Nikolas Coupland & Adam Jaworski (eds.), *Sociolinguistics: A Reader*, 179–184. London: Macmillan Education UK. [https://doi.org/10.1007/978-1-349-25582-5\\_15](https://doi.org/10.1007/978-1-349-25582-5_15).



---

One Book Watches the Other:  
Using Mules and Men to Determine Distinctive Features of Barracoon

Leah Nodar  
Purdue University, West Lafayette, USA

In 1860, the *Clotilda* arrived in Mobile Bay, the last known slave ship to land in America. After the Civil War and emancipation, the Africans from the ship founded Africatown, a semi-isolated community north of Mobile, Alabama (Diouf 2007). The group was known locally for its unique variety of English (Diouf 2007, Allen Jones 1979), of which the longest and most detailed portrayal occurs in Zora Neale Hurston's *Barracoon* (2018). Hurston spent several months of 1927-8 interviewing Kossula, aka Cudjo Lewis, but her manuscript was not published, in part because Hurston wrote the dialogue in her hallmark style of dense orthographic indication of dialect features (Hurston 2018:xv, xxiii-xxiv).

*Barracoon*'s recent publication after so many decades gives us a unique window onto early Africatown speech. Here, I focus on the methods used to analyze possible spoken features when the dialect evidence is largely limited to a single historical semi-literary text. Two questions are considered: first, is Hurston generally reliable in her portrayal of vernacular dialogue? And if so, what does she indicate in *Barracoon* to be distinctive linguistic features of Kossula's speech?

To understand the language of *Barracoon*, I use another Hurston work, *Mules and Men* (2008 [1935]). This book is also non-fiction, anthropological, interview-based, and dialogue-heavy. However, the speakers here are Black Southerners from Eatonville, FL; Polk County, FL; and New Orleans, LA. Thus these dialect portrayals are from the most comprehensively studied dialects in sociolinguistics, African American English (Wolfram and Schilling 2016:217).

To check Hurston's accuracy, I first pull three 200-clause random samples of *Mules and Men* dialogue, one sample for each location. This dialogue text is compared to narration text, as Hurston's version of standard English, and I extract orthographic patterns that differ between them. These patterns are cross-checked against linguistic feature compendiums (e.g. Bailey and Thomas 1998, Kautzsch 2008, Wolfram and Schilling 2016) and there matched with independently attested features of Black and/or Southern and/or widespread non-standard English.

This survey finds 22 phonetic patterns and 31 morphosyntactic patterns, of which fifty match to independently attested vernacular linguistic features; two are eye dialect; and only one is not found. I conclude that Hurston's presentation of Black Southern English outside of Africatown overwhelmingly agrees with independent linguistic evidence about these dialects, and therefore choose to rely also on her portrayal of Africatown English.

Second, I use the *Mules and Men* orthographic patterns as a baseline for Hurston's portrayal of Southern Black English, and compare a similar 200-clause sample of *Barracoon*, looking for portrayals that differ between the two texts. Several patterns stand out, with some occurring only in *Barracoon* (verbs with enclitic vowels), some in *Mules and Men* but not *Barracoon* ("ah" for "i"/"y," suggesting [ai]-monophthongization), and some in both but under different conditions, such as copula absence, which occurs in past tense and first person only in *Barracoon*. The unique dialect and history of Africatown is worth study in itself, and it also adds to our understanding of the development of early AAE.

- Allen Jones, Clara Eva. (1979, June 6). 127. Clara Eva Allen Jones [interview]. Oral History Project; Mobile Public Library Local History and Genealogy.
- Bailey, Guy and Erik Thomas. 1998. Some aspects of African-American Vernacular English phonology. In Mufwene, Salikoko S., John R. Rickford, Guy Bailey, and John Baugh, eds. *African-American English: Structure, history and use*. New York, NY: Routledge.
- Diouf, Sylviane A. (2007). *Dreams of Africa in Alabama: The slave ship Clotilda and the story of the last Africans brought to America*. New York, NY: Oxford University Press.
- Hurston, Zora Neale. (2008 [1935]). *Mules and men*. New York: HarperCollins Publishers, Inc.
- Hurston, Zora Neale. (2018). *Barracoon: the story of the last "black cargo."* New York, NY: HarperCollins Publishers.
- Kautzsch, Alexander. 2008. Earlier African American English: morphology and syntax. In Edgar W. Schneider, ed. *Varieties of English 2: The Americas and the Caribbean*. Berlin: Mouton de Gruyter.
- Roche, Emma. (1914). *Historic sketches of the south*. New York: Knickerbocker Press.
- Wolfram, Walt, and Natalie Schilling. (2016). *American English*. 3rd edition. West Sussex, UK: John Wiley and Sons, Inc.

---

In search of "literary" Taiwanese: Negation in a Taiwanese hand puppet theater performance

Caroline Norfleet  
The Ohio State University, Columbus, USA

Traditional Taiwanese hand puppet theater is generally described as being performed in a literary register of Taiwanese Southern Min (henceforth Taiwanese) (Wu 2003: 101). There has been, however, little description of

what “literary Taiwanese” is. Rather, when applied to Taiwanese, the term “literary” appears in the context of literary and colloquial character readings: etymologically related syllables which entered Taiwanese as part of different linguistic strata. Colloquial readings entered the Southern Min speaking regions of China with the earliest Han settlers in the Qin, Han and Jin dynasties, while literary readings evolved from the speech of educated scholars during the later Tang dynasty (Klötter 2005: 18). This distinction is not sufficient to describe the language of hand puppet theater, as many syllables in the language do not make a literary-colloquial distinction, and it additionally does not address factors such as syntax and word choice. This project investigates the use of negation in the performance *Liông Bùn Tîn* (“Dragon Gate Formation”) in order to better understand language use in Taiwanese hand puppet theater and provide a partial description of this aspect of literary Taiwanese.

The negative morphemes used in *Liông Bùn Tîn* can be divided into literary and colloquial categories, as shown in Table 1. The colloquial negatives, making up approximately 25% of negation in the performance, are identical to those used in everyday Taiwanese speech. The literary negatives, making up the remaining 75%, more closely resemble negation as used in Literary Chinese (*wenyan*, the written variety used by speakers of various Sinitic and non-Sinitic languages in the Chinese cultural sphere until the early 20th century) (Denecke & Nguyen 2017: 511). Each literary negative corresponds to a colloquial negative: in two cases, *bô-bû* and *buē-bī*, they represent literary and colloquial readings as described above, while the other two pairs would not be considered such. Items in each pair mainly have the same function, but certain specific functions are performed only by literary or colloquial morphemes: for instance, only colloquial negatives serve as verb complements.

The use of literary or colloquial negation is determined not only by these grammatical restrictions but also by context. When the distinction between a literary and colloquial form of a verb can be made, literary negatives tend to co-occur with literary verbs and vice versa. Additionally, high-status characters use literary negation more frequently than lower-status characters, and literary negation occurs more frequently in monologues than in dialogues. It is concluded that literary language is not the sole medium of puppet theater, but rather that the choice between literary and colloquial language is made with the intention to convey information about the plot and characters. It is clear also that, while syllables representing literary character readings have been incorporated into spoken Taiwanese without any literary connotation, in this context associations of education and status still remain.

Denecke, W. & Nguyen, N. (2017). Shared literary heritage in the East Asian Sinographic Sphere. In W. Denecke, W. Li & X. Tian (Eds.), *The Oxford handbook of classical Chinese literature* (pp. 510-532). Oxford University Press.

Klötter, H. (2005). *Written Taiwanese*. Harrassowitz Verlag.

Ministry of Education, R.O.C. (Director). (1997). *Long Men Zhen 龍門陣 [Dragon Gate Formation]* [Film]. Caituan Faren Li Tianlu Budaixi Wenjiao Jijinhui Hui 財團法人李天祿布袋戲文教基金會.

Wu, S. (2003). Hand puppet theater performance: Emergent structures and the resurgence of Taiwanese identity. In P. R. Katz & M. A. Rubenstein (Eds.), *Religion and the formation of Taiwanese identities* (pp. 99-121). Palgrave Macmillan.

## Linguistic innovation or ancestral feature? The case of *tsuzamen* and *tsam* in Hasidic Yiddish

Chaya Nove

University of California Berkeley, Berkeley, USA

In Yiddish, the morpheme *tsuzamen* ('together') functions syntactically both as an adverb and as a separable prefix that, when attached to certain verbs, forms words with new meanings. Some of these formations are semantically transparent (e.g., *tsuzaman* + *nemen* ['take'] 'collect'), while others are more opaque (e.g., *tsuzamen* + *shtoyesn* ['kick'], 'collide') (Jacobs 2005). In the conversational Yiddish of contemporary Hasidim in New York, there are two distinctive variants for these two syntactic functions: While *tsuzamen* is used in the adverbial capacity, the prefixed form is invariably *tsam*. This syntactic duality can be seen in data drawn from sociolinguistic interviews with 44 Hasidic Yiddish speakers in New York. Moreover, in Hasidic Yiddish, *tsam* can also be productively affixed to verbs to add the meaning 'excessive' (e.g., *tsam* + *lakhn* ['laugh'] 'laugh excessively', *tsam* + *shlogn* ['beat'] 'beat excessively'), but *tsuzamen* cannot have this meaning. Since neither the form *tsam* nor its sense of 'excessive' is described in the Yiddish literature, the question arises whether these are recent innovations or merely overlooked features of Central Yiddish, the dialect spoken in the Transcarpathian region of Eastern Europe, the ancestral homeland of present-day Hasidim. The recently created Corpus of Spoken Yiddish in Europe (CSYE) (Bleaman and Nove, submitted), based on Holocaust testimonies recorded by the Shoah Foundation, is enabling the discovery and analysis of previously undescribed variability in prewar Yiddish dialects. For the present study, nearly 400 tokens of *tsuzamen* and its variants were extracted from testimonies by 91 speakers sampled in the CSYE. In these data, 24% of the variants uttered by speakers of Central Yiddish surface as *tsam*. The data also reveal several additional, less frequently occurring forms of *tsuzamen* (including *tsuzam*, which may be an intermediate form). When the Central Yiddish tokens of *tsuzamen* are analyzed by region (Poland vs. Transcarpathia), *tsam* emerges as the dominant form in Transcarpathia (62% of 233 tokens), while in Poland it does not appear at all. Moreover, there is evidence of the use of *tsam* as 'excessive' in prewar Transcarpathia. This study, which traces the emergence of a novel split in the

syntactic function of tsuzamen-tsam in Hasidic Yiddish, illustrates how data from the CSYE can provide a crucial link between pre- and postwar Yiddish dialects, enabling the comparison of contemporary Yiddish to homeland varieties that no longer exist. This is especially important for the study of Hasidic Yiddish, whose ancestral dialect is severely understudied due to scholarly bias. This study also models how the CSYE can be used to discover patterns of variation in prewar Yiddish dialects that can broaden our understanding of dialect formation.

Beaman, Isaac L., and Chaya R. Nove. submitted. "The Corpus of Spoken Yiddish in Europe: Goals, Methods, and Applications." *Language Documentation & Conservation*.

Jacobs, Neil G. 2005. *Yiddish: A Linguistic Introduction*. Cambridge: Cambridge University Press.

---

## D1 feature maintenance in a D2 region: Older New Yorkers in Toronto

Jennifer Nycz<sup>1</sup>, Matthew Dearstyne<sup>2</sup>

<sup>1</sup>Georgetown University, Washington, DC, USA. <sup>2</sup>Institute for Health Metrics and Evaluation, Washington, DC, USA

Second dialect acquisition (SDA) studies often quantify "success" in acquiring new dialect (D2) variants, comparing mobile speaker use rates to those of nonmobile speakers. However, use is an imperfect proxy for acquisition: speakers might avoid D2 features for many reasons (stigmatization, for example). Continued use of D1 features, meanwhile, may allow them to convey important aspects of identity in their new region. Studies of how mobile speakers use D1 features are rare (Bortoni-Ricardo 1985; Bowie 2000), but more such work is needed to understand the dynamics of accent change (or lack thereof) due to mobility. Accordingly, our paper examines how native New Yorkers in Toronto use two New York City English (NYCE) features, (r) and (oh), and whether use of NYCE variants can be linked to expressions of New Yorker identity.

Data come from 15 participants interviewed as part of a larger study of New Yorkers in Toronto; all were aged 51+, as older speakers are most likely to have acquired the classic NYCE variants prior to relocation (Becker 2016). Interviews were recorded to 44.1 kHz/16-bit wav files using an Audio-Technica lavalier mic and a Zoom H4N Pro, transcribed (ELAN 2018), then submitted to FAVE-Align scripts (Rosenfelder et al, 2011) to generate segmented Praat textgrid files (Boersma & Weenink, 2021). Formants were extracted from stressed vowels >50 msec, then Lobanov-normalized. Coda r tokens were coded as realized or deleted. Annotation units were coded for place-based topic and affect: Locality was coded primarily as Toronto or NYC. Affect was coded as neutral by default, and as positive or negative in the presence of overt affect markers. Linguistic variables and content were coded independently by two RAs, with a third coder serving as a tie-breaking vote when needed. Mixed-effects models were used to assess the impact of linguistic and stylistic factors on variant use and to compare speakers to one another via random effects.

We find considerably higher rates of rhoticity (>80%) than those found for comparable non-mobile speakers (Becker 2014). (oh) height varies more, with eight speakers producing (oh) at a height above a benchmark indicating raised (oh). For both variables, the ordering of linguistic factors conditioning NYCE-variant appearance mirror those found for non-mobile New Yorkers. To examine the stylistic function of these variables, we focus on TH, a man who has lived in Toronto for 47 years after moving from Queens at age 15, and who has the lowest r-index and highest (oh). TH exhibits an overall r-index of just over 50, but is most rhotic (r-77) when talking negatively about NYC topics. Meanwhile, TH uses his highest (oh) in positive talk and lowest (oh) in negative talk about NYC-topics. These patterns suggest TH's partial maintenance of D1 forms is not due to inability to acquire D2-like patterns, so much as the continued usefulness of these forms in conveying aspects of his place-linked identity. We discuss what these findings imply for studies of SDA and how "acquisition" may be better quantified and compared in mobile speakers.

Becker, Kara. 2014. (r) we there yet? The change to rhoticity in New York City English. *Language Variation and Change* 26, 141–168. doi:10.1017/S0954394514000064

Becker, Kara. 2016. Linking community coherence, individual coherence, and bricolage: The co-occurrence of (r), raised BOUGHT and raised BAD in New York City English. *Lingua*. Guy & Hinskens 2016.

Bortoni-Ricardo, Stella Maris. 1985. *The urbanization of rural dialect speakers: a sociolinguistic study in Brazil*. Cambridge, UK: Cambridge University Press.

Bowie, David. 2000. *The effect of geographic mobility on the retention of a local dialect*. Philadelphia, PA: University of Pennsylvania dissertation.

ELAN (Version 6.1) [Computer software]. 2021. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from <https://archive.mpi.nl/tla/elan>

Rosenfelder, Ingrid; Fruehwald, Josef; Brickhouse, Christian; Evanini, Keelan; Seyfarth, Scott; Gorman, Kyle; Prichard, Hilary; Yuan, Jiahong; 2022. FAVE (Forced Alignment and Vowel Extraction) Program Suite v2.0.0 /zenodo.

Boersma, Paul & Weenink, David (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.42, retrieved 15 April 2021 from <http://www.praat.org/>

---

A San Diegan Debate: Take 8 or Take the 8,  
Investigating the use of the determiner “the” before numeric freeways in San Diego  
Brianna O’Boyle  
San Diego State University, San Diego, USA

California has gained attention as a notable dialect region, with recent research on the Californian Vowel Shift exemplified by Podesva et al. (2015) and D’Onofrio et al. (2019). However there are linguistic features within California that remain relatively unexplored. This paper aims to investigate one of those features: usage of the determiner the before numeric freeways (e.g., referring to Interstate 5 as the 5). Geyer (2001) mentions this is a feature of Southern California English and Bucholtz et al. (2007) find people are aware of it as a regional feature. The is salient enough to have reached mainstream media. In Saturday Night Live’s recurring skit “The Californians,” which parodies and highlights noticeable features of California English (Pratt & D’Onofrio 2017) actors frequently mention freeway names with the. To the best of my knowledge this is the first study using formal sociolinguistic methodology to examine the use of this variable in Southern California, specifically within San Diego. This paper also seeks to examine second-dialect acquisition through usage of the among transplants to San Diego. According to Chambers (1992) acquisition of lexical features occur faster than phonological elements, and they are acquired more quickly in the first stages of dialect acquisition—this may allow us to see the beginnings of dialect acquisition among recent transplants.

Ten speakers were interviewed: four San Diegan natives, four long-term transplants to San Diego (20+ years), and two recent transplants to San Diego (<1 year): despite a small sample size the results were remarkably clear. It was hypothesized Native San Diegans would have high usage of the, long-term transplants would have some usage of the but primarily for roads located in San Diego, and recent transplants would have low usage of the. Data was collected through sociolinguistic interviews with modules designed around topics of transportation to elicit more tokens of the variable. Additionally, a map task was designed where participants were shown a map of San Diego, and a map of an unfamiliar region, and asked to navigate from various locations using freeways. The data was then analyzed with mixed-effect regressions using speaker as a random factor. Three factors were found to be statistically significant: San Diegan native vs transplant status, age, and map presence.

The age effect suggests change in apparent time toward increased use of the. A couple of anecdotes from interviewees support the hypothesis that the is an innovation by describing it as something that diffused from Los Angeles to San Diego in the 1980s. Participants also favored the most with the presence of the map of San Diego, and to a lesser degree with the map of the unknown region. Recent transplants only used the with roads in San Diego meanwhile their long-term counterparts used the for roads in San Diego and unfamiliar regions. Overall San Diegan natives favored the, long-term transplants also favored the though to a lesser degree, and recent transplants disfavored the. Usage of the has become, in part, engrained into the dialectal grammar of the long-term transplants.

- Bucholtz, M., Bermudez, N., Fung, V., Edwards, L., & Vargas, R. (2007). Hella Nor Cal or Totally So Cal? *Journal of English Linguistics*, 35(4), 325–352. <https://doi.org/10.1177/0075424207307780>
- Chambers, J. K. (1992). Dialect acquisition. *Language*, 68(4), 673. <https://doi.org/10.2307/416850>
- D’Onofrio, Annette, Teresa Pratt & Janneke Van Hofwegen (2019). Compression in the California Vowel Shift: Tracking generational sound change in California’s Central Valley. *Language Variation and Change* 31.2:193–217.
- Geyer, G. (2001). “The” Freeway in Southern California. *American Speech*, 76(2), 221–224. <https://doi.org/10.1215/00031283-76-2-221>
- Podesva, Robert, Annette D’Onofrio, Janneke Van Hofwegen, & Seung Kyung Kim (2015). Country ideology and the California Vowel Shift. *Language Variation and Change* 27.2:157–186.
- Pratt, Teresa & Annette D’Onofrio (2017). Jaw setting and the California Vowel Shift in parodic performance. *Language in Society* 46.3:283–312.

---

It’s GIFing gay: Reaction images and GIFs on gay Twitter  
Caolan O’Neill  
Pembroke College, University of Oxford, Oxford, United Kingdom

In digital spaces, reaction GIFs and images function as multimodal linguistic objects (Kuo, 2019); however, aside from their capacity to digitally embody gestures or facial expressions, their use is critically understudied. To address this gap, this research probes how reaction GIFs and images are marshalled to performatively enact gender- and sexuality-based identities among a community of practice (Eckert & McConnell-Ginet, 1992) of eight (mostly British) self-identified cisgender gay men. Through both quantitative and qualitative analyses, this research shows that reaction GIFs and images are associated with an indexical field involving a web of stances, qualities, and characteristics (Eckert, 2008), that enable the performance of a multifaceted “gayness” online.

Data derives from a carefully constructed, manually collected corpus of Twitter exchanges (N=837) between the eight men that include reaction GIFs (N=675) or images (N=162), spanning an almost ten-year period. These data were analysed via a computer-mediated discourse analysis (Herring, 2004), which drew on Du Bois' (2007) stance triangle. Across the interactional data, these men repeatedly use GIFs/images with affective or epistemic stances depicting flamboyant or feminine subjects, enacting particular versions of gayness via visual depiction of overt femininity or hetero-/cis-non-conformity. Four (feminine) embodied archetypes/personae emerge: the Sassy Queen, the Hun, the Battle-Axe, and the Flamboyant Queer, each with specific affective or linguistic collocations. Examples of the latter include lexical and morphosyntactic features from African American Language (e.g., the lexical item *yas*, copula absence in where you at?, etc.) in performances of the Sassy Queen, orthographic representations of regional British English accents (e.g., TH-fronting orthographically represented in the respelling of McKeith as *as <MCKEEEEF>*) in constructions of the Hun persona, and so on. The variable deployments of these embodied personae reflect core vs peripheral status within the group and reveal the qualities and sensibilities that these men highly value or identify with, including campness (Sontag, 1964), outspokenness and resoluteness. Their use also (re)produces gendered, racial, class-based, etc. stereotypes, raising questions about the potential harmful or appropriative nature of these performances.

While the participants most often use reaction images and GIFs which (visually/linguistically) index femininity, this analysis reveals that they frequently do so to mock or poke fun at each other, which is stereotypically more characteristic of homosocial male interactions (cf. Coates, 1996). It is thus shown that these men reinforce homosocial bonds and do being masculine in their deployments of the various feminine/queer embodied personae, while simultaneously being subversive with respect to masculinity by doing it through feminine associated language/resources.

This research ultimately demonstrates how contemporary individuals' performances of gender and sexuality manifest across asynchronous, computer-mediated interactions, wherein cyberbodies, free from offline limitations, can more fully embody archetypal personae using digital objects (and the linguistic variation which is embedded in/co-occurs with these objects) to do being gay.

- Coates, J. (1996). *Women talk: Conversation between women friends*. Blackwell.
- Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 139–182). John Benjamins. <https://doi.org/10.1075/pbns.164.07du>
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453–476. <https://doi.org/10.1111/j.1467-9841.2008.00374.x>
- Eckert, P., & McConnell-Ginet, S. (1992). Think practically and look locally: Language and gender as community-based practice. *Annual Review of Anthropology*, 21, 461–490. <https://doi.org/10.1146/annurev.an.21.100192.002333>
- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behaviour. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338–376). Cambridge University Press. <https://doi.org/10.1017/CBO9780511805080.016>
- Kuo, R. (2019). Animating feminist anger: Economies of race and gender in reaction GIFs. In D. Ging & E. Siapera (Eds.), *Gender hate online: Understanding the new anti-feminism* (pp. 173–193). Palgrave Macmillan. [https://doi.org/10.1007/978-3-319-96226-9\\_9](https://doi.org/10.1007/978-3-319-96226-9_9)
- Sontag, S. (1964). Notes on 'camp'. *Partisan Review*, 31, 515–30.

---

## Finding order (or not): Hierarchical Density Based Clustering of New Yorkers' vowels.

Kovid Pal Odouard      Michael Newman<sup>3</sup>

New York University, New York, USA. Queens College/CUNY, Flushing, NY, USA

Cluster analyses reveal inherent structures in data through unsupervised means and are used for exploratory analyses and to mitigate potential unconscious researcher biases across many fields. Clustering of community members by their linguistic features should, in principle, reveal the sociolinguistic organization of a community prior to or instead of feature correlations with social categories. However, albeit with some exceptions, this tool has been infrequently used by sociolinguists.

A contrast between recent studies in New York and London reveals promises and challenges. Cole and Strycharczuk's (2022) k-means clustering of suburban Londoners' vowels subject to diphthong shift divides participants along a more Cockney vs. more standard binary. Haddican, et al.'s (2022) DIANA (i.e., divisive) clustering of five evolving vowels in New York however reveals seven groups. The differences may lie in the far more ethnically diverse NYC sample or that the vowels involved in (reversing) London diphthong shift may act more coherently than the ones changing in NYC.

This project launch explores how clustering vowel systems can be pursued in a diverse community. The results are expected to relate to the level of generality in which orderly heterogeneity may be found. It consequently provides another window of the question focus versus diffuseness of sociolinguistic norms (Le Page & Taboret-

Keller 1985). It also can provide a better understanding of how similar linguistic behaviors relate (or not) to the social structure of a community.

Potential data input include the mean vowel realizations of 14 vowel classes from 217 New Yorkers from the CUNY Corpus of New York City English (CONYCE). We employ two refinements on prior efforts:

1. Use of bark scaling to reduce the distorting effects of greater absolute differences relative to perceptual consequences as we move up and back in vowel space.
2. Use of Hierarchical Density Based clustering-HDBSCAN (Campello, et al. 2020), which unlike prior methods a) identifies outliers, b) is tolerant of higher dimensions and so multiple vowel classes, c) does not assume gaussian regularity; i.e., does not assume that clusters will be spherical and d) reveals smaller clusters embedded in larger clusters but does not force data into clusters.

Initial analyses with minimum clusters (as required by HDBSCAN) set at 5 individuals show the following.

1. Participants' full vowel systems do not seem amenable to clustering. The algorithm provides a single large cluster with a small number of outliers, no neat divisions when applied to all 14 vowels.
2. We do get clusters individual vowel classes (e.g., price, thought, prize, particularly robust when only F1 or F2 are entered — see Figure 1).
3. Smaller subsystems (e.g., kit and face backing and lowering — See Figure 2) confirming findings in Haddican, et al. (2022) but with a substantial number of outliers.

Next steps include:

- \* Testing with additional vowels and combinations.
- \* Identifying social characteristics of the clusters
- \* Using demographically stratified input

We expect that responses from sociolinguists can help refine the sociophonetic and statistical dimensions of the project and perhaps extend it to other types of communities.

Campello, Ricardo, Peer Kröger, Jörg Sander, Arthur Zimek. 2020. Density-based clustering. *WIREs Data Mining Knowl Discov.* 2020;10:e1343. <https://doi.org/10.1002/widm.1343>

Cole, Amanda and Strycharczuk, Patrycja. 2022. Dialect levelling and Cockney diphthong shift reversal in South East England: The case of the Debden Estate. *English Language & Linguistics*, 26(4), 621-643. doi:10.1017/S1360674321000290

Haddican, Bill, Cecilia Cutler, Michael Newman & Christina Tortora. 2022. Cross-Speaker Covariation across Six Vocalic Changes in New York City English. *American Speech*, 97 (4): 512–542.

Le Page, Robert & Tabouret-Keller. 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*.

---

## The social embedding of lectal coherence among mobile speakers

Livia Oushiro

UNICAMP, Campinas, Brazil

Lectal coherence has long been a topic of interest in sociolinguistics (e.g., Labov 2006 [1966]), and recent research has delved into more comprehensive investigations (e.g., Beaman & Guy 2022; Erker 2017; Guy & Hinskens 2016; Tamminga 2019). This study examines the social embedding of lectal coherence among 35 Brazilian Portuguese mobile speakers from rural Northeast to an urban area in the Southeastern state of São Paulo, to discuss whether internal migrants in a dialect contact situation simultaneously adopt typical host community variants.

Four sociolinguistic variables were analyzed in the corpus, balanced for speakers' gender, age of arrival and length of residence in São Paulo: (i) /r/-fronting—the realization of coda /r/ (porta 'door') as aspirated or tap/retroflex (N=2,847); (ii) /s/-fronting—the realization of coda /s/ (festa 'party') as post-alveolar or alveolar (N=1,787); (iii) /r/ non-deletion—deletion or realization of non-morphemic /r/ (amor 'love,' excluding the infinitival morpheme as in amar 'to-love') (N=3,372); and (iv) /s/ non-deletion—deletion or realization of non-morphemic coda /s/ (mas 'but,' excluding the plural morpheme as in gatos 'cats') (N=3,822). The second variants are more common in the host community, with the first two variables reflecting North-South identities and the latter two indicating rural-urban differences. Data were analyzed in Spearman correlation tests in R, utilizing usage rates and probabilities in logodds of Southern/urban variants from logistic regression analyses. The models included Speaker and internal variables as fixed effects, lexical item as random effect, and excluded social variables.

The consistent findings reveal significant positive correlations among three out of the six possible pairings: /r/- and /s/-fronting ( $\rho=.578$ ,  $p<0.001$ ); /r/ and /s/ non-deletion ( $\rho=.545$ ,  $p<0.001$ ); and /s/-fronting and /r/ non-deletion ( $\rho=.431$ ,  $p<0.01$ ) (Figure 1). Further analyses based on social groups show that, regarding Gender, women are more coherent than men, as women exhibit covariation for these three pairs and between /s/-fronting and /s/ non-deletion; in contrast, men only demonstrate coherence in the use of /r/- and /s/-fronting (Figure 1a). Regarding Age of Arrival, speakers who arrived earlier (before 19 y.o.) display higher coherence than those who arrived later (after 20 y.o.), as the former exhibit significant covariation for four pairings (the above three pairs and

between /r/-fronting and non-deletion) (Figure 1b). Lastly, regarding Length of Residence, speakers who have lived longer in the new community (10+ years) are more coherent than those with a shorter length of residence (-9 years), the former exhibiting covariation for three pairs and the latter for none (Figure 1c).

The overall findings demonstrate that migrants exhibit a coherent pattern of adopting or rejecting Paulista (fronted /r/ and /s/) and urban traits (/r/ and /s/ non-deletion), and suggest that similar phonological processes (fronting and non-deletion) promote lectal coherence. The results also indicate that dialect contact initially disrupts lectal coherence, as migrants don't adopt new variants simultaneously. Lectal coherence is restored over time and is contingent on speakers' gender and age of arrival. This study thus enhances our comprehension of how coherence is preserved, disrupted, and attained.

Beaman, Karen V.; Guy, Gregory R. (eds.) (2022) *The coherence of linguistic communities: Orderly heterogeneity and social meaning*. New York: Routledge.

Erker, Daniel (2017) Contact, co-variation, and sociolinguistic salience: What Mister Rogers knows about language change. *University of Pennsylvania Working Papers in Linguistics* 23(2), 68–78.

Guy, Gregory R.; Hinskens, Frans (2016) Linguistic coherence: Systems, repertoires and speech communities. *Lingua* 172—173, 1—9.

Labov, William (2006 [1966]) *The social stratification of English in New York City*. Cambridge: Cambridge University Press.

Tamminga, Meredith (2019) Interspeaker covariation in Philadelphia vowel changes. *Language Variation and Change* 31, 119—133.

---

## Investigating the link between intelligibility and attitudes towards varieties of French

Kaitlyn Owens, Monica Nesbitt, Jeffrey Lamontagne  
Indiana University, Bloomington, USA

**Introduction.** Whereas speakers from France note that Canadian French is incomprehensible (e.g. Schafroth, 2009) and have negative attitudes towards it (e.g. Slezák, 2015), an empirical study probing the relationship between intelligibility and attitudes has not been conducted for French varieties. In this study, we build upon methodologies of previous studies investigating intelligibility and attitudes (Schüppert et al., 2011; 2015) to explore their link in Canadian French (CF) and Hexagonal French (HF; France). We test intelligibility using a word-recognition task to assess whether French participants have difficulty comprehending CF and elicit attitudes using a perception task to determine if attitudes align with intelligibility results. 56 participants completed the study (20 Canadian; 36 French).

**Intelligibility.** In this forced-choice task, participants identified whether an oral stimulus matched a written stimulus. Stimuli were created in pairs: one undergoes a phonological processes characteristic of CF, and its pair is a phonetically similar minimal pair that was selected based on the given CF phonological process. Participants heard talkers from both dialects. We predict token identification and reaction time using mixed-effects regression models on 4480 decisions. Models include participant dialect, stimulus dialect, and their interaction; participant and stimulus are random intercepts. The reaction time model also includes identification as a main effect and its interactions with participant dialect and stimulus dialect. Figure 1 illustrates that participants incorrectly identified CF stimuli more often than HF stimuli ( $p=0.0216$ ). Figure 2 shows that participants took longer to respond to CF stimuli than HF stimuli ( $p=0.0381$ ). Participants also took longer responding to CF stimuli which were incorrectly identified in comparison to other stimuli ( $p=0.0007$ ). Participant dialect is not a significant predictor in either model.

**Attitudes.** Participants answered questions measuring perceived intelligibility, solidarity, and status of CF and HF talkers using a sliding scale. We use mixed-effects linear regressions on 448 responses per question with participant dialect, stimulus dialect, and their interaction; participant is a random intercept. Our questions targeting perceived intelligibility demonstrate French participants found it more effortful to understand Canadian talkers ( $p=0.0005$ ). Additionally, French participants rated Canadian talkers lower on the solidarity dimension ( $p=0.0082$ ). No differences were found for status.

**Discussion.** Our results show that CF is not exceptionally difficult or effortful for French participants to comprehend, despite attested notions of incomprehensibility in the literature. While CF stimuli were incorrectly identified more often and have longer reaction times than HF stimuli, it is true for both participant groups. We posit this is because the competitor stimuli were selected based on CF, rather than HF. Further, attitudes results demonstrate that French participants possess negative attitudes towards CF based on solidarity, yet there are no significant differences according to status; we postulate this reflects an “othering” of Canadian talkers by French participants, which is not present for Canadian participants towards French talkers. Additionally, French participants perceive that CF is less intelligible than HF. Given the asymmetry between intelligibility results and attitudes, we propose that French speakers' attitudes towards CF influence notions of intelligibility, rather than it being influenced by comprehension issues.

Schafroth, Elmar. (2009). Die französische Standardsprache in Québec. In Ursula Reutner (Ed). 400 Jahre Quebec. Kulturkontakte zwischen Konfrontation und Kooperation. Heidelberg: Winter, 45–72.

- Schüppert, Anja & Charlotte Gooskens. (2011). Investigating the role of language attitudes for perception abilities using reaction time. *Dialectologia* II. 119–140.
- Schüppert, A., Hilton, N. H., & Gooskens, C. (2015). Swedish is beautiful, Danish is ugly? Investigating the link between language attitudes and intelligibility. *Linguistics*, 53(2), 275-304.
- Slezák, Edith. (2015). Prestige et stigmatisation dans les variétés du français canadien. *Zeitschrift für Kanada-Studien* 35: 8-24.
- 

Northern Maine as a transition zone:  
Evidence from rhoticity in Southern Aroostook County English

Katharina Pabst  
Radboud University, Nijmegen, Netherlands

Previous dialectological work has yielded inconclusive results with regard to the question how Northern Maine fits into the surrounding dialect regions. Kurath (1939:17) groups it with nearby New Brunswick, but most large-scale projects, including the Linguistic Atlas of New England (LANE, which Kurath bases his claims on), the Dictionary of Regional English (DARE), and the Atlas of North American English (ANAE), collected little to no data from the area; even very recent projects based on crowdsourced data include only a handful of speakers from Northern Maine (Kim et al. 2019, Stanford 2019). This project addresses this gap in the research literature by examining rhoticity (i.e., the realization of post-vocalic /r/ in coda position), one of the most salient features in the area, and what it can tell us about the historical development and current state of the local variety.

Data for this project come from a word- and sentence-list task conducted with 28 white and two Native American speakers from Southern Aroostook County, Maine (sample stratified by age and gender, word list adapted from Stanford et al. 2012). All test words where postvocalic /r/ occurred in coda position were auditorily coded as ‘present’ or ‘absent’ by two coders; whenever coders did not agree, tokens were coded as ‘ambiguous.’ For the statistical analysis, these tokens were grouped with the absent tokens and coded as ‘vocalized.’ The data were further coded for a variety of linguistic (word, style, position, preceding vowel, lexical frequency) and social factors (individual speaker as well as their age, gender, education, and local affiliation), adapted from Nagy & Irwin (2010). The final analysis is based on 1,309 tokens, which were analyzed using distributional analysis, conditional inference trees, and mixed effects logistic regression.

Results show that speakers of Southern Aroostook County English are overwhelmingly r-ful, with 93.5% of all tokens coded as present. Only three tokens (0.2%) were unambiguously absent; all of these occurred in instances of the proper noun New Hampshire, which has likely become lexicalized. However, there are a sizable number of ambiguous tokens (n=82). Most of these come from speakers born before 1973, but there is substantial individual variation among speakers. This, in combination with the fact that several surrounding communities are (still) r-less, provides strong evidence that non-rhoticity is a minority variant (likely a remnant from early settlers) and is quickly falling out of use, in line with supra-local trends (see Kim et al. 2019, Stanford 2019). The mix of rhotic and non-rhotic pronunciations makes Southern Aroostook County a “transition zone” (Chambers & Trudgill 1998:105) between Eastern New England and Atlantic Canada, providing an explanation for why previous work has found competing evidence for and against Northern Maine patterning alongside the two dialect regions.

- Chambers, J. K. & Trudgill, P. (1998). *Dialectology*. Second edition. Cambridge, UK: Cambridge University Press.
- Kim, C., Reddy, S., Stanford, J. N., Wyszogrod, E., & Grieve, J. (2019). Bring on the crowd! Using online audio crowd-sourcing for large-scale New England dialectology and acoustic sociophonetics. *American Speech*, 94(2), 151–94.
- Kurath, H. (1939). *Handbook of the linguistic geography of New England*. With the collaboration of Marcus L. Hansen, Bernard Bloch, and Julia Bloch (Original work published 1939). Washington, DC: American Council of Learned Societies. (Reprinted in 1954).
- Nagy, N., & Irwin, P. (2010). Boston (r): Neighbo(r)s nea(r) and fa(r). *Language Variation and Change*, 22(2), 241–78.
- Stanford, J. N. (2019). *New England English: Large-scale acoustic sociophonetics and dialectology*. New York, NY: Oxford University Press.
- Stanford, J. N., Leddy-Cecere, T. A., & Baclawski, Jr., K. P. (2012). Farewell to the founders: Major dialect changes along the east-west New England border. *American Speech*, 87(2), 126–69.
- 

Regressive voicing assimilation in Modern Greek dialects  
Panayiotis Pappas, Adeola Babayode-Lawal, Laurens Bosman  
Simon Fraser University, Vancouver, Canada

We will discuss the dialectal variation of a morphophonological voicing assimilation pattern in Modern Greek that has received little attention in the literature. The forms /ton/, /tin/, /to/ are accusative singular forms of the masculine, feminine and neuter gender, either for the definite article or for the clitic object pronoun. For example: /ton polemo/ "the war", /ton pira/ "I took him." In Standard Modern Greek (Arvaniti, 1999; Holton et al., 2012;



Kainada, 2011), the /n/ of the masculine and feminine forms causes a following voiceless stop to become voiced, while it undergoes place assimilation. Thus, the forms above are pronounced [tombolemo] and [tombira] respectively.

It has been reported (Kontosopoulos, 2001; Vadasi, 2007) that in some northern Greek varieties, the /t/ of the article or pronoun is also voiced, meaning that in these varieties, the examples above are pronounced [dombolemo] and [dombira]. However, there are no detailed accounts of this pattern. We examine it in a dataset constructed from 208 transcribed interviews of first-generation Greek Canadians (Anastassiadis et al., 2017) who moved to Canada between 1945 and 1975 as young adults (median age: 22 years of age). The sample contains 120 (58%) men & 88 (42%) women born between 1935 and 1965 and represents 5 of the 15 major varieties of Modern Greek: Athenian (standard), Northern, Peloponnesian, South - Eastern and Eastern (Trudgill, 2003).

We have taken 20-minute segments of each interview which have been automatically annotated on the basis of the transcriptions by using a Modern Greek model for the Montreal Forced Aligner (McAuliffe et al., 2017). We have extracted 9,045 tokens of these constructions in phrase initial position, which we are manually coding for the presence or absence of voicing in the initial segment of the determiner or pronoun. So far, we have examined 1000 of these tokens and found corroborating evidence for the pattern as only the masculine and feminine forms have a voiced /t/ (figure 1), whereas the /t/ of the neuter form remains unvoiced. A phonologically interesting finding is that both types of assimilation (progressive and regressive) occur even when the nasal is deleted, that is we find instances of examples such as [doblemo] and [dobira]. In terms of geographical distribution, we find voicing not only in speakers from Northern Greek but also in some speakers from Central Greece. Finally, voicing is more prevalent among women. We hypothesize that this is the case because the men were employed outside the household and had broader social networks within the Greek community that encouraged a more standard pronunciation, whereas the women worked mostly in the house and had much narrower social networks.

Anastassiadis, A., Ralli, A., Gekas, A., Pappas, P. A., Papanagioutou, C., Siotou, A., Tsimpouris, C., & Tsolakidis, S. (2017).

Immigration and Language in Canada: Greeks and Greek Canadians. [Electronic Database]. <https://immigrec.com/en>. Arvaniti, A. (1999). Standard Modern Greek. *Journal of the International Phonetic Association*, 29(2), 167–172.

Holton, D., Mackridge, P., Philippaki-Warbuton, I., & Spyropoulos, V. (2012). *Greek: A Comprehensive Grammar of the Modern Language* (2nd ed.). Routledge. <https://doi.org/https://doi-org.proxy.lib.sfu.ca/10.4324/9780203802380>

Kainada, E. (2011). The acoustics of post-nasal stop voicing in Standard Modern Greek. <https://doi.org/10.13140/2.1.1146.0483>

Kontosopoulos, N. (2001). *Διάλεκτοι και ιδιώματα της Νέας Ελληνικής* [Dialects and varieties of Modern Greek] (3rd ed.). Grigori Publications.

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. *Proceedings of the 18th Conference of the International Speech Communication Association (INTERSPEECH 2017)*, Stockholm, Sweden 20 - 24 August 2017, 498–503.

Trudgill, P. (2003). Modern Greek dialects: A preliminary classification. *Journal of Greek Linguistics*, 4, 45–63.

Vadasi, K. (2007). Το ρηματικό κλιτικό παράδειγμα στη Διάλεκτο της Δεσκάτης και των περιχώρων της [The verbal clitic phrase in the dialect of Deskati] [Aristotle University]. <https://doi.org/10.13140/RG.2.1.1769.5441>

---

## Social and Phonological Factors in (r) Variation in 1930s Inland Massachusetts

Joe Pater<sup>1</sup>, James Stanford<sup>2</sup>

<sup>1</sup>University of Massachusetts, Amherst, USA. <sup>2</sup>Dartmouth College, Hanover, USA

Bloch's 1935 dissertation [1] presents a seemingly untapped wealth of data and observations about the (r) variable in New England based on LANE records [2, 3]. This paper builds on his work in a study of the 50 informants in the areas Bloch classifies as Western and Central Massachusetts, which lie at the crossroads of the North-South and East-West New England isoglosses in ANAE [4]. Our findings:

1. (r) had a complex social distribution in this area in the 1930s: non-rhoticity was prevalent amongst upper middle class urban speakers in Northampton and Springfield and amongst rural speakers to the east and north, while to the west rhoticity prevailed.

2. (r) was subject to a gradient phonological constraint that mirrors a categorical constraint for contemporary Eastern Massachusetts (r)

Bloch chose 145 words with final or preconsonantal (r) that occur in most informants' records, divided into three phonological classes: stressed vocalic (e.g. nurse), unstressed vocalic (e.g. letter), and post-vocalic (e.g. part). For each informant and word, he registered whether the fieldworker transcribed (r) as fully retroflex, weakly retroflex or non-retroflex.

We calculated a retroflexion score for each informant for each phonological class that ranges from 0 for no retroflexion to 1 for consistent full retroflexion (weak retroflexion contributed 0.5 to the numerator in the proportion-like formula). Means and standard deviations for each class in each county appear in (1). Berkshire has the highest rate of retroflexion, and Franklin and Worcester the lowest, with Hampshire and Hampden intermediate. Across all of them, stressed vocalic is the class that is most often retroflexed. McCarthy [5] claims that retroflexion

occurs always and only in stressed vocalic position in late 20th century Eastern Massachusetts English. The 1930s inland Massachusetts data demonstrate that the preference for stressed nucleus rhoticity can also hold as a tendency (see relatedly [6]). Individual informant data, as well as fieldworker notes and recordings, confirm that the patterns of variation hold at an individual as well as community level.

A mixed effects logistic regression with integer-rounded retroflexion scores as a dependent variable appears in (2). More education predicts lower retroflexion ( $p = 0.034$ ), and stressed vocalic position predicts higher retroflexion ( $p = 0.096$ ). Informants that LANE labels “cultured” are predicted to have lower retroflexion scores, though the effect is not significant. The association of lower retroflexion with relatively “higher class” informants is confirmed in other LANE records. This contrasts markedly with the situation in NYC thirty years later [7].

This study lays the groundwork for two ongoing projects. The first examines archival oral histories [8] and contemporary data [9] to trace the subsequent changes in the socio-phonology of (r) in this area. The second examines the vowel systems of these 50 informants to better understand the role of this area in the genesis of the Northern Cities Shift [10] and the relationship to other developments in Western New England [11]. A general conclusion is that not only the LANE recordings [12], but also the transcriptions provide valuable insight into variation and change.

- [1] Bloch, Bernard. 1935. The treatment of Middle English final and preconsonantal R in the present-day speech of New England. PhD dissertation, Brown University.
- [2] Kurath, Hans. 1939. Handbook of the Linguistic Geography of New England. Providence: Brown University for the American Council of Learned Societies. (2d ed., 1973).
- [3] Linguistic Atlas of New England Records (MS 330). Special Collections and University Archives, University of Massachusetts Amherst Libraries.
- [4] Labov, William, Sharon Ash and Charles Boberg. 2006. The Atlas of North American English: Phonetics, Phonology and Sound Change. Berlin and New York: Mouton de Gruyter.
- [5] McCarthy, John J. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics* 23, 169-195
- [6] Nagy, Naomi and Patricia Irwin. 2010. Boston (r): Neighbo(r)s nea(r) and fa(r). *Language Variation and Change* 22, 241-78.
- [7] Labov, William. 1966. The social stratification of English in New York City. Washington DC: Center for Applied Linguistics.
- [8] Oral history collection, Historic Northampton. <https://www.historicnorthampton.org>.
- [9] Nesbitt, Monica, and James Stanford. 2021. Structure, chronology, and local social meaning of a supra-local vowel shift: Emergence of the Low-Back-Merger Shift (LBMS) in New England. *Language Variation and Change* 33, 269-95.
- [10] Boberg, Charles. 2001. The Phonological Status of Western New England. *American Speech* 76.
- [11] Roberts, Julie. 2007. Vermont lowering? Raising some questions about /ai/ and /au/ south of the Canadian border. *Language Variation and Change*, 19, 181-197.
- [12] Johnson, Daniel Ezra and David Durian. 2017. New England. In R. Hickey (Ed.), *Listening to the Past: Audio Records of Accents of English* (Studies in English Language, pp. 257-297). Cambridge: Cambridge University Press.

---

## Is MaxEnt Cumulativity too restrictive? Evidence from a variety of NYC AAE

Joe Pater

University of Massachusetts Amherst, Amherst, USA

Probabilistic constraint-based models of grammar provide an important bridge between phonology and sociolinguistics [1, 2]. A particularly tight link is provided by Maximum Entropy Grammar (MaxEnt) [3], a generalization of logistic regression that uses constraints with numeric weights. Some recent evidence favoring MaxEnt over Stochastic Optimality Theory [4] is its ability to better model the effect of cumulative constraint violations in French variable schwa alternations French [5]. This paper shows that the predictions of the MaxEnt model adopted in that study may be too strong, given data from coronal stop deletion (CSD) in a variety of AAE spoken in New York City in the 1960s [6]. It goes on to provide initial comparisons amongst alternative frameworks.

Analyzed in constraint-based terms, the CSD data provide an example of cumulativity that is exactly parallel to the one in the French schwa alternations. In (1), the proportion of realized coronal stops in four morpho-phonological environments for two groups of preadolescent speakers is provided. These show two now-familiar effects: the stops are realized more often before vowels than consonants (e.g. CT#V vs. CT#C), and more often when they are the phonological exponent of the past tense morpheme (e.g. C+T#C vs. CT#C). The bottom row shows the effect of cumulativity, where both constraints apply; this is the environment with the least deletion.

The table in (2) provides the constraint violations in a standard phonological approach to CSD [1]. When the stop is realized, it violates a structural constraint against the coronal-final cluster, \*CT. When it is deleted, it violates Max, the general constraint against deletion. MaxPreV is violated when the stop is prevocalic, and Realize-Morph is violated when the stop would have been the realization of the past tense morpheme.

When the Aces data are used to train a MaxEnt model, the model's predictions closely match the empirical data (MaxEnt Aces in (3)). For the 1390 Lames data, the model's predictions diverge markedly (MaxEnt Lames in (3)). The challenge of the Lames data is that the cumulative effect of MaxPreV and Realize-Morph is too weak given the effects of each constraint on its own. Viewed as a regression model, this is evidence of an interaction. The phonological equivalent of an interaction term is a conjoined constraint [7].

The table in (3) shows the result of fitting Stochastic OT and MaxEnt models with and without a conjunction between MaxPreV and Realize-Morph. Without, Stochastic OT fails to successfully model the Aces data (OT Aces), but matches the Lames data well (OT Lames), the reverse of the MaxEnt outcome. With the conjoined constraint, both Stochastic OT (OTConj) and MaxEnt (MaxEntConj) succeed on both datasets. The success of the MaxEnt model depends on a negatively weighted conjoined constraint, which counterbalances the cumulative interaction. This is potentially problematic. Versions of MaxEnt that allow negative weights often generate unattested patterns [1]. In this case it would allow anti-cumulativity, where the joint effect of two constraints is less than the effect of a single one.

- [1] Coetzee, Andries and Joe Pater. 2011. The place of variation in phonological theory. In John Goldsmith, Jason Riggle, and Alan Yu (eds.), *The Handbook of Phonological Theory* (2nd ed.). Blackwell. 401-431.
- [2] Fruehwald, Josef. 2022. The study of variation. Chapter 27 in *The Oxford History of Phonology*.
- [3] Sharon Goldwater and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*.
- [4] Boersma, Paul. 1997. How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21. 43-58.
- [5] Smith, Brian W. and Joe Pater. 2020. French schwa and gradient cumulativity. *Glossa: a journal of general linguistics*.
- [6] Labov, William. 1973. The linguistic consequences of being a lame. *Language in Society*.
- [7] Smolensky, Paul 1993. Harmony, markedness, and phonological activity. Paper presented at Rutgers Optimality Workshop 1.

---

## Code-switching datasets for NLP: Using intonation for elementary units

Rebecca Pattichis<sup>1</sup>, Rena Torres Cacoullos<sup>2</sup>

<sup>1</sup>University of California, Los Angeles, USA. <sup>2</sup>Penn State University, University Park, USA

Code-switching (CS)—going back and forth between languages—is increasingly garnering attention in Natural Language Processing (NLP) (Indra Winata et al. 2022:1; Solorio et al. 2021), however, models and metrics are inadequate due to lack of suitable training data for NLP tasks (Seza Doğruöz et al. 2021). In particular, metrics of CS complexity have been word-level based (e.g. Gambäck & Das 2016). Such word-level metrics entail that CS is equally likely between any two words and assume that single-word and multi-word items behave the same. Here, we adapt NLP metrics to account for prosodic and syntactic constraints on CS, and calculate them separately for distinct mixing types. We draw on a bilingual community-based speech corpus, transcribed in Intonation Units (IUs). IUs are prosodic chunks “uttered under a single, coherent intonation contour” (Du Bois et al. 1993:47). Importantly, words in the same IU have a closer syntactic relationship than words at IU boundaries (Croft 1995:849-864). We sample five speakers (N=4.8 hours, 41,000 words, 14,000 IUs).

CS is prosodically structured: bilinguals tend to prosodically separate their two languages. The IU-Boundary constraint states that bilinguals prefer CS across IU boundaries (cf. Torres Cacoullos & Travis 2018: 51). CS is four times more likely at IU boundaries than within IUs and, at IU boundaries, CS is twice as likely following intonational completion than continuing intonation (Figure 1). The IU-Boundary constraint corresponds to the well-established syntactic Equivalence constraint, which requires local equivalence of word order in the two languages around a switch point (Poplack 1980). In addition, bilinguals treat multi-word CS differently from single-word incorporations, or lone items, with respect to switch points. While multi-word switches are placed at cross-language equivalence points, lone items tend to follow recipient-language word order (Poplack 2018) (e.g., for Spanish-English NPs, "los big sizes" 'the big sizes', with multi-word CS after the determiner, vs. "unos desks grandes" 'some big-postdesks', with a lone noun and postnominal adjective).

The Multilingual Index (M-Index) (Barnett et al. 2000) measures the multilinguality of a corpus, with values from 0 (=monolingual) to 1 (=a balance of languages). The Integration Index (I-Index) (Guzman et al. 2017) measures the probability (observed rate) of CS. We adapt these metrics to the prosodic structure of CS, using IUs rather than individual words as tokens. We separately calculate the I-index for switches across and within IU boundaries, first for multi-word CS and then including lone items (Table 1). Individual speakers differ in M-Index, ranging 0.52-0.98, and, independently, in I-Index for across-IU CS, ranging 0.03-0.16. Nevertheless, they uniformly show low I-indices for within-IU CS, ranging 0.0-0.01. When lone items are included, the I-index increases, modestly for across-IU, ranging 0.04-0.18, but greatly for within-IU (300% for 03), now ranging 0.01-0.04. Thus, merging lone items with multi-word CS misses the IU-Boundary constraint, which applies uniformly despite

individual differences in multilinguality (M-Index) and CS rate (I-Index). These results highlight the IU as a suitable token level for future construction of transcribed CS data sets and synthetic data generation.

- Barnett, R. et al. (2000). The lides coding manual: A document for preparing and analyzing language interaction data version 1.1– July 1999. *International Journal of Bilingualism* 4(2):131–271.
- Croft, W. (1995). Intonation units and grammatical structure. *Linguistics*63, 839-882.
- Seza Dođruöz, A., Sitaram, S., Bullock, B.E., & Toribio, A.J. (2021). A survey of code-switching: Linguistic and social perspectives for language technologies. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*(Volume 1: Long Papers), 1654–1666.
- Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription. In *Talking data: Transcription and coding in discourse research*, J. A. Edwards & M. D. Lampert (Eds.), 45–89. Lawrence Erlbaum.
- Gambäck, B. & Das, A. (2016). Comparing the level of code-switching in corpora. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 1850–1855, Portorož, Slovenia. European Language 595 Resources Association (ELRA).
- Guzmán, G., Ricard, J., Serigos, J., Bullock, B. & Toribio, A. (2017). Metrics for modeling code-switching across corpora. *Interspeech*2017, 67–71.
- Indra Winata, G., Fikri Aji, A., Yong, Z.-X., and Solorio, T. (2022). The decades progress on code-switching research in NLP: A systematic survey on trends and challenges. *arXiv e-prints*.
- Poplack, S. (1980). “Sometimes I’ll start a sentence in Spanish y termino en español”: toward a typology of code-switching. *Linguistics*18, 581-618.
- Poplack, S. (2018). *Borrowing. Loanwords in the speech community and in the grammar*. Oxford UP.
- Solorio, T., Chen, S., Black, A.W., Diab, M., Sitaram, S., Soto, V. (2021). *Proceedings of the Fifth Workshop on Computational Approaches to Linguistic Code-Switching(CALCS)*.
- Torres Cacoullous, R. & Travis, C.E. (2018). *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge UP.

---

## Sociolinguistic Variation in American Sign Language in Washington State

Yadi Peng

University of Washington, Seattle, USA

Washington State (WA) has a long-standing deaf and hard of hearing (DHH) community which have contributed to the richness of sociolinguistic variation in American Sign Language (ASL). Since Washington State School of the Deaf (WSD) opened in 1886, they have modified their pedagogical approaches due to language policies and language attitudes: they taught only oral English at some times and both English and ASL at other times [6]. On the other hand, organizations like the Cascadia Black Deaf Center suggest a history of Black signers in the state, but no record has shown their enrollment in the deaf educational system.

A deficiency of studies including Black signers in WA has left a knowledge gap. I will study the question: How are race, age, and schooling associated with lexical and phonological variation among near-native signers?

Lexical variation has been found to correlate with race, age, and schooling. McCaskill and colleagues [4] found that Black signers produced a large number of signs that White signers did not. Age is an important factor. First, multiple variants could coexist as new technology (e.g. microwaves) emerges when a generation grows up. Second, age was also found to correlate with potentially racist variation of signs such as AFRICA and JAPAN [3]. Third, older signers may have adopted negative attitudes towards ASL. Unlike DHH students at deaf schools, those who have been mainstreamed might have learned English-based communication systems [1].

ASL has five parameters: handshape, orientation, location, movement, and non-manual markers [5]. Phonological variation in handshape, location, and number of hands is constrained by linguistic environment, region, race, age, and schooling. Thumb extension (extended vs. closed) in FUNNY and location (high vs. low) in KNOW vary systematically with regard to signers’ region and age [3]. [4] showed that older Black signers who attended segregated schools produced more two-handed variants (e.g. THANK-YOU) and used larger signing spaces. I will recruit 30 near-native ASL users, balancing the numbers in each group of age, gender, race (Black or White), and type of school attended (deaf, mainstream, or both). Elicitation tasks will involve three types of stimuli motivated by [3]. Type 1 will be pictures, indicating signs potentially unique to WA (e.g., STRAWBERRY, SALMON). Type 2 will be written English to elicit variation in lexicalized fingerspelled signs (e.g., #ALL, #SURE). Type 3 will be images that elicit signs with initialized variants (e.g., FAMILY, CULTURE). Signers will be videotaped while demonstrating the signs.

Signers’ age, race, and schooling are the main predictors, which will be coded as categorical variables. Following the transcription conventions in [2], I will code the variation as categorical data. As multiple phonological variants might be nested in one lexical variant, I will build a multinomial logistic regression model to account for the

collected data and to predict future data. Given previous research, older, Black signers who have gone to deaf schools are expected to show more variants closer to citation forms [3][4].

[1] Hill, J.C. 2012. "Language Attitudes in the Am. Deaf Community".

[2] Liddell, S. K., & Johnson, R. E. 1989. *ASL: The phonological base*.

[3] Lucas et al. 2001. "Sociolinguistic Variation in ASL."

[4] McCaskill et al. 2011. "The Hidden Treasure of Black ASL: Its History and Structure".

[5] Valli et al. 2011. "Linguistics of ASL: An Introduction".

[6] WSD [Washington School for Defective Youth]. 1887–1889. Second biennial report of the trustees of the Washington School for Defective Youth. 1906-1908. Biennial report of the superintendent. 1910-1912. Biennial report of the superintendent.

---

## Spanish rhotic production and dialectal variation in Mexican Heritage Speakers

Ileana Perez

The Ohio State University, Columbus, USA

Heritage speakers are often defined as members of a linguistic minority who grew up exposed to their home language and the majority language (Montrul 2010). Few studies have looked at Heritage speaker's Spanish language use, often comparing it to monolingual speakers. Nevertheless, comparing Heritage speakers to monolingual speakers in phonetic studies can be challenging because of extralinguistic factors such as speakers' identities that may influence Heritage speakers' Spanish production (Kissling 2018). This study further explores Heritage speakers' production of dialectal rhotics, specifically looking at contexts where they occur.

Spanish has two contrastive rhotics, a tap and a trill, that present much dialectal variation including breathy, fricated, assibilated, lateralized or deleted realizations (Bradley et al., 1999, Bradley & Willis 2012, Caranza 2006). In Mexico, dialectal rhotics such as the assibilated rhotic, can be found (Mazarró & De Anda 2019, 2020). Studies looking at rhotic production amongst English-Spanish bilingual and heritage speakers have focused on whether participants maintain the contrast between tap and trill, all finding that these speakers are able to keep a distinction between them (Cummings Ruiz et al., 2020, Kissling 2018, Henriksen 2015). Only one study has analyzed Heritage speakers' rhotic variation from a dialectal perspective and found that, Heritage speakers are able to produce dialectal rhotics such as the assibilated rhotic (Cummings Ruiz et al., 2020). However, Cummings Ruiz et al., (2020) did not examine what factors influence dialectal rhotics productions amongst Heritage speakers. I build upon previous literature to further explore Heritage speakers' production of dialectal rhotics, focusing on what contexts favor them, while examining whether they maintain the tap/trill contrast, including how it is manifested.

This study looks at the rhotic production of 6 Heritage Spanish speakers of Mexican descent that were born and currently reside in Southern California. Each speaker completed a short interview, the Bilingual Language Profile questionnaire (Birdsong et al., 2012), a picture description task, and a narration task. Based on acoustic cues, 705 tokens of rhotic productions from the picture description and narration were classified as either tap, trill, approximant tap, approximant trill, fricated tap, fricated trill, assibilated, lateralization or deletion. Tokens were also coded for syllable position (onset, coda; only for taps), word position (initial, medial, final), and utterance position (initial, medial, final), in order to understand the role of these factors on rhotic production for Heritage speakers.

For expected tap realizations, approximant taps made up most tokens for all syllable, word, and utterance positions. Assibilated rhotics were found in coda positions, including word final position, and trills made up less than 2% of the data. For expected trill realizations, in utterance-initial position, approximant trills and fricated trills made up most of the data, 72.2%, while in non-utterance initial contexts, fricated trills made up most of the tokens. These results suggest that Heritage speakers maintain the tap and trill contrast, however their productions tend to have different degrees of weakening and frication. Heritage speakers also produce dialectal rhotics, more precisely assibilated ones, but at lower rates than studies looking at monolingual speakers.

Bradley, T. G., Authier, J., Bullock, B., & Reed, L. (1999). Assibilated rhotics in Ecuadorian Spanish. In *Formal Perspectives on Romance Linguistics. Selected Papers from the 28th Linguistics Symposium on Romance Languages*. Amsterdam, John Benjamins Publishing Company (pp. 57-71).

Bradley, T. G., & Willis, E. W. (2012). Rhotic variation and contrast in Veracruz Mexican Spanish. *Estudios de fonética experimental*, 43-74.

Carranza, L. M. V. (2006). On the phonetic realization and distribution of Costa Rican rhotics. *Revista de filología y lingüística de la Universidad de Costa Rica*, 291-309.

Cummings Ruiz, L. D., & Montrul, S. (2020). Assessing Rhotic Production by Bilingual Spanish Speakers. *Languages*, 5(4), 51.

Kissling, E. M. (2018). An Exploratory Study of Heritage Spanish Rhotics: Addressing Methodological Challenges of Heritage Language Phonetics Research. *Heritage Language Journal*, 15(1), 25–70.

Mazarró, N., & González de Anda, R. G. (2019). The perception-production connection: /tʃ/ deaffrication and rhotic assibilated in Chihuahua Spanish. *Issues in Hispanic and Lusophone Linguistics series: Recent Advances in the Study of Spanish Sociophonetic Perception*. Edited by Whitney Chappell. Amsterdam: John Benjamins Publishing Company

---

## Variation at the Lexicon-Syntax interface in Spanish: dialectal differences in copular sentences with adjectival predicates

Isabel Pérez-Jiménez, Norberto Moreno-Quibén, Silvia Gumiel  
Universidad de Alcalá, Alcalá de Henares, Spain

In the literature on Spanish copular verbs, it has been claimed, from different theoretical perspectives, that *estar* is a stage-level copula: in copular sentences with *estar*, a bounded property is attributed to a spatiotemporally located stage of the subject of predication (aspectual approaches to the *ser-estar* distinction) or else the comparison of stages of the subject of predication is necessary to attribute the property to that subject (modes-of-predication approaches to the *ser-estar* distinction). Understanding *estar* as a stage-level copula explains the ungrammaticality of sentences such as \**La fiesta estuvo posible*; (lit. The party was<sub>-estar</sub> possible) in Standard Spanish (vs. *La fiesta fue posible*; lit. The party was<sub>-ser</sub> possible;), since modal adjectives are predicated of eventive subjects and events are a kind of semantic entity that by definition has no stages.

However, a different use of *estar* has also been described in the literature, as illustrated in the following example: *El dueño de este lugar [restaurant] tiene otro en el centro que está más pequeño* “The owner of this place has another one downtown, which is<sub>-estar</sub> smaller” (adapted from Gutiérrez 1994: 124; Mexico). In this *innovative use* (Silva Corvalán 1986) or *perspectivized use of estar* (Gumiel-Molina, Moreno-Quibén, Pérez-Jiménez 2020, 2023, Moreno Quibén 2022), the property is not necessarily understood as bounded, nor does its attribution imply the comparison between stages of the subject (“the restaurant”). This innovative/perspectivized use, according to the literature, is typical of American Spanish, except for the Rioplatense dialectal area (Uruguay and Argentina).

However, not all lexical-syntactic classes of adjectives show a similar geographical distribution and frequency in this structure. Thus, for Mexico and Central America, Gutiérrez (1994), Ortiz-López (2000), Brown and Cortés-Torres (2012), García Márkina (2013) note that the adjectives found in this context are, in order of frequency of occurrence, age adjectives, followed by size and evaluative adjectives of different kinds. Other adjectives, such as color adjectives, seem to be rejected. Gutiérrez (1994) connects this fact with the meaning of the innovative structure, which expresses the speaker's subjective point of view about the attribution of the property, so that the meaning of the adjective must be compatible with that meaning component.

In this context, the goal of this talk is twofold. The empirical goal of the talk is to offer an exhaustive and updated characterization of the extent of this structure in Latin American Spanish, considering both its geographical distribution in the different American dialectal varieties and the lexical-syntactic classes of adjectives that appear as predicates in innovative *es-tar*-sentences. Our study will be based on an empirical analysis of the Preseea corpus. Taking these two parameters together (dialectal area where the innovative structure is documented / lexical-syntactic class of the adjectival predicate), we obtain the following generalization, formulated as an implicative hierarchy, regarding the extent of the innovative *estar*-structure in Latin American Spanish. According to this hierarchy if, for example, relational predicates are attested in the innovative structure in a specific area, then, the rest of preceding adjectival classes are also attested in the innovative *estar*-construction in that dialectal area.

Generalization: age adjectives (documented in all areas in the innovative construction) > evaluative adjectives of the ‘Other’ class (all areas) > evaluative adjectives of personal judgement and maximum degree (documented in Mexican & Central American, Caribbean, Andean areas) / aesthetic evaluatives (documented in Mexican & Central American, Caribbean, Rioplatense areas) > dimensional/ property/ adverbial adjectives used as predicates (documented in Mexican & Central American, Caribbean areas) > relational and modal adjectives used as predicates (documented in Mexican & Central American area).

Therefore, the so-called innovative use of *estar* is a robust phenomenon of variation characteristic of the Mexican & Central American area, which occurs to a lesser extent in the Caribbean area and extends non-systematically to other dialectal areas. With respect to the lexical-syntactic classes of adjectives, age adjectives and some sub-classes of evaluative adjectives (aesthetic adjectives, and the heterogeneous class that we have termed ‘Other’) occur in innovative examples in all dialectal areas. Let us illustrate with a modal adjective:

I: pero, pero bueno, yo insisto, lo menos importante es el toro ¿eh?, lo menos importante para los antitaurinos es que se maten o no los toros, creo que hay mucho / mucho aderezo y algunos otros objetivos que realmente son su // eh eso , su objetivo (...) espero que tenga muchos alumnos que se pongan a teclear por evitar eso, **está imposible** eh, bueno (Mexico, PUEB\_H33\_086)

‘I: but, but well, I insist, the least important thing is the bull, eh, the least important thing for the anti-bullfighting people is whether the bulls are killed or not, I think there is much / much pretending and some other objectives that

are really their // eh that, their objective (...) I hope you have many students who will start typing to avoid that, **it is. ESTAR impossible** eh, well'

From the theoretical point of view, in the light of the generalizations established, a critical evaluation will be made of the existing proposals in the literature that explain the existence and properties —both syntactic and semantic— of the innovative structure. The distribution pattern found will be explained by elaborating the idea defended in Moreno-Quibén (2022) that the argument structure of lexical items, specifically adjectives, is subject to syntactic variation. The fact that certain classes of adjectives may have in some varieties a dative experiencer as an argument, on which the attribution of the property can be based, would explain the grammaticality of these examples since no stages of the subject are accessed in the predication relation. The presence of the experiencer is ultimately responsible of the subjective meaning of the construction in the Spanish varieties in which it is found.

- Brown, E. L. & M. Cortés-Torres (2012). 'Syntactic and Pragmatic Usage of the [estar + Adjective] Construction in Puerto Rican Spanish: ¡Está brutal!', in K. Geeslin and M. Díaz-Campos (eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, Somerville, MA, Cascadia Proceedings Project, pp. 61-74.
- Demonte, V. (1999). 'El adjetivo. Clases y usos. La posición del adjetivo en el sintagma nominal', in I. Bosque and V. Demonte (dirs.), *Gramática descriptiva de la lengua española*, vol. I, Madrid, Espasa-Calpe/RAE, pp. 129-215.
- García-Márkina, Y. (2013). *Recherches sur l'opposition entre ser et estar en espagnol. Historique de la question, et application à l'étude des variations dans leurs emplois en espagnol spontané contemporain au Mexique*, Doctoral Dissertation, l'Université Sorbonne Nouvelle – Paris 3.
- Gumiel-Molina, S., N. Moreno-Quibén & I. Pérez-Jiménez (2020). 'Variación dialectal en las oraciones copulativas en español: los adjetivos de edad y tamaño', in Hernández Hernández, E. and P. Martín Butragueño (eds.), *Elementos de variación y diversidad lingüística: qué es una palabra*, CSIC/COLMEX.
- Gutiérrez, M. J. (1994). *Ser y estar en el habla de Michoacán*, Publicaciones del Centro de Lingüística Hispánica, 38, México, UNAM/Instituto de Investigaciones Filológicas.
- Moreno-Quibén, N. (2022). 'Estructuras de predicación y alternancia de cópula', Doctoral Dissertation, Universidad de Alcalá.
- Ortiz López, L.A. (2000). 'Extensión de estar en contextos de ser en el español de Puerto Rico: ¿evaluación interna y/o contacto de lengua?', *BAPLE*, pp. 99-118.
- Silva-Corvalán, C. (1986). 'Bilingualism and language change: The extension of estar in Los Angeles Spanish', *Language*, 62, pp. 587-608.

---

## Variable Realizations of Coronal Plosives in a French Urban Youth Vernacular: A Case Study

Scarlet Peterson

University of Illinois at Urbana-Champaign, Urbana, USA

In this paper, I present the early stages of a socio-phonetic analysis of variable realizations of coronal plosives /t/, /d/ preceding /i/, /y/ in a corpus of Parisian French urban youth vernacular [1] [2]. In this variety, plosives present realizations ranging from frication to palatalization and affrication [3] [4]. Previous studies also indicate that these realizations can index characteristics such as speaker age, socioeconomic status, national origins, and cultural background [5] [6], while others assert that anatomically conditioned physiological factors, rather than social influence, might be at play [3]. Guided by these hypotheses, I present a case study building on third-wave models of socio-phonetic variation [7] [8]: it explores the phonetic and discursive aspects of social indexicality as “the dynamic result of the self-expression and self-differentiation” [9] of adolescent identities in multicultural peer group interactions.

Twenty-five instances of /t/ and /d/ followed by /i/ and /y/ were selected from the speech of three male adolescents recorded in the same peer group conversation with the fieldworker (Table 1). Each speaker had the highest possible network score, indicating a diverse peer circle [2]. The duration and the Center of Gravity (CoG) of each consonant release were measured in Praat [10], with CoG used as indicative of place of articulation [11]. Analyses of Variance (ANOVA) and post-hoc t-tests showed significant differences between the three speakers' duration and CoG values (Table 2). Post-hoc analyses showed that SP41 and SP44 patterned together. Their average release durations were longer and CoG values were lower, suggestive of more complex and more posterior articulation than for SP39 who patterned separately from them both. Preliminary auditory analyses confirmed more affricated and palatalized realizations of /t/ and /d/ in the former two speakers' speech, while fricative-type releases were prevalent in SP39's speech.

Despite similarities in speech, SP44's and SP41's social positionings were diametrically opposed. SP41, a monolingual youth of European immigrant descent, positioned himself as affluent, contrasting himself with his peers and echoing class-based cultural tensions between Paris and its working-class suburbs: “hey, it's not my fault if they're poor; I paid for that trip”. SP44 and SP39, fluent speakers of Algerian Arabic, spoke for their banlieue with pride and solidarity, with SP44 leading and SP39 supporting the exchange (“we in the banlieue, in fact, we're the good guys and they [the 'Parisians'] are the victims”...“I swear, he's right!”) [8]. When SP41 attempted to

comment, he was systematically interrupted by SP39. In this peer discussion group, SP44 demonstrated authority on matters of adolescent banlieue culture, SP39 was supportive, while SP41's views were sidelined.

Preliminary conclusions suggest that, rather than macro-social or speaker-specific influence, micro-phonetic variation in these three speakers' pronunciation is best interpreted as self-expression put in the service of dynamic processes of negotiation and identity construction. As these adolescents navigate the variation within the wide, linguistic diversity of their multiethnic community, they both blend in with, and distinguish themselves from, each other and their community in their own unique ways.

- [1] Gardner-Chloros, P., Cheshire, J., and Secova, M. (2014). Multicultural Paris French (MPF) Corpus, <http://www.mle-mpf.bbk.ac.uk/Data.html>.
- [2] Gardner-Chloros, P., & Secova, M. (2018). Grammatical change in Paris French: in situ question words in embedded contexts. *Journal of French Language Studies*, 28(2), 181–207. <https://doi.org/10.1017/s0959269518000091>.
- [3] Berns, J. (2013). Friction between phonetics and phonology: The status of affricates. *LOT*.
- [4] Devilla, L., & Trimaille, C. (2010). Variantes palatalisées/affriquées en français hexagonal: Quel(s) statut(s) sociolinguistique(s) pour quel destin? In M. Iliescu, H. Siller-Runggaldier, & P. Danler (Eds.), *XXVe CILPR Congrès International de Linguistique et de Philologie Romanes* (pp. 4-99–4–108). De Gruyter. <https://doi.org/10.1515/9783110231922.4-99>
- [5] Trimaille, C., Candea, M., & Lehka-Lemarchand, I. (2012). Existe-t-il une signification sociale stable et univoque de la palatalisation/affrication en français ? Étude sur la perception de variantes non standard. *SHS Web of Conferences*, 1, 2249–2262. <https://doi.org/10.1051/shsconf/20120100122>
- [6] Spini, M., & Trimaille, C. (2017). Les significations sociales de la palatalisation/affrication à Marseille: Processus ségrégatifs et changement linguistique: *Langage et Société*, N° 162(4), 53–78. <https://doi.org/10.3917/l.s.162.0053>
- [7] Eckert, P. (2012). “Three waves of variation study. The emergence of meaning in the study of sociolinguistic variation”, *Annual Review of Anthropology* 41, 87–100.
- [8] Lew-Hall, L., Moore, E., & Podesva, R. (eds.) (2021). *Social meaning and linguistic variation: theorizing the third wave*. Cambridge, Cambridge University Press.
- [9] Fagyal, Z. & Davidson, J. (2021). Sociophonetics. In Gabriel, C., Randall, G., and Meisenburg, T. (eds). *Manual of Romance phonetics and phonology*, (eds), Berlin, New York: De Gruyter, 343-374.
- [10] Boersma, Paul & Weenink, David (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.10, retrieved 3 May 2023 from <http://www.praat.org/>
- [11] Gordon, M., Barthmaier, P., & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32(2), 141–174. <https://doi.org/10.1017/S0025100302001020>

---

## (Heritage) Russian case-marking: Variation and paths of change

Julia Petrosov, Naomi Nagy  
University of Toronto, Toronto, Canada

Russian has a complex morphological marking system, with six cases and 87 noun classes, making case-marking potentially challenging ground for heritage speakers. Indeed, morphological levelling, “probably the best described feature of language loss,” has been substantiated experimentally. Polinsky (2006:250) showed that, in an experimental task, Heritage Russian speakers living in the USA produced canonical case markers for only 13% of prepositional oblique nominals (instrumental, prepositional, dative, genitive, or accusative case marking on a noun in a prepositional phrase). In contrast, Łyskawa & Nagy (2020) found 94% use of canonical case markers in conversational speech from Heritage Russian speakers in Canada. Why this difference?

Nagy (2015) outlines several possible accounts for the stark difference in these outcomes. The current paper makes three important contributions to this task:

1. We circumscribe the context to oblique nominals in Heritage Russian speech from the same Toronto corpus, mirroring the domain of investigation reported by Polinsky.
2. We compare oblique nominals produced by Homeland Russian speakers (residing in Moscow or St. Petersburg) in conversational contexts, using data from the Russian National Corpus (Institute of Russian Language, Russian Academy of Sciences, 2003). This allows us to compare the degree of variability in monolingual vs. multilingual speakers.
3. We examine the roles of lexical frequency, disentangling whether differences in morphological marking are tied to heritage speakers' lexicon size.

~100 (non-nominative) NPs were extracted from conversational speech of 26 heritage and 6 homeland [NN1] speakers and coded for match between prescribed and produced case (“match rate”), syntactic environment and lexical frequency. Predictions were tested that heritage speakers produce canonical case marking more often in more frequent nouns and in pronouns compared to nouns. Mixed Effects Models of Heritage and Homeland Russian compared the effects of linguistic factors across groups, finding similarity in the conditioning of variability for Heritage and Homeland speakers.



Small intergenerational differences in rate of use of canonical forms (see Figure 1) suggests some morphological levelling, evident through the higher rate of mismatch on nouns by speakers of second and third generations, but less than reported in Polinsky (2006). Yet, the 39 lexical classes used by these speakers cover ~97% of Russian nouns, alleviating concern that heritage speakers suffer from a reduced vocabulary or select only “easy” nouns in conversation (cf. Pechkina & Nagy, 2017). However, the data revealed an interesting trend: the later the generation since immigration, the more speakers’ mismatches are restricted to rarer lexical items (see Figure 2).

Comparing multiple generations of heritage speakers allows us to trace more precisely the way that lexical structure interacts with ongoing changes in a little-described variety of Russian.

- Institute of Russian Language, Russian Academy of Sciences. (2003). Russian National Corpus. <https://ruscorpora.ru/>
- Lyskawa, P., & Nagy, N. (2020). Case marking variation in heritage Slavic languages in Toronto: Not so different. *Language Learning*, 70(SI), 122–156. <https://doi.org/10.1111/lang.12348>
- Nagy, N. (2015). A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua*, 164, 309–327. <https://doi.org/10.1016/j.lingua.2014.04.012>
- Pechkina, A., & Nagy, N. (2017). Heritage Russian speakers’ use of noun cases in relation to frequency and entropy of noun class [ROP 299].
- Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, 14, 191–262.

---

## “Mission Diyari”: polysemy & contact-induced semantic change

Josh Phillips<sup>1,2</sup>, Jack Sullivan<sup>1</sup>, Will Wegner<sup>1</sup>

<sup>1</sup>Yale University, New Haven, USA. <sup>2</sup>Stanford University, Stanford, USA

In Diyari (Karnic: Central Australia), the lexical item *marla* functions as an adjectival intensifier, roughly translatable as ‘very, truly’ (see 1). This function is also attested for cognate items in various related languages (e.g. Arabana *arla*). *marla* additionally sees use as a comparative morpheme (2) — a grammatical category all but unattested in Australian Languages — as well as an cessative adverb, realising an aspectual function that corresponds to English ‘anymore’ under the scope of a negative operator (3). The colexification of these three uses is—so far as we know—unattested elsewhere on the continent (in fact, a recent survey has shown that negative polarity items are all but unattested in Australian Languages.) In this paper, we propose a sociohistorically-grounded account of the trajectories of semantic change that have given rise to the synchronic polysemy that is exhibited by Diyari *marla*. Taking these results as a case study, we consider various methodological desiderata in diagnosing contact-induced semantic change.

**CONTACT.** In the mid-19th century, German Lutherans established a mission and school in Diyari country. Diyari was selected as the language of their ministry given that it was understood by—and served as a *lingua franca* for—a number of different tribal groups in the area (Hoffman 2008). As a consequence, vernacular literacy among Diyari speakers was significant and the language possesses a large corpus of written materials from the 19th and 20th centuries, including Bible translations, dictionaries and letters. For Kneebone (2005: 7), “‘Mission languages’ are characterised by structural standardisation ... The functional range of such languages is engineered and restricted according to the aims of the mission”. The deliberate construction of “Mission Diyari” thus produced a new language variety, functionally and lexically distinct from its pre-contact form. At various places, particular structures appear to have been developed by missionaries as solutions for particular translational problems (Stockigt 2016; Moore 2019).

**SEMANTIC CHANGE.** We argue that sustained contact between indigenous communities and European settlers and the concomitant shifts in usage contexts precipitated significant restructuring of Diyari grammar. Consequently, we propose an account of the recruitment and reanalysis of *marla* across various stages which has given rise to the synchronic multifunctionality exemplified in (1–3). This extension in *marla*’s distribution (“polysemy copying”) can partially be understood within the framework of replica grammaticalization (in the sense of Heine & Kuteva 2012, see also Epps & Law 2019 a.m.o.) Given the attestation of numerous close relatives in addition to the existence of a diachronic corpus, Diyari provides a unique opportunity within Australian linguistics to assess the profound effects of language contact on the grammars of these languages. The specific phenomenon reported here also provides a case study of the effects of a colonial language contact scenario where instigators of a semantic change were in fact L2 speakers, apparently imposing changes on the language in which they were missionizing.

- Austin, P. K. (2011). *A grammar of Diyari*, South Australia. SOAS University of London, London, 2nd edition.
- Epps, P. and Law, D. (2019). Contact-induced semantic change. In Darquennes, J., Salmans, J. C., and Vandenbussche, W., editors, *Language Contact: An international handbook*, vol. 1, pages 38–52. De Gruyter, Berlin.
- Heine, B. and Kuteva, T. (2012). *Language contact and Grammatical change*. Cambridge University Press, Cambridge.
- Hoffman, D. (2008). *Language Contact in Australia*. Master’s thesis, Universität Konstanz.
- Kneebone, H.-M. (2005). *The Language of the Chosen View: the first stage of graphization of Dieri by Hermannsburg missionaries, Lake Killalpaninna 1867-80*. PhD thesis, University of Adelaide, Adelaide, SA.

- Moore, D. (2019). German Lutheran Missionaries and the linguistic description of Central Australian languages 1890-1910. PhD thesis, University of Western Australia, Perth, WA.
- Schweiger, F. (1984). Comparative: A neglected category in Australian languages? Technical Report 18, SIT, Launceston, Tas.
- Stockigt, C. (2016). Pama-Nyungan morphosyntax: Lineages of early description. PhD thesis, University of Adelaide, SA.

---

## An ocean apart: Pronoun Choices in Turkish Heritage Speakers

Ilaria Porru

This study aims to explore pronoun interpretation and preference between Turkish heritage populations in Germany and in the United States. This comparative approach is valuable due to the different linguistic environments of these groups: in Germany, Turkish is prevalent in public life and media whereas, in the U.S., it mainly exists within familial settings (Iefremenko et al. 2021). We investigate whether the difference in active L1 use modulates the degree of cross-linguistic influence of the societally dominant language.

With regard to third person pronouns, Turkish (*o*), German (*er/sie/es*) and English (*he/she/they*) obey principle B and cannot take local antecedents. However Turkish, unlike German and English, allows pronouns to be omitted. Turkish also allows the inflected reflexive pronoun *kendisi* as a subject, used interchangeably with *o* (Çınar & Çakır, 2019; Enç, 1986; Özsoy, 1987, Gračanin-Yüksek et al., 2020, Rudnev, 2011). This flexibility in pronoun usage sets Turkish apart from German and English.

According to some researchers (e.g., Sorace and Filiaci, 2006; Haznedar, 2010), bilingual speakers of pro-drop languages tend to overuse overt pronouns compared to their monolingual peers. Turkish-German bilinguals have not consistently exhibited this pattern (Gračanin-Yüksek, 2019; Azar et al., 2020).

This study's research questions are the following:

1. Do different Turkish heritage speaker populations use overt and null pronouns differently?

To answer this question, participants will be presented with sentences containing either *o* or *pro* and will identify potential referents for the pronouns. We expect participants with more active use of Turkish to interpret *o* as indicating topic shift, while speakers with higher use of the societally dominant languages may exhibit a more heterogeneous interpretation.

2. In heritage speakers' pronominal system, does the presence of the reflexive *kendisi* signal topic shift?

Based on the fact that the reflexive pronoun *kendisi* can be used in free variation with *o*, we might expect that it also signals topic shift. We will test this with an interpretation task which will show whether participants assign topic shift to the reflexive pronoun *kendisi* to the same extent as they do with *o*.

3. Are there differences in the use rate of overt pronouns between heritage speakers in Germany and in the US?

We will assess this through a preference task which will ask participants to rate sentences with omitted and overt pronouns.

4. Is there a difference between the pronominal systems of heritage speakers in Germany compared to heritage speakers in the United States?

5. Does active use of the heritage language influence pronoun interpretation?

At the end of the experimental task, we will administer the Heritage Language Experience (HeLEx) online questionnaire (Tomić et al. (2023), which focuses on active use of their spoken languages.

In summary, this study focuses on pronoun interpretation and use in Turkish heritage speakers in two distinct sociolinguistic contexts and with different levels of active use. The goal is to contribute to a deeper understanding of how active use may shape pronoun preferences in heritage bilingual populations.

---

## The surprising pathways of change: Lessons from the vernacular

Shana Poplack, Nathalie Dion  
University of Ottawa, Ottawa, Canada

The gradual ousting of one of two or more expressions of the “same thing” is the most straightforward and intuitive scenario for change. But resulting shifts in rates of competing variants may stem from other sources as well. Even where they do signal change, rarely is it confined to the rise or decline of a variant, and even more rarely will it go to completion, especially in the time spans typically available to linguistic inquiry. These facts highlight the importance of focusing on the transition period. What transpires in this interval? How do changes spread across grammatical sectors, and what does their trajectory reveal about the productivity, dispersion and potential for survival of the variants involved?

Drawing on the Comparative Variationist framework (Poplack & Meechan 1998; Tagliamonte 2002), we address these questions through analysis of three corpora of spoken Quebec French, which together cover an apparent-time

span of a century and half, offering a virtually unprecedented amount of time to trace the existence and progression of slow-moving changes in speech. Quebec French lends itself particularly well to this endeavour, since it is widely assumed to have changed profoundly.

Sustained analyses of a number of morphosyntactic variables in these materials did in fact turn up some spectacular changes in the distribution of alternating variants, but the configuration of environmental factors affecting their selection is more revealing. Such conditioning can be construed as the grammar of the variability, and its varying dispositions may be marshalled at different points in time to identify and classify types and pathways of change. This exercise uncovered remarkable stability at the core grammatical level. But a recent deep dive into these competing trajectories revealed alterations of many more subtle types that throw the standard enterprise of relying on rates to infer change into doubt. Among the most potentially misleading are rising rates with loss of productivity, stable rates with change of function, and mismatches between rates and conditioning. Importantly, only some of these qualify as structural changes, others are more accurately viewed as changes in the “textual habitat” (Szmrecsanyi 2016). Indeed, the vagaries of the linguistic contexts hosting the variants (whether receding, expanding, or serving as “last bastions”) play a crucial role in these and other previously undocumented developments. The same is true of extra-linguistic contexts, since some variants may be totally absent from the speech of some community members, while increasing or waning in that of others.

In this paper we review these developments, largely invisible to any but quantitative variationist analysis over the longue durée, and propose a more holistic approach to the identification and assessment of change. This would incorporate not only rates, but also conditioning and context, while simultaneously taking account of productivity and dispersion. Only once the various types of change have been apprehended in the stream of spontaneous speech and accurately identified, can we begin to address questions about where they occur (if at all), when and under what conditions.

Poplack, S. & Meechan, M. 1998. How languages fit together in codemixing. *International Journal of Bilingualism* 2(2): 127-138.

Szmrecsanyi, B. 2016. About text frequencies in historical linguistics: Disentangling environmental and grammatical change. *Corpus Linguistics and Linguistic Theory* 12(1): 153-171.

Tagliamonte, S. 2002. Comparative sociolinguistics. In J. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *Handbook of Language Variation and Change*, 729-763. Malden/Oxford: Blackwell Publishers.

---

## Codeswitching variation in Chanka Quechua: Palindromic switches

Natalie Povilonis

Indiana University, Bloomington, USA

This paper analyzes Quechua-Spanish codeswitching in Andahuaylas, Peru, with the goal of legitimizing its use by contemporary speakers. While Quechua has become more socially acceptable in recent years, speakers maintain a strong ideology valuing ‘pure’ Quechua ([citation redacted]). Codeswitching (with Spanish) is stigmatized and rejected as ‘contaminated’ Quechua, despite its prevalence among bilinguals and Quechua-dominant alike. This delegitimization of variable Quechua discourages bilingual speakers from maintaining Quechua, accelerating community shift to Spanish.

Aligned with this study’s goals, variationist approaches have demonstrated the systematic nature of codeswitching in numerous language pairs (Poplack 2018). Specifically, this study considers Quechua-Spanish palindromic switches, also called mirror-image constructions, which have been documented for Finnish-English (Poplack et al. 1989), Tamil-English (Sankoff et al. 1990), and other language pairs. In palindromic switches, speakers use both the Spanish preposition and the equivalent Quechua suffix in the same mainly Quechua phrase. For example, in ‘suyawarqanku hasta chayaranay-kama’ “they waited for me until I arrived-until”, hasta in Spanish introduces the subordinate phrase, which is closed by the Chanka -kama suffix with a similar meaning (all words except hasta are Quechua). This study considered only phrases with Quechua -kama and occasionally the Spanish counterpart hasta, though this type of duplication occurs with a range of Spanish preposition-Quechua suffix pairs of similar meaning. As the Quechua morpheme is always present, this codeswitching does not necessarily indicate a proficiency gap, notwithstanding popular belief to the contrary.

Data in this study included 300 sentence tokens from sociolinguistic interviews with 46 speakers (21 monolinguals; 25 bilinguals) in Chanka Quechua. Use of palindromic switches was found to be variable, with absence of codeswitching (the Spanish preposition) the preferred variant: the population overall maintains 73.3% Quechua-only phrases. At the same time, the effect of Spanish proficiency is clear: only fluent bilinguals exhibit palindromic switches. The only four bilinguals without switches had prescriptivist metalinguistic training in Quechua, which may have led them to avoid codeswitching. The remaining 21 bilinguals had at least one palindromic switch, up to rates of 100%. Conversely, eighteen of 21 almost-monolingual Quechua speakers had zero Spanish switches. The remaining three of the Quechua-dominant speakers had one palindromic switch token each,

but this coincided with a longer borrowed phrase—when the object of the preposition was also borrowed from Spanish.

These results indicate a contact effect of Spanish morphosyntax on Quechua usage: speakers with increased Spanish proficiency systematically show more Spanish preposition switches, though they never replace the Quechua affix. While the data set needs to be expanded to find results of statistical significance, this work contradicts perceptions that Quechua is monolithic, without Spanish influence. It shows the vibrant range of ways that bilingual speakers utilize their prepositional repertoires. Results thus facilitate the creation of accurate pedagogical materials that recognize this linguistic diversity. This will encourage speakers to maintain their unique Quechua, including occasional codeswitching, rather than switching completely to Spanish.

Poplack, S. (2017). *Borrowing: Loanwords in the Speech Community and in the Grammar*. Oxford University Press.

Poplack, S., Wheeler, S., & Westwood, A. (1989). Distinguishing language contact phenomena: evidence from Finnish-English bilingualism. *World Englishes*, 8(3), 389-406.

Sankoff, D., Poplack, S., & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language variation and change*, 2(1), 71-101.

---

## Old wine in new bottles: Tokyo vowel devoicing

Dennis Preston<sup>1</sup>, Terumi Imai<sup>2</sup>

<sup>1</sup>University of Kentucky, Lexington KY, USA. <sup>2</sup>National Institute for Language and Linguistics (NINJAL), Tachikawa, Japan

Author (2004) is a 30,000 token GoldVarb study of 42 Tokyo respondents; the data were stratified by style, status, age, gender, and linguistic variables. The major findings confirmed and challenged earlier studied linguistic factors and established base-line social ones:

- 1) The conflict over devoicing frequency was resolved in favor of (i) over (u).
- 2) Affricates & fricatives are separate influences in pre- & postvocalic positions.
- 3) Devoicing is not prohibited between two fricatives.
- 4) Allophone vs phoneme consonant identity affects devoicing rate.
- 5) Bound morpheme boundaries as well as no boundary promote devoicing.
- 6) Low-pitch and final position pitch-accent patterns promote devoicing.
- 7) Previous opportunities for devoicing as well as actually devoiced ones demote devoicing.
- 8) The stylistic frequency order is conversation > reading passage > word list.
- 9) Younger males devoice most and younger females least, but there are no significant gender differences for older and middle-aged groups.
- 10) There were no significant social class differences.

These data are reinterpreted in mixed-effects, conditional inference tree, and random forest procedures (Tagliamonte & Baayen 2012) and suggest the following additions:

- 1) Sibilant and non-sibilant fricatives are distinct in devoicing rates.
- 2) Preceding affricates and following affricates pattern with fricatives for the former and stops for the latter, suggesting a phonetic rather than phonemic interpretation.
- 3) The association of low pitch and devoicing is not strongly promoting.
- 4) There is an interaction of gender, age, and style: young men devoice more and young women less but both in more conservative (word list & reading passage) styles.
- 5) Random effects of respondents was significant but reflected the demographic profile of the individual respondents with new actual scores for status and age rather than the previous tripartite groupings.

This last is a sociolinguistic puzzle: Why would young men increase use of a standard variable and young women reduce it in formal environments? Perhaps younger age groups are now so distant from the concerns for standard devoicing that they can select and rework it, assigning age and gender social meaning. The only statistical hint of this lies in the conditional inference tree identification of higher status older speaker increased devoicing, but only for (u).

Whatever that interpretation, the question of age grading versus change remains. The newer mixed model format allows careful investigation of marginal means, ones that show trends in age, gender, and style supportive of change. Recently, real time additions have been made to the apparent time work. Respondents classified as "Young" now "Middle-Aged" and previously "Middle-Aged" now "Older" have been re-interviewed, and analyses of their data show no rate changes. When panel-study respondents cannot be found, others who match the "Young" age of the original now in the "Middle" category and previously "Middle-Aged" now "Older" have been identified and are being interviewed; data are also being collected from the new "Young" group. So far, the likelihood of a linguistic change with surprising new age and gender social meanings assigned to an old standard variable seems supported.

- Imai, Terumi. 2004. (National Institute for Japanese Language and Linguistics), Vowel devoicing in Tokyo Japanese: A variationist approach. Doctoral dissertation: Michigan State University. East Lansing MI USA.
- Tagliamonte, Sali A. and Harald Baayen. 2012. Models, forests, and trees of York English: was/were variation as a case study for statistical practice. *Language Variation and Change* 24,2:135-176.

---

## Are there really Pirates in Beijing? Perceptual Dialectology of Rhoticization in Mandarin

Travis Putman, Brianna O'Boyle  
San Diego State University, San Diego,

Many varieties of Mandarin exhibit a process of rhotacization in the rime of a syllable. Hu (2022) analyzes this process as the suffixation of a sound *er* (conventionalized pinyin) to a rime. While this process exists in many varieties of Mandarin, such as Standard Mandarin, Chengdu, Jilu, among others, it is more frequently used in and associated with the Beijing dialect. As some of these dialect regions spread into the North(east) of China like Beijing and Jilu and there is an established association of rhotacization with the Beijing dialect, we wish to investigate the extent of the regional association of this feature further. We ask whether there exists a salient North-South regional distinction based on the feature of rhoticity.

To answer this question, we follow Plichta & Preston (2005) and replicate a similar methodology to test the regional association of this rhoticity feature along a North-South latitude. We will present native Mandarin speakers, located in the United States and China, with varying degrees of rhotacization in the rime of one Mandarin word spoken by male and female speakers from Beijing and Jinan. Using speakers' recordings taken from the ManDi corpus (Zhao & Chodroff, 2022), we truncate the rhotic segment at two points: before the rhotic segment and halfway through the rhotic segment; therefore, resulting in a three-step continuum of rhoticity, ranging from non-rhotic to fully rhotic. The participants will listen to recordings of the speakers, at each step of rhoticity, three times in a random order; thus, there are 18 trials total in the test block of the survey. After listening to an utterance they will place the location of a speaker on a map with five cities, Beijing, Jinan, Nanjing, Wenzhou, and Xiamen, which range along a North-South dimension.

Demographic information such as gender, region, and age combined with participants' responses to the survey will be analyzed to investigate which cities are commonly associated with rhoticity. If there is an association with rhoticity and Northern cities, we expect that participants will choose Jinan and Beijing, rather than southern cities, in conditions where recordings have longer rhotic segments. If there does not exist an association of the northern region as a whole, we expect rhotacization to only increase associations with Beijing, in line with previous research. We hypothesize that the former will hold true and that in recordings with a longer duration of rhoticity, we will see association with both northern cities, Beijing and Jinan, increase. As for correlation with demographic information, we intend to use mixed effects linear regression to assess the factors of gender, age, and participant's dialect region on the perception of a speaker's dialect region.

If the results are as expected, this would provide evidence for the case that there is a distinct perceptual difference between northern and southern varieties of Mandarin. We expect participant gender and age to have no statistically significant effect on responses following Plichta & Preston (2005), however, we do expect participants' dialect region to have an effect on perception.

Hu, H. (2022). *The Sociolinguistics of Rhotacization in the Beijing Speech Community* [Doctoral Dissertation].

Plichta, B., & Preston, D. R. (2005). The /ay/s Have It The Perception of /ay/ as a North-South Stereotype in Uninvited States English. *Acta Linguistica Hafniensia: International Journal of Linguistics*, 37(1). 107-130.

Zhao, L., & Chodroff, E. (2022). The ManDi Corpus: A Spoken Corpus of Mandarin Regional Dialects. *Proceedings of the 13th Conference on Language Resources and Evaluation*. 1985-1990.

---

## The intersection of ethnicity and social class in language variation and change

Gan Qiao, Catherine Travis  
Australian National University, Canberra, Australia

Ethnicity has long captured the attention of sociolinguists, with extensive consideration of the ways in which ethnic minorities respond to patterns of variation and change in the wider community, as well as how they use language to mark ethnic orientation (e.g., Labov 2001; Hoffman and Walker 2010). In such work, ethnicity is typically considered independently of other social factors. In this talk, we demonstrate a fundamental intersection between ethnicity and social class, and argue that patterns of ethnic variation are illuminated by taking account of class affiliations of ethnic groups.

The evidence comes from quantitative patterns of variation in a corpus that brings together sociolinguistic interviews with 230 Australians recorded in Sydney at two time points, in the 1970s and the 2010s. The speakers

represent some of Australia’s largest ethnic groups—Anglo-Celtic, Italian, Greek and Chinese Australians—and are stratified according to age, gender and social class. Transcriptions totaling some 1.25 million words provide large amounts of data for linguistic analysis, allowing for a testing of ethnicity over time and in relation to other social characteristics.

Previous work on this same dataset has found that ethnic groups are differentiated socially, with, for example, Greek Australian teenagers in the 1970s tending to produce more “cultivated” vowels (Horvath 1985), and Chinese Australian young adults in the 2010s also tending to produce vowel realisations that are closest to those of women and middle-class Australians (Grama, Travis and Gonzalez 2021). Here, we test the broader applicability of this differentiation by asking not how different social groups pattern linguistically, but how speakers with similar linguistic behaviour are grouped socially (cf., Horvath and Sankoff 1987), based on a set of variables drawn from phonetic, morphosyntactic and discourse levels (FLEECE and FACE vowels, prevocalic *the*, *-ing*, word-final *-er*, existential *there*’s, modals of obligation, quotatives).

We conduct independent regression analyses for each variable, analysing the linguistic predictors only, and extract the speaker random intercepts from each model, as a metric against which to meaningfully compare across speakers (cf., Drager and Hay 2012). Although some ethnic differences are observed, these rarely correlate with measures of ties to ethnic heritage, such as social network, language proficiency, or ethnic orientation. To compare the social makeup of groups who pattern similarly linguistically, we conduct divisive cluster analysis of the random intercepts from each model (following the methodology outlined in Haddican et al. 2021). This reveals clear age clustering for both time periods, with ethnic groups clustering with their age cohort. We further find that Greek Australians in the 1970s and Chinese Australians in the 2010s tend to cluster more with their middle-class Anglo counterparts, while the same is not so for Italian Australians (see Figure 1 for Anglo, Chinese and Italian Australians, 2010s). This indicates that ethnic minorities may use variation patterns to mark, not their ethnic heritage, but their affiliations to the broader community, as a way of “sounding Australian” (Horvath 1985: 176), and highlights the need to interpret ethnic differences in intersection with other social variables.

- Drager, Katie and Jennifer Hay. 2012. Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change* 24(1): 59-78. <https://doi.org/10.1017/S0954394512000014>
- Grama, James, Catherine E. Travis and Simon Gonzalez. 2021. Ethnic variation in real time: Change in Australian English diphthongs. In Hans Van de Velde, Nanna Haug Hilton, and Remco Knooihuizen (eds), *Studies in Language Variation*, 292-314. Amsterdam: John Benjamins. <https://www.jbe-platform.com/content/books/9789027259820-silv.25.13gra>
- Haddican, Bill, Michael Newman, Cecelia Cutler and Christina Tortora. 2021. Aspects of change in New York City English short-a. *Language Variation and Change* 33(2): 135-163. <https://doi.org/10.1017/S0954394521000120>
- Hoffman, Michol F. and James A. Walker. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22(1): 37-67. <https://doi.org/10.1017/S0954394509990238>
- Horvath, Barbara. 1985. *Variation in Australian English: The sociolects of Sydney*. Cambridge: Cambridge University Press.
- Horvath, Barbara and David Sankoff. 1987. Delimiting the Sydney speech community. *Language in Society* 16(2): 179-204. <https://www.jstor.org/stable/4167835>
- Labov, William. 2001. *Principles of linguistic change: Social factors*, vol. 2, 3 vols. Oxford: Blackwell.

---

## The impact of social information on VOT shadowing by nonbinary speakers

Jack Rechsteiner

Michigan State University, East Lansing, USA

Social information can impact the degree to which one speaker phonetically converges with another speaker (Babel 2012, Gasiorek et al. 2015). There is also evidence that nonbinary speakers alter their speech due to social factors, including the threat of being misgendered (Gratton 2016). The present study builds upon these findings by analyzing 45 nonbinary American English speakers in an online phonetic shadowing task. We investigate the impact of social information on nonbinary speakers’ convergence toward extended voice onset time (VOT) in word-initial English /p, t, k/. We hypothesized that participants exposed to a model speaker stated as having a nonbinary identity would show higher rates of convergence than participants exposed to a model speaker who was not stated as being nonbinary.

Extended VOT was examined because it has been documented as producing convergence, it is not an explicit stereotype of gender, and it has no phonological perception consequences for voiceless stops in English (Shockley 2004, Nielsen 2011, Schertz 2021). The experiment was a shadowing input-driven elicitation task where words were presented to participants one at a time throughout three phases and participants recorded themselves speaking the word aloud. Phase 1 (Baseline) elicited participants’ baseline productions by presenting written instructions and words without auditory exposure. Phase 2 (Exposure) presented participants with auditory instructions and words read aloud by the model speaker. Phase 3 (Post-exposure) presented participants with written words without accompanying audio. In the Exposure phase, participants were given auditory instructions from one of three

conditions: the model speaker explicitly identifies themselves as nonbinary (Nonbinary Condition), the model speaker does not give information about their gender (Neutral Condition), the model speaker explicitly identifies themselves as cis (Cis Condition). The model speaker was the same across all conditions, and recordings were identical in all conditions except for the gender identity information given. Participants were distributed evenly across conditions with 15 participants per condition.

Results were analyzed using a linear mixed-effects model with VOT as the dependent variable, an interaction term between fixed effects of Experiment Phase and Condition, a fixed effect for initial stop, and random intercepts for speaker and word. VOT values are shown in Figure 1, and the output of the statistical model is shown in Table 1. Unexpectedly, all conditions saw participant VOT values diverging from the model talker during Exposure compared to their Baseline, shown in Figure 1. Crucially, however, the results found that participants diverged the most in the Cis Condition ( $p < 0.001$ ,  $-15.23$  ms ( $-9.55$ - $5.68$ )) and the least in the Nonbinary Condition ( $p = 0.03$ ,  $-6.52$  ms ( $-9.55$ + $3.03$ )). While the hypothesis predicted that participants would show higher rates of convergence to a nonbinary model speaker, these patterns of divergence still suggest that nonbinary participants align their speech most closely to a model talker who is stated to be nonbinary. These findings agree with previous work which argues that nonbinary speakers pattern more like each other in queer contexts, suggesting that there is a nonbinary speech community (Gratton 2016, Rechsteiner & Sneller 2023).

- Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 40(1), 177–189. <https://doi.org/10.1016/j.wocn.2011.09.001>
- Gasiorek, J., Giles, H., & Soliz, J. (2015). Accommodating new vistas. *Language & Communication*, 41, 1–5. <https://doi.org/10.1016/j.langcom.2014.10.001>
- Gratton, C. (2016). "Resisting the Gender Binary: The Use of (ING) in the Construction of Non-binary Transgender Identities," *University of Pennsylvania Working Papers in Linguistics*, 22(2). Retrieved from <https://repository.upenn.edu/pwpl/vol22/iss2/7>
- Nielsen, K. (2011). Specificity and abstractness of VOT Imitation. *Journal of Phonetics*, 39(2), 132–142. <https://doi.org/10.1016/j.wocn.2010.12.007>
- Rechsteiner, J. & Sneller, B. (2023). "The Effects of Topic and Part of Speech on Nonbinary Speakers' Use of (ING)," *University of Pennsylvania Working Papers in Linguistics*, 29(1). Retrieved from <https://repository.upenn.edu/pwpl/vol29/iss1/19>
- Schertz, J., Johnson, E. K., & Paquette-Smith, M. (2021). The independent contribution of voice onset time to perceptual metrics of convergence. *JASA Express Letters*, 1(4), 045205. <https://doi.org/10.1121/10.0004373>
- Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception & Psychophysics*, 66(3), 422–429. <https://doi.org/10.3758/bf03194890>

---

### You best button [bʌrən] it up:

An American English change-in-progress with people identified as Black taking the lead

Chiara Repetti-Ludlow, Renee Blake  
New York University, New York, USA

In American English, the pronunciation of words like ‘button,’ with the underlying post-tonic string /tən/, is variable. Previous research has found that although the traditional pronunciation of these words is with [ʔn], as in [bʌʔn], there are also pronunciations with [ən], as in [bʌən], which could represent either a case of stable variation or a change-in-progress (Eddington & Brown 2021; Davidson et al. 2021). Repetti-Ludlow (forthcoming) found that the realization of /ən/ as [ən] on Long Island, New York is indeed a change-in-progress. Additionally, she found several instances of /t/ realized as [r], but only in the environment of [ən], resulting in the pronunciation [bʌrən]. The current study builds on this research by expanding the data pool from 60 participants from a predominantly white Long Island speech community to 108 participants of different regional and racial identity backgrounds.

For this work, we continued to utilize the online production study developed by Repetti-Ludlow using PCIBex. The study consisted of asking participants to complete semantic differential tasks and rapid word lists, which included target words and fillers. A phonetic analysis of the collected tokens (N=6857) was conducted to determine the realization of /t/ ([ʔ] or [r]) and the realization of /ən/ ([ən] or [n]). Finally, logistic mixed-effects regressions were carried out in R to determine significant predictors of these realizations.

Like Repetti-Ludlow, we find the only significant predictor for /t/ realization is how /ən/ is realized, with [n] generally correlated with [ʔ], and [ən] correlated with [r]. We also find that [ən] occurs more with more frequent words, and is produced more by younger participants. With the addition of speakers from different regions of the United States, we find speakers from the Northeast and South more likely to produce [ən] than speakers from the West, and Black-identified participants more likely to produce [ən] than non-Black participants.

The findings of this study have implications for both phonetics and sociolinguistics. Phonetically, the change in one realization (e.g., /ən/ > [ən]) leaves open the path for another change (e.g., /t/ > [r]). Sociolinguistically, the results suggest that the realization of /ən/ in words like ‘button’ is undergoing a change-in-progress, with younger

speakers producing [ən] more than older speakers. Furthermore, results suggest that this change is more advanced in the Northeast and South of the U.S., and critically, the social variable of racial identification is significant. Black-identified speakers appear to be leading in the change to [ən], coupled with [r]. These findings inform our understanding of how, and through whom, linguistic changes progress through time and across communities.

Davidson, Lisa, Shmico Orosco, and Sheng-Fu Wang. (2021). The link between syllabic nasals and glottal stops in American English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 12(1): 1-23.

Eddington, David & Earl Brown. (2021). A production and perception study of /t/-glottalization and oral releases following glottals in the United States. *American Speech* 96 (1): 78-104.

Repetti-Ludlow, Chiara. (forthcoming). Variation in post-tonic /tən/ in words like ‘button’: A change in progress on Long Island.

---

## Morphologically and topically conditioned [t] – [k] variation in Olelo Niihau

Travis Richardson

Georgetown University, Washington, DC, USA

Olelo Niihau is a dialect of the Hawaiian language spoken by approximately 500 speakers on Niihau. Niihau remains the only place on the Hawaiian archipelago where the language predominates over English for daily interactions (Wong, 2010). Niihau differs from the standardized variety of Hawaiian; one such difference involves the dialect’s retention of the Proto-Polynesian distinction between /t/ and /k/, which appear to be in complementary distribution in Niihau (Blust, 2004; Wong, 2010). /t/ is absent in the phonological inventory of standardized Hawaiian. The factors that condition the use of one variant over the other remain to be discovered. The present study addresses this gap in understanding by asking: What potential linguistic and social factors could condition the selection of [t] or [k]? This study is a quantitative analysis of the variation as exhibited by two native Niihau speakers. The data originate from a televised news segment from Hawaii News Now. Audio from the YouTube clip was imported to ELAN and transcribed in its entirety. Each token of /t/ and /k/ was tagged for the following linguistic and social variables hypothesized to affect variation: word class (lexical or functional), following phonological environment (front, central, or back vowel), speaker, and topic of conversation. The topic of conversation was coded as Niihau when the interviewees were discussing island life or the state of the Niihau dialect, and it was coded as Hawaii when the interviewees were discussing the general condition of language revitalization in Hawaii as a whole. While both interviewees vary in their use of [t] and [k], [t] is by far the predominant variant; the [k] variant comprised only 21% of the voiceless stop tokens. Multivariate regression analysis on the data indicates that neither the individual Niihau speaker nor the phonological environment in which the voiceless stop occurred predict the variation in a statistically significant way. This study offers two preliminary factors that condition the variable use of [k] over [t] in Niihau: lexical category and conversational topic. Niihau appears to have a subtly different determiner system from standard Hawaiian, as [t] is categorically used in Niihau for this word class. Minor variation with [k] exists for functional categories like conjunctions and prepositions. Reduplication also includes word-internal [t] – [k] variation (kupuatupuna, ‘great-grandparents’), providing nascent empirical support to Blust’s (2004) claim that multiple alveolar stops cannot appear within the same word. As for topically based variation, 75% of [k] tokens are uttered when the conversation concerns life in Hawaii more broadly, as opposed to more narrow discussions of Niihau and its dialect. The shift to [k] might be explained if one considers the potential for convergence (Bell, 1984) to an (imagined) audience of (assumedly) standard Hawaiian speakers. The results have implications for the language policy of revitalizing languages, as the Niihau speakers appeal to the retention of historical distinctions in their dialect as evidence for Niihau being a ‘purer’ form of the language compared to standardized Hawaiian.

Bell, A. (1984). Style as audience design. *Language in Society*, 13(2), 145-204.

Blust, R. (2004). \*t to k: An Austronesian sound change revisited. *Oceanic Linguistics*, 43(2), 365-410.

ELAN (Version 6.2) [Computer software]. (2021). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from <https://archive.mpi.nl/tla/elan>

Wong, A. K. (2010). Pukaiki kula maniania no Niihau, na ka Niihau [Doctoral dissertation, University of Hawai‘i at Mānoa].

---

## Co-occurrence and absence of feature: investigating combinatorial effects of stereotypical Southern French features on speech perception.

Célia Richy

Queen Mary University of London, London, United Kingdom

This paper contributes to work on co-occurrence of features and speech perception, examining combinatorial effects of Southern French features on perception of accented speech and formality.



Work on speech perception shows a bi-directional relationship between linguistic features and speakers. On the one hand, social information has been shown to influence the interpretation of linguistic features (e.g., Rubin 1992, Niedzielski 1999, D’Onofrio 2018). On the other hand, linguistic features have also been shown to index different social information (e.g., velar and alveolar variants for the variable (ING) in American English, Labov 2001, Campbell-Kibler 2007, 2009)). Variants themselves convey multiple meanings, depending on several parameters; what Eckert (2008) refers to as the ‘indexical fields’ of variables. Social meanings associated to linguistic features and particular variants are not fixed in production nor in perception, as they are context-dependent (see Hilton and Jeong 2019). Of particular importance here is work on linguistic coherence, co-variation, and speech perception. This research has shown that combinations of linguistic features in a speech signal can affect (or block) the social meaning of another specific feature in this signal (e.g., Levon 2014, Pharo & Maegaard 2017). It has also been found that features belonging to different registers and conveying distinct meanings in isolation can become harmonious when co-occurring (Grondelaers & van Hout 2016). That is, ‘dissonant’ clusters of features can co-occur without altering the perception of the signal. The present study builds on these past findings and explores the potential cumulative perception of features and the effect of an expected yet absent feature.

This study tests the effects of co-occurrence of Southern features on the perception of spoken Southern French. In particular, it investigates the perception of word-internal schwas in relation to their co-occurrence with different realisations of nasal vowel /ã/ ([ã] vs [ãŋ]). Whereas nasal vowels in standard French are realized as fully nasalized [ã̃], these vowels are stereotypically realized with a nasal coda [ã̃N] in Southern varieties. Word-final schwas (e.g., [laka] – laque) are ideologically associated with Southern French, in that final schwas are retained more frequently in Southern French compared to in non-Southern varieties. However, non-final schwas are indexically linked to both formality and Southern varieties (e.g., Durand et al. 1987, Pustka 2011). Therefore, this study looks at the different perceptions of word-internal schwa depending on its co-occurrence with different realizations of /ã/ (Southern [ãŋ] vs standard [ã], see table 1). Listeners’ perceptions are examined under the prism of perceptual harmony, compositionality, cumulation and combinativity of meanings. I also test whether the absence – rather than presence – of [ãŋ] in Southern French speech can impact speech perception, using a within- (linguistic conditions) and between-subject (social prime condition) design. Results show an incremental effect of schwa presence and [ãŋ] on the perception of accentedness, while the perceived degree of formality is affected by the presence/absence of schwa. I argue that, although the effect of feature absence remains unclear, this aspect of perception deserves more attention and investigation as it relates to the violation of listener expectations.

- Campbell-Kibler, K. (2007). Accent, ING, and the social logic of listener perceptions. *American Speech* 82(1). 32-64
- Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change* 21(1). 135-156
- D’Onofrio, A. (2018). Controlled and automatic perceptions of a sociolinguistic marker. *Language Variation and Change*, 30(2), 261-285
- Durand, J., Slater, C. & Wise, H. (1987). Observations on schwa in southern French. *Linguistics*, 25(5): 983–1004.
- Eckert, P. (2008). Variation and the indexical field. *Journal of sociolinguistics*, 12(4), 453-476.
- Grondelaers, S., & Van Hout, R. (2016). How (in) coherent can standard languages be? A perceptual perspective on co-variation. *Lingua*, 172, 62-71.
- Hilton, K. & Jeong, S. (2019). The role of context in sociolinguistic perception. *Linguistics Vanguard*, 5(s1), 20180069
- Labov, W. (2001). *Principles of Linguistic Change*. vol. 2, Social factors. Oxford: Blackwell
- Levon, E. (2014). Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society*, 43(5), 539-566.
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of language and social psychology*, 18(1), 62-85.
- Pharo, N., & Maegaard, M. (2017). On the influence of coronal sibilants and stops on the perception of social meanings in Copenhagen Danish. *Linguistics*, 55(5), 1141-1167.
- Pustka, E. (2011). L’accent méridional: représentations, attitudes et perceptions toulousaines et parisiennes. *Lengas. Revue de Sociolinguistique*, 69, 117–152
- Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates’ judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33(4):511–31

---

## Sociophonetic variation in Breton: Analyzing the effects of social factors, language contact, and speaker attitudes

Dakota Robinson  
University of California, Berkeley, Berkeley, USA

Breton, a Celtic language of northwestern France with an estimated 200,000 speakers, faces severe endangerment following massive language shift toward French since the early twentieth century (Broudic 2009). Language planning efforts to revitalize Breton have focused heavily on standardization with the intent of unifying the population under common norms, though an effect of this has often been a sentiment of division between traditional speakers who acquired Breton natively and new speakers who have learned the standardized variety

through immersion or bilingual schooling. One result of this generational divide has been a relative lack of sociolinguistic inquiry into variation within the rapidly shrinking population of native Breton speakers, in contrast with a more significant literature investigating the differences between traditional and new speakers (e.g., Davies-Deacon 2020, Hornsby & Quentel 2013, Kennard 2019), in spite of a relatively long tradition of regional dialectological research and some sociolinguistic literature that has highlighted the ways in which Breton-speaking communities do not actually behave as generational monoliths (Kennard 2019; Timm 2003). The current study aims to contribute to the effort to describe spoken Breton from a sociophonetic perspective by addressing the following research questions: (1) To what extent do social factors such as age and gender influence linguistic behavior, and (2) To what extent do speakers' attitudes toward the Breton language and its revitalization predict their use of more conservative phonetic variants?

The current study analyzes rhotic segments, which have been observed to vary widely within and between speakers, and for which little descriptive consensus exists in the literature. Productions vary between alveolar taps and trills, alveolar approximants, and uvular fricatives and trills (which are argued to be the result of language contact with French) (Press 1986).

Data come from the Brezhoneg war an Dachenn corpus, a database of recorded sociolinguistic interviews and translations from French to Breton by native speakers from diverse social and regional backgrounds (Blanchard & Thomas 2011). Rhotic segments will be extracted from the verbal translation task, in which the interviewer says 84 sentences one at a time in French for speakers to translate to Breton. Each token will be coded for place and manner of articulation based on auditory judgment and spectrogram and waveform analysis. Speaker age, sex, educational background, place of origin, and occupation will be included as potential predictors of rhotic production. Additionally, information provided during the sociolinguistic interview portion of each recording will be used for qualitative analyses of language attitudes, which may predict a speaker's use of more conservative variants, i.e., those that are not the uvulars shared by standard French.

In addition to offering a richer description of contemporary spoken Breton, results from this study may inform future studies of language change in the context of Breton and other minoritized languages. Existing models of language variation and change do not always account for patterns attested in endangered varieties (e.g., Kasstan 2019), and detailed sociolinguistic analyses such as the present study may also elucidate the effects of language planning strategies on language maintenance and revitalization.

- Blanchard, N. & Thomas, M. (2011). Brezhoneg war an dachenn: Enquête de dialectologie et sociolinguistique bretonnes. Bibliothèque numérique du Centre de recherche bretonne et celtique (CRBC), <https://bibnumcrbc.huma-num.fr/document/4700>.
- Broudic, F. (2009). Parler breton au XXI<sup>e</sup> siècle: Le nouveau sondage de TMO Régions. Emgleo Breiz.
- Davies-Deacon, M. (2020). New speaker language and identity: Practices and perceptions around Breton as a regional language of France (Unpublished doctoral dissertation). Queen's University Belfast.
- Hornsby, M. & G. Quentel. (2013). Contested varieties and competing authenticities: neologisms in revitalized Breton. *International Journal of the Sociology of Language*, 223, 71-86.
- Kasstan, J. (2019). Emergent sociolinguistic variation in severe language endangerment. *Language in Society*, 48(5), 685–720.
- Kennard, H. (2019). Morphosyntactic and morphophonological variation in Breton: a cross-generational perspective. *Journal of French Language Studies*, 29(2), 235–263.
- Press, I. (1986). *A grammar of modern Breton* (No. 2). Berlin: Mouton de Gruyter.
- Timm, L. A. (2003). Breton at a crossroads: Looking back, moving forward. *E-Keltoi: Journal of Interdisciplinary Celtic Studies*, 2 (1), 2.

---

## You Know I'm from the Chi: Identity construction, negotiation, and placemaking among Black Chicagoans

AC Rochelle Thomas  
Northwestern University, Evanston, IL, USA

This study concerns how Black Chicagoan speakers of African American English (AAE) employ various linguistic features in the process of shaping and performing their race and place identities, while also accounting for the metalinguistic beliefs and language ideologies that speakers hold about said facets of identity. Current research trends in sociolinguistics have come to a nuanced understanding of identity and performance as ongoing processes rather than set features or static characteristics, and there are several levels when a speaker projects their identity during discourse and interaction [1][2][3]. Similar work in raciolinguistics and Critical Race Theory (CRT) also brings to the table more complex understandings of race as a constantly evolving, context-dependent process [4][5]. As aptly put in the 2018 LSA statement on race: “Language & race ideology are co-constructed and co-negotiated between community, individual, and researcher” [6]. The goal, then, is to unearth the manner in which the intersections of race and place are constructed and negotiated by speakers. The scope of the study is both on the

individual and larger community level, through explicit interrogation of practices and ideologies by means of sociolinguistic interviews and small focus groups, allowing for both quantitative and qualitative analyses of interview data.

The primary research question of the current study is twofold: how do Black Chicagoans perform their identity as it pertains to both race and place, and how do they think they perform said identity? Specifically, I will assess:

1. some (morphosyntactic, phonetic, phonological) features of AAE typically covered in linguistic analyses of race performance [7][8], and

2. features or characteristics that speakers themselves identify

In particular, I am interested in the contrast between the two; what differences come to light when looking at research-based predetermined features versus features explicitly mentioned by participants? What impact does this difference have on the analyses that can be done? While much work has been done on AAE, little has focused on Chicago as a Black locale; it boasts a sizable Black population from the Great Migration while also being the biggest urban center of the Midwest. In focus groups and individual interviews, participants will be asked about their identity and expression, but they will also be asked, explicitly, about their linguistic actions and placemaking [9]. While investigations into social factors are currently happening in the field [10][11], an emergent analysis of participant-defined features is novel, and will be used to pivot towards the linguistic features that speakers themselves consider to be the most salient in constructing their identities.

This study plans to recruit at least 15 participants via the snowball method, branching from the author's own social networks. Sociolinguistic interviews will be 1:1 with the author, largely narratives about language behavior and ideologies, which will be further expanded upon in the small focus group, where participants will converse with the researcher and each other, to construct a larger group identity. Data will be transcribed and quantitatively analyzed for morphosyntactic, phonetic, and phonological measurements with similar qualitative discourse analysis.

- [1] Blake, R. (2016). *Toward heterogeneity: A sociolinguistic perspective on the classification of Black people in the twenty-first century*. In *Raciolinguistics: How language shapes our ideas about race*. Oxford University Press New York.
- [2] Calder, J., & King, S. (2022). Whose gendered voices matter?: Race and gender in the articulation of /s/ in Bakersfield, California. *Journal of Sociolinguistics*, 26 (5), 604–623. <https://doi.org/10.1111/josl.12584>
- [3] D'Onofrio, A. (2020). Personae in sociolinguistic variation. *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(6), e1543.
- [4] Carlton Parsons, E. R. (2016). Interfaces between critical race theory and sociocultural perspectives. In *Power and Privilege in the Learning Sciences* (pp. 46-67). Routledge.
- [5] Alim, H. S., Rickford, J. R., & Ball, A. F. (2016). Introducing raciolinguistics. In *Raciolinguistics: How language shapes our ideas about race* (pp. 1–30). Oxford University Press New York.
- [6] Hudley, A. H. C., Mallinson, C., Bucholtz, M., Flores, N., Holliday, N., Chun, E., & Spears, A. (2018). Linguistics and race: An interdisciplinary approach towards an LSA statement on race. *Proceedings of the Linguistic Society of America*, 3 (1), 8–1.
- [7] Green, L. J., Yu, K. M., Neal, A., Whitmal, A., Powe, T., & Özyıldız, D. (2022). Range in the Use and Realization of BIN in African American English. *Language and Speech*, 65 (4), 958–1006.
- [8] Thomas, E. R. (2015). Prosodic Features of African American English. In S. Lanehart (Ed.), *The Oxford Handbook of African American Language* (pp. 420–436). Oxford University Press Oxford, UK/New York.
- [9] Hunter, M. A., Pattillo, M., Robinson, Z. F., & Taylor, K. Y. (2016). Black placemaking: Celebration, play, and poetry. *Theory, Culture & Society*, 33(7-8), 31-56.
- [10] D'Onofrio, A., & Benheim, J. (2020). Contextualizing reversal: Local dynamics of the Northern Cities Shift in a Chicago community. *Journal of Sociolinguistics*, 24 (4), 469–491.
- [11] King, S. (2021). Rethinking race and place: The role of persona in sound change reversal. *Journal of Sociolinguistics*, 25 (2), 159–178.

---

## Examining variation and grammaticalization of the English expression a couple of in the Canadian prairies

Maria Rodrigo-Tamarit, Nicole Rosen  
University of Manitoba, Winnipeg, Canada

Variants couple and couple a of the expression a couple of have been recorded in written form since 1876 (Charles Henry Davis, 1876:163, cited in OED Online, 2018b7b), and 1906 (Green, 1906:254, cited in OED Online, 2018a), respectively. Despite the relatively prolonged use of these variants, to our knowledge, very few studies have examined (a couple of) systematically (Kjellmer, 2007; Wood, 2019), with two having analyzed the social variation of this variable (Lee & Raines, 2018; Author, 2019). Results from Lee and Raines' study showed

higher presence of a couple of overall, although the use of couple increased throughout the decades examined, suggesting a possible change in progress towards preposition deletion.

The present study examines three variants of the variable (couple), namely couple of (“I had a couple of aunts that lived” (054)), couple a (“the US team would do a couple a miles” (054)) and couple (“they moved here a couple years ago” (096)). Data was extracted from 124 sociolinguistic interviews with English speakers from two Canadian provinces collected between 2011 and 2019. Additionally, social and linguistic factors potentially impacting preposition deletion among this group of speakers are examined, namely participants’ gender (women 60.5% and men 39.5%) birth year (1929-2000), rurality (rural 48.4% and urban 51.6%), socio-economic status (professional 58% and non-professional 42%), province (Alberta 47% and Manitoba 53%), and the variable’s preceding context (fully pronounced 81.6% and reduced 18.4%), following phonological context (consonants 57.2% and vowels and glides 42.8%), and following lexical items (time expressions 61.4% and other 38.6%).

Logistic regression models show that participants’ birth year significantly predicts the use of this variable, namely, the younger the participants are, the more likely they are to use the variant couple and the less likely they are to use couple a and couple of, consistent with a generational change (Labov, 1994:84). Overall results also show a preference for couple (73.8%) over couple a (15.8%) and couple of (10.3%), a pattern which emerges in almost all factors tested. This pattern towards preposition deletion supports the notion that the variable is undergoing phonetic reduction or ‘erosion’ (Heine et al., 1991: 15), indicative of grammaticalization. The variant couple is also preferred when following time expressions, more frequent and presenting less variability than other expressions, when compared to couple of. Since high frequency has been tied to grammaticalization as being “both indicator and facilitator” (Brems, 2011:116), these results further confirm the grammaticalization of the expression. Relatedly, the higher presence of couple of in Lee and Raines (2018) and Wood (2019) reflects the difference in type of data analyzed, where written displayed less deletion than in natural speech. Furthermore, the least formal radio program in Lee and Raines contained the highest deletion rate, reinforcing the finding that preposition deletion increases in informal/natural speech.

Overall, we show that the understudied (a couple of) variable is undergoing a grammaticalization change in progress, supporting predictions made in Lee and Raines (2018) that natural speech would see more of deletion than the observed in their corpus.

- Brems, L. (2011). Grammaticalization, delexicalization and subjectification in SN- and TN-constructions. In *Layering of Size and Type Noun Constructions in English* (pp. 79–122). De Gruyter Mouton.
- Charles Henry Davis. (1876). *Narrative of the North Polar Expedition U.S. Ship “Polaris,” Captain Charles Francis Hall, Commanding Edited under the Direction of the Hon. G. M. Robeson, Secretary of the Navy, by Rear-Admiral C.H. Davis, U.S.N. (United States, Navy Department, & United States North Polar Expedition, Eds.)*. Government Printing Office. <https://doi.org/10.1038/016225a0>
- Green, H. (1906). *At the actors’ boarding house, and other stories* (First edit). The Nevada publishing co.
- Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A Conceptual Framework*. The University of Chicago Press.
- Kjellmer, G. (2007). A couple hours – On non-occurrence of of after partitive English nouns. *Zeitschrift Für Anglistik Und Amerikanistik*, 55(3), 293–312. <https://doi.org/10.1515/zaa.2007.55.3.293>
- Labov, W. (1994). *Principles of Linguistic Change. Volume 1: Internal Factors*. Blackwell.
- Lee, S., & Raines, T. (2018). “A couple (of) weeks ago”: Syntactic variation of a quantifying expression. Poster Session Presented at the 92nd Annual Meeting of the Linguistic Society of America (LSA 2018). <https://www.linguisticsociety.org/node/8208/schedule>
- OED Online. (2018a). *coupla*, n. In *Oxford English Dictionary*. Oxford University Press.
- OED Online. (2018b). *couple*, n. In *Oxford English Dictionary*. Oxford University Press.
- Author (2019). “Couple hours” or “couple of hours”? A variationist study on quantifier a couple of in Southern Alberta. Individual Paper Presented at the Change and Variation in Canada (CVC) 11.
- Wood, J. L. (2019). A couple (of) changes in the ‘Brown family’: British and US English compared. In Ken Ramshøj Christensen, Henrik Jørgensen, & Johanna L. Wood (Eds.), *The Sign of the V: Papers in Honour of Sten Vikner* (pp. 759–787). Aarhus University Library. <https://doi.org/10.7146/aul.348.123>

---

“¡A[ʎ]á no, pero aquí usamos la elle!”: Articulation as a sociolinguistic sign of identity in Bolivia

Miguel Roman, Rachel Sangster Garza  
Indiana University, Bloomington, USA

Yeísmo is a phonemic change that is characterized as the loss of the lateral palatal approximant /ʎ/, merging with various realizations that are associated with the grapheme *y*. This results in a loss of distinction between minimal pairs such as *se cayó* ‘they fell over’ and *se calló* ‘they fell silent.’ Regardless of the progress of this well-attested change throughout the Spanish speaking world, some regions maintain the phonetic contrast, known as lleísmo (García Mouton & Molina Martos, 2012; Gómez & Molina Martos, 2013; Lipski, 1989; 1996; Navarro, 1993;

Orzechowski & Seara, 2014; Rost Bagudanch, 2013; 2017). In other regions where lleísmo has been maintained, research on perception indicates that it is associated with regional identity, especially in the Andean regions of Peru and Ecuador (Cole, 2021; Godenzzi, 2013). As a generally understudied variety of Spanish that composes part of the Andean region, the present study focuses on perceptions of lleísmo in Bolivian Spanish, and asks the following questions: (1) What types of speakers do Bolivians associate with lleísmo? (2) What social values do Bolivians ascribe to the use of lleísmo? And yeísmo? (3) Does lleísmo hold overt prestige and/or mark solidarity or identity in Bolivia?

Twenty-nine Bolivian speakers of Spanish completed a matched guise activity via Qualtrics. During the task participants listened to audio – recorded by a female and male Bolivian speaker of Spanish, around 50 years of age – containing productions of lleísmo or yeísmo and assigned social characteristics to each recording. Each actor recorded 25 words containing the grapheme *ll*, produced as both /k/ and /j/ (e.g. olla ‘pot’ or detalle ‘detail’) within carrier phrases. Each listener heard 40 target stimuli and 10 distractors containing productions of intervocalic assibilated /r/ and responded to a social characteristics matrix as well as indicated the region they thought the person was from and the age of the speaker. One-Way ANOVAs were conducted to identify any significant differences in perception of /k/ and /j/, followed by linear and logistic regressions with the social characteristics as the dependent variable. Finally, qualitative data was collected which considered the participants’ perceptions of yeísmo v. lleísmo usage.

The results of statistical analysis suggest that yeísmo and lleísmo significantly impact Bolivian listeners’ social perception. In particular, productions of [k] were significantly associated with the characteristics of professional ‘professional’, simpático ‘nice’, educado ‘educated’ and use by older individuals. Regarding region of origin, [j] use was associated with being from another region in Bolivia or another country. These results indicate that use of lleísmo carries overt prestige that ties into what is perceived as “standard” or careful speech. Moreover, lleísmo can reinforce perceptions of regional identity, although more research is still required to determine the status of lleísmo in Bolivian speech as well. The results of this study have broader implications in the field of sociolinguistics as they demonstrate how identity acts as a driving force for maintenance of older linguistic forms.

- Cole, M. (2021). Sociolinguistics and Language Shift in Ecuador: A Case Study of Palatals. [Dissertation, Indiana University]. ProQuest Dissertations and Theses Global.
- García Mouton, P., & Molina Martos, I. (2012). The /k/-/j/ merger (yeísmo) in central Spain: advances since the ALPI. *Dialectología*. Special issue, III, 23-42.
- Godenzzi, J. C. (2013) Resistencias al yeísmo en los Andes: Convergencia lingüística y expresión identitaria. In Gómez, R., & Martos, I. M. (Eds.), *Variación yeísta en el mundo hispánico* (pp. 295-309). Iberoamericana Editorial Vervuert.
- Gómez, R., & Martos, I. M. (Eds.). (2013). *Variación yeísta en el mundo hispánico*. Iberoamericana Editorial Vervuert.
- Lipski, J. M. (1989). Spanish yeísmo and the palatal resonants: Towards a unified analysis. *Probus*, 1(2), 211-223.
- Lipski, J. M. (1996). *El español de América*. Madrid: Cátedra.
- Navarro, T. (1993). Nuevos datos sobre el yeísmo en España. In R. Páez Patiño & I. Chaves Cuevas (Eds.), *Thesaurus: Boletín del Instituto Caro y Cuervo: Muestra antológica 1945-1985*, I: Lingüística; II: Historia de la literatura, filología y análisis literario; III: Historia cultural, cultura popular, poesía latina, discursos. Inst. Caro y Cuervo.
- Orzechowski Dias, E. C., & Seara, I. C. (2014). Descripción fonético-acústica de las variantes del yeísmo presentes en el español de rioplatenses y de colombianos de la región de la costa atlántica y del valle del cauca. *Arco/Libros*, 399-418.
- Rost Bagudanch, A. (2013). La transcripción fonética en estudios dialectales: propuestas en el caso del yeísmo. *Revista de filología española*, 93(1), 165-192.
- Rost Bagudanch, A. (2017). Variation and Phonological Change: The Case of yeísmo in Spanish. *Folia Linguistica: Acta Societatis Linguisticae Europaeae*, 51(1), 169–206.

## From ‘Cuck’ to ‘Prostitute’: Crosslinguistic Variation of Dekasegi in Brazilian Portuguese and Japanese Online Discourse

Miguel Germán Román

Indiana University, Bloomington, USA

Although earlier work delineates a theory of indexicality (Silverstein 2003) and the mapping of linguistic features onto an indexical field of related meanings (Eckert 2008), Beaton & Washington (2015) are the first to apply this framework to semantic items. When citing indexicality as a practice of identifying with perceived characteristics that are associated with enacting group identity, Eckert (2008) highlights the need for research that examines the indexical field in a language contact environment. In this investigation, I examine the use of the term *dekasegi* “temporary migrants”, in online blogs, forums and Tweets in Brazilian Portuguese and Japanese. By contrasting the indexical values and dominant senses in both languages, this study demonstrates how attitudes and language contact in a migrant

situation allowed for the integration of a new lexical item into a community of Japanese heritage speakers. Notably, it also shows how the term varies in its usages crosslinguistically.

Indexicality applies to linguistic features, following base  $n$ -th order meanings, with further  $n + 1$  order social meanings added to the base (Silverstein 2003). Eckert (2008) proposes that  $n + 1$  meanings can be mapped onto an indexical field, representing speaker traits as social types, permanent qualities, and momentary stances. Beaton & Washington (2015) apply this framework to lexical items, demonstrating how the indexical field can represent amelioration in pejorative words like *favelado*. Following the groundwork from these researchers, I define base ( $n$ -th) meanings and competing ( $n + 1$ ) social meanings before constructing an indexical field to represent their differing distributions.

The term in question, *dekasegi*, originates from Japanese, where it means ‘one who leaves their home for work to save money’ (Morioka 2000). It was later borrowed into Portuguese as *decasségui*, where it denotes someone who migrates to Japan as a manual laborer (Houais et al. 2009). Apart from notable differences, where the Portuguese meaning includes Japan as a necessary destination, both of these words denote a migrant worker.

After extracting a total of 155 tokens in Brazilian Portuguese and 74 in Japanese, 44 representative usages of *dekasegi* in Brazilian Portuguese and 15 representative usages of *dekasegi* in Japanese were analyzed qualitatively and used to create the indexical fields of Figures (1) and (2). Discussions of *dekasegi* by Brazilian Portuguese speakers show a large range of negative (Example 1) and positive (Example 2) evaluations associated with the term. The expansion of positive and negative permanent qualities appears to be influenced by the large number of neutrally valenced social types. On the other hand, *dekasegi* appears to primarily denote ‘temporariness’ in Japanese, as in Example (3) below. Although there are connections made with international populations, the prevalent sense of ‘temporariness’ is further exemplified in a novel use (Example 4) relating to prostitution on Japanese Twitter. This new meaning is only accessible based on the prominent sense of temporariness for *dekasegi* in Japanese. Even though they share the same  $n$ -th order meaning, the indexical field surrounding these two terms shows how prevalent  $n + 1$  order meanings influence new uses in both languages.

Beaton, M. E., & Washington, H. B. (2015). Slurs and the indexical field: the pejoration and reclaiming of *favelado* ‘slum-dweller.’ *Language Sciences*, 52, 12–21.

Eckert, P. (2008). Variation and the indexical field. *Socioling.* 12, 453–476.

Houaiss, A., Villar, M., & Franco, F. Manoel de Mello. (2009). *Dicionário Houaiss da língua portuguesa* [Houaiss Portuguese Language Dictionary]. 1a ed. Rio de Janeiro: Objetiva.

Morioka, K. (2000). *Shūeisha kokugo jiten* [Shūeisha Japanese Language Dictionary]. Dai 2-han. Tōkyō: Shūeisha.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Lang. Commun.* 23, 193-229.

---

## Age and other social factors impacting generic third person pronoun use in nonnative English speakers

Ell Rose

Swarthmore College, Swarthmore, USA

It has been repeatedly found that nonnative English speakers tend to use generic singular they less frequently than L1 English-speakers to refer to a person of unknown gender, but that generic singular they use is nonetheless on the rise among nonnative speakers (Reilly-Thornton, 2021; Smith, 2022; Stormbom, 2019; Zhang & Yang, 2018). It has been argued that this difference between native and nonnative speakers is due to four factors:

- \* limited exposure to English outside of the classroom
- \* sociolinguistic factors in a speaker’s native language and home culture
- \* the classroom’s (and sometimes local culture’s) emphasis on grammatical accuracy
- \* the lag of English textbooks behind innovations taking place in English-speaking societies (Smith 2022, Zhang & Yang 2018).

However, relatively little research has been done to conclude whether individual speakers’ identities and social positions affect nonnative speakers’ generic pronoun choices. This project launch proposes a study to determine whether the social factors that affect native speakers’ use of generic pronouns—primarily age, gender, and number of transgender or nonbinary acquaintances—similarly affect nonnative speakers (Ackerman et al., 2018; Conrod,

2019; Hekanaho, 2020; Hernandez, 2020; Camilliere, 2021). I hypothesize that the three factors listed above which correlate with native speakers' use of generic singular they will also correlate with nonnative speakers' use of the pronoun. I also hypothesize that groups of nonnative English speakers in these social categories will use generic singular they more consistently than a similarly situated group of native speakers, and groups of those outside of these social categories will more consistently not use singular they. Given the factors above, exposure to, and acceptance of, new, controversial language may have to be a more deliberate process for nonnative speakers than for their native counterparts, which could mean that the difference in pronoun choice between nonnative speakers who have use for singular they and those that do not would be greater.

To test this hypothesis, I will conduct an online large-scale fill-in-the-blank Qualtrics survey with a robust demographic survey, including native language(s), familiarity with English, age, pronouns, gender, LGBTQ+ identity, location, and number of transgender or nonbinary acquaintances. Stimuli will include sentences with normed gender-neutral definite NP antecedents (from Mizersky et al., 2014) and blanks in place of possessive pronouns. This methodology deviates from the studies referenced above as this project uniquely facilitates organic generation of written generic pronouns. They will be randomized and interspersed with an equal number of filler sentences with blanks in place of prepositions. Preliminary results from an ongoing pilot survey (n=61) indicate that there may be a stronger relationship between age and generic pronoun use among nonnative English speakers than native English speakers (Figs. 1 and 2). This project seeks to verify and expand on these findings and determine whether and to what extent age and other social factors impact nonnative English speakers' pronoun use, and how or whether they may differ from native speakers.

- Ackerman, L., N. Riches, & J. Wallenberg. (2018). Coreference dependency formation is modulated by experience with variation of human gender. The 92nd Annual Meeting of the Linguistic Society of America (LSA2018) – Salt Lake City, UT (January 2018).
- Camilliere, S., Izes, A., Leventhal, O., & Grodner, D. (2021). They is Changing: Pragmatic and Grammatical Factors that License Singular they. Proceedings of the Annual Meeting of the Cognitive Science Society, 43. Retrieved from <https://escholarship.org/uc/item/3tc9s9b0>
- Conrod, K. (2019). Pronouns raising and emerging. Unpublished manuscript. PhD dissertation, University of Washington.
- Hekanaho, L. (2020). Generic and nonbinary pronouns: Usage, acceptability and attitudes. Helsingin yliopisto. <https://helda.helsinki.fi/handle/10138/321581>
- Hernandez, E. (2020). Pronouns, Prescriptivism, and Prejudice: Attitudes toward the Singular 'They', Prescriptive Grammar, and Nonbinary Transgender People (Version 1). Purdue University Graduate School. <https://doi.org/10.25394/PGS.12231095.v1>
- Mizersky, J., Gygax, P.M., Canal, P. et al. (2014). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behav Res* 46, 841–871. <https://doi.org/10.3758/s13428-013-0409-z>
- Reilly-Thornton, A. (2021). Gendered pronouns in chinese EL2 speech: The case of epicene pronouns. *TEANGA, the Journal of the Irish Association for Applied Linguistics*, 28, 247–277. <https://doi.org/10.35903/teanga.v28i.697>
- Smith, C. (2022). Japanese university students' exposure to and attitudes toward singular they. *Bulletin of Faculty of International Studies, Niigata University of International and Information Studies (新潟国際情報大学 国際学部 紀要)*. [https://nuis.repo.nii.ac.jp/?action=repository\\_action\\_common\\_download&item\\_id=3496&item\\_no=1&attribute\\_id=22&file\\_no=1](https://nuis.repo.nii.ac.jp/?action=repository_action_common_download&item_id=3496&item_no=1&attribute_id=22&file_no=1)
- Stormbom, C. (2019). Language change in L2 academic writing: The case of epicene pronouns. *Journal of English for Academic Purposes*, 38, 95–105. <https://doi.org/10.1016/j.jeap.2019.02.001>
- Zhang, X., & Yang, H. (2021). Gender voices in Chinese university students' English writing: A corpus study. *Linguistics and Education*, 64, 100935. <https://doi.org/10.1016/j.linged.2021.100935>

---

## A culture of labour: indexing 'blue-collar' through a lower /s/ COG in Manitoba, Canada

Nicole Rosen<sup>1</sup>, Alexandra Pfiffner<sup>2</sup>

<sup>1</sup>University of Manitoba, Manitoba, Canada. <sup>2</sup>University of California, Berkeley, Berkeley, USA

In this paper we show that /s/ COG values in Manitoba are substantially lower than are seen in other North American English varieties, including in Canada (c.f. Podesva & Van Hofwegen 2014, Hazenberg 2016). We argue that these low values reflect a culture of labour, drawing on Podesva and Van Hofwegen (2014)'s 'country' persona, but are also bolstered by a Ukrainian substrate. We also find evidence of change in progress, where COG values are lowering in apparent time in the city of Winnipeg, and are furthermore lowest in the city as compared to the rural regions.

Sibilants have been well-studied as a locus of variability in gendered speech, with a particular focus on the center of gravity (COG) of /s/. A continuum emerges from cisgender men with the lowest mean COG to cisgender women with the highest mean COG (Podesva & Van Hofwegen 2014, Hazenberg 2016). Note that while gay men in rural California have a higher COG than straight men, these values are not as high as in San Francisco, which Podesva & Van Hofwegen attribute to the suppression of the higher values due to prevailing anti-LGBTQ sentiment in the area.

The interplay of ‘country’ rurality and the gay population in rural California shows that there are interesting implications for how sibilance does social work in speech, such as that lower COG may index more ‘country’, heterosexual male speech.

The present study further explores COG variation, focusing on four communities in Manitoba. Our data come from a corpus of words lists with 23 speakers from the Interlake, 21 from the Southern region, 36 from Steinbach, and 37 from Winnipeg. A total of 7,591 /s/ tokens were extracted and their COG measured in Praat (Boersma & Weenink 2023), with the window centered on the midpoint of the sibilant. Our results show that overall, COG measurements are much lower in these Canadian Prairie communities in comparison to other documented varieties of North American English, with averages ranging from 4422-5612 Hz among female speakers, and 3119-4113 Hz among male speakers. When further broken down by community, Winnipeg was significantly different ( $p=0.03$ ) with the lowest mean COG. Table 1 shows the averages by gender in each community.

We argue that while the overall low COG measurements in the Canadian Prairies reflect Podesva’s ‘country’ indexation, additional factors are at play in the region. Manitoba, with a strong ethnically Ukrainian population, has an added Ukrainian substrate, which also displays very low COG values (see Ukrainian L1 speakers in Figure 2). Furthermore, there is a long tradition of blue-collar and agricultural labour in Manitoba, perhaps most notably represented by the 1919 Winnipeg General Strike. Noting that the lowest COG values are found in younger speakers in the city of Winnipeg, we further suggest that there is a change currently in progress towards lower COG, playing on the area’s long culture of labour.

Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer (Version 6.3.02). [Computer program].  
<http://www.praat.org>

Hazenbergh, E. (2016). Walking the straight and narrow: linguistic choice and gendered presentation. *Gender and Language*, 10(2), 270-294.

Pfiffner, A. & Rosen, N. (2023). Stop voicing in two settler communities on the Canadian Prairies. To appear in *Proceedings of ICPHS 2023*.

Podesva, R. J. & Van Hofwegen, J. (2014). How conservatism and normative gender constrain variation in inland California: The case of /s/. *University of Pennsylvania Working Papers in Linguistics*, 20(2), 129-137.

---

## Embodying Energy: Correlating Pitch and Gesture in Performances of Social Personae

Parker Scarpa

Reed College, Portland, USA

This project investigates embodiment as a stylistic practice in performances of social personae through a quantitative analysis of pitch and gesture. Under the framework of the indexical field, the social meaning of a variable is argued to be underspecified, drawing from a field of possible meanings to take on specificity within a particular social context (Eckert 2008). Under this framework, studies such as Mendoza-Denton (2011) and D’Onofrio (2018) have identified the social persona as an effective vehicle through which to study these context-specific social meanings because of its ability to efficiently represent a complex, ideologically-ingrained stylistic package. Furthermore, recent studies have examined the relationship between embodied practices and stylistic variation, arguing that the configuration of the body itself affects the meaning of a variable (Pratt 2018; Calder 2019; Podesva 2021). However, little work has applied this embodied approach to the framework of the indexical field within the context of the social persona.

I contribute data from an experimental study in which two participants performed impressions of five social personae: Business Professional, Fitness Instructor, Surfer, Valley Girl, and Yoga Teacher. During the experiment, participants were instructed to read from short passages while imitating the speech patterns, gestures, and mannerisms of the personae in such a way that their partner could identify it. I adopted this novel methodology to promote exaggerated performances, emphasizing the most salient features of these personae; the intent was to encourage extreme stylistic variation, not authentic depictions.

My analysis investigates the correlation between manual gesture and pitch. To do this, I recorded visual and auditory data tracking each performer’s use of pitch and gesture. I used the DARLA web interface’s semi-automated alignment feature to transcribe auditory data into word-level segments, from which I extracted a total of 2,314 mean pitch measurements. These measurements were then categorized according to the type of gesture being performed at the moment of utterance: resting, one-handed, or two-handed.

My within-subject analysis, which fit mixed-effect regression models to participants’ pitch measurements, identified main effects of persona and gesture on pitch. The effect of persona affirms the previously-attested salience of the persona as a vehicle for stylistic variation, and the nature of the relationship between persona and pitch aligns with standard linguistic and gender ideologies, such that participants used a lower pitch for personae they identified as male and a higher pitch for those they identified as female. However, the independent effect of gesture on pitch, such that two-handed gestures correlated with higher mean pitch measurements regardless of persona, suggests that



pitch and gesture are simultaneously recruited to convey social meanings beyond the stylistic variation attributed to the personae.

I argue that, while pitch itself is indexically linked to gender identity, it takes on new meaning in the presence of embodied actions, such that the simultaneous recruitment of pitch and gesture is iconically linked to ideologies of bodily energy (Esposito & Gratton 2022). Crucially, this embodied approach gives greater insight into the social meaning of variation within the framework of the indexical field.

Calder, J. (2019). The fierceness of fronted /s/: Linguistic rhematization through visual transformation. *Language in Society*, 48(1), 31–64.

D’Onofrio, A. (2018). Personae and phonetic detail in sociolinguistic signs. *Language in Society*, 47(4), 513–539.

Esposito, L., & Gratton, C. (2022). Prosody and ideologies of embodiment: Variation in the use of pitch and articulation rate among fitness instructors. *Language in Society*, 51(2), 211–236.

Mendoza-Denton, N. (2011). The semiotic hitchhiker’s guide to creaky voice: Circulation and gendered hardcore in a chicana/o gang persona. *Journal of Linguistic Anthropology*, 21(2), 261–280.

Podesva, R. J. (2021). The role of the body in language change. In E. Moore, L. Hall-Lew, & R. J. Podesva (Eds.), *Social Meaning and Linguistic Variation: Theorizing the Third Wave* (pp. 363–381). Cambridge University Press.

Pratt, T. (2018). *Affective Sociolinguistic Style: An Ethnography of Embodied Linguistic Variation in an Arts High School* [Stanford University].

---

## Decoding Social Inequality in Historical Documents. Unraveling Traces in the Early Modern Era

Christa Schneider

University, Bern, Switzerland. University, Salzburg, Austria

In the early modern state of Bern (Switzerland), certain criminals were brought to the Tower of Bern where they stood trial. The records of these trials were collected in the Bernese Tower Books and are currently accessible in the State Archive of Bern. Unfortunately, systematic research on the Bernese Tower Books from historical, linguistic, or jurisprudential perspectives is still lacking. The absence of research may be attributed to the size of the corpus (approximately 250,000 pages) and to the type of handwriting used in the documents (German Kurrent). Moreover, investigations concerning historical dialects and vernaculars in Switzerland are generally scarce. The following project aims to address these gaps.

One of the primary objectives of the project Crimes of the Past is to identify and describe the use of historical Bernese Swiss German, the vernacular which served as the spoken language in the early modern city-state of Bern. A pilot study has already revealed significant differences between the spoken and the written language. Interestingly, this points towards the early development of the Swiss medial diglossic situation, which can now be tracked back at least 500 years. Around 1550, early Bernese Swiss German was used as a spoken language, while another variant of German served as the written language (cf. Sonderegger 2003: 2849).

In the past few months and supported by recent developments in Digital Humanities, a sample of 30’000 pages (1547-1747) has been digitized using ScanTents and was subsequently automatically transcribed with the assistance of Transkribus[1]. This program uses advances from AI to transform handwritten text first into structured images and then into digital text. Based on principles of Deep Learning, a language model was later trained which is specifically suitable to improve the recognition and transcription of the Tower Book text. This approach has resulted in automatic transcriptions of the handwritten texts with a character error rate of approximately 8-10%. With this improvement it became significantly easier accessing the Tower Book sample and since very recently, a first batch of data is ready for analysis.

In this presentation, I will focus on the traces of early Bernese Swiss German within the written language of the Tower Books and address some key questions: How prevalent is the spoken language in this corpus and what purpose did it serve? Through selected trial records, I will demonstrate how different degrees of language variation reflect differences among the accused individuals in terms of education, social status, and gender. There are, for instance, two uneducated young men living in an orphanage, accused of sodomy. Analyses of the trial records have shown that these men experience a completely different linguistic treatment in front of the court compared to the aristocrat, who is accused of murdering his maid.

[1] <https://readcoop.eu/transkribus/?sc=Transkribus>

Sonderegger, Stefan (2003): Aspekte einer Sprachgeschichte der deutschen Schweiz. In: Werner Besch et al. (Eds.): *Handbücher für Sprach- und Kommunikationswissenschaften, Sprachgeschichte, Band 2.3*. De Gruyter: Berlin, New York

---

## "I don't even take note of the bad English lyrics": The listening subjects of international K-pop

Ian Schneider

University of Kentucky, Lexington, USA

As Korean popular music (K-pop) grows internationally, so too has the use of English within K-pop song lyrics (Ahn, 2021). Yeo (2018:108) posited a register of K-pop English as "instances of non-native-like English in K-pop song lyrics" intended to index cosmopolitan yet distinctly Korean personae. While some scholars and international listeners describe this English use as "grammatically incorrect" (Lawrence, 2010) or "funny" (Yeo, 2018) limited work has considered the ideological dimensions of international listeners of K-pop as "listening subjects" (Inuoe, 2006). This paper therefore explores the enregisterment (Agha, 2005) of K-pop English through the metalinguistic commentary of six articles published on five online English-language K-pop news websites. How does such commentary reflect the ideological discourses of English-language lyrics among international K-pop fans?

This study analyzes 84 instances of metalinguistic commentary from 6 online news articles about notable English lyrics in K-pop songs. These comments comprise of 68 reported song lines from 52 different songs and 41 different artists. A thematic analysis (Braun & Clarke, 2006) coded and consolidated the ideological positions of each instance of metalinguistic commentary into four themes: Disjunct, Innuendo, Cringe, and Positive (Table 1). "Innuendo" involved comments that suggested covert sexual or sensual undertones within the lyrics (Excerpt 1). "Disjunct" entailed comments about song lyrics diverging from commenters' syntactic or discursive expectations as English users (Excerpt 2). "Cringe" involved general expressions of vicarious embarrassment over violated social expectations or "mismatch between the intended self-presentation and the self-presentation actually created" (Wöhrle, 2021:10) (Excerpt 3). Finally, "Positive" involved authors expressing appreciation or otherwise positive sentiments toward the lyric (Excerpt 4).

While this thematic analysis serves to forward discussions on English in K-pop lyrics often described in vague terms like "non-native-like" (Yeo, 2018) or "grammatically incorrect" (Lawrence, 2010), further ideological analysis through (retro) contextualization (Makoni, 2014) also yields insights into international K-pop fans as "perceiving subjects" (Park, 2022) or how subjects choose to see, hear, or perceive those (K-pop idols) who are speaking or acting (Inuoe, 2006). For example, commentary themed as "Innuendo" point to fans' broken illusion around K-pop idols' manufactured personae that often juxtapose sexualization and innocence (Oh, 2018) (Excerpt 1). And commentary themed as "Disjunct" often suggest ideologies of linguistic purism (or the group policing of correct uses of English or Korean) among international K-pop fan communities (Chun, 2017) (Excerpt 2).

This study furthers analysis of what scholars and fans alike have termed "Engrish" (Ikeshima, 2005) within K-pop lyrics first by conducting a thematic analysis of published metalinguistic commentary of English lyrics in K-pop and then connecting that metacommentary to the ideological positioning of international K-pop fans as listening subjects. Language use is not neutral, and no perspective on language comes from nowhere (Irvine & Gal, 2000). Therefore, this study builds upon a growing call to interrogate the historical embeddings of metapragmatic discourses among digitally mediated pathways (Lo & Park, 2017) within the context of a popular yet under-researched site of metalinguistic ideology – K-pop.

Agha, A. 2005. 'Voice, footing, enregisterment.' *Journal of linguistic anthropology*, 15(1), pp.38-59.

Ahn, H., 2021. English and K-pop: The case of BTS 1. In Low, E.L. & Pakir, A. (eds.) *English in East and South Asia*. New York: Routledge. pp.212-225.

Braun, V. and Clarke, V. 2006. 'Using thematic analysis in psychology.' *Qualitative research in psychology*, 3(2), pp.77-101.

Chun, E.W. 2017. 'How to drop a name: Hybridity, purity, and the K-pop fan.' *Language in Society*. 46(1), pp.57-76.

Ikeshima, J.H. 2005. 'Some Perspectives on the Phenomenon of "Engrish"'. *The Keiai journal of international studies*, (15), pp.185-198.

Inuoe, M., 2006. *Vicarious language: Gender and linguistic modernity in Japan*. Berkeley: University of California Press.

Irvine, J & Gal, S. 2000. 'Language Ideology and Linguistic Differentiation.' In Kroskrity, P.V. (ed.). *Regimes of Languages*.

*Ideologies, Politics, and Identities*. Houston: School of America in Research Advanced Seminar Series, pp. 35-83

Lawrence, B.C. 2010. 'The verbal art of borrowing: Analysis of English borrowing in Korean pop songs.' *Asian Englishes*, 13(2), pp.42-63.

Lo, A. and Park, J. 2017. 'Metapragmatics of mobility.' *Language in Society*, 46(1), pp.1-4.

Makoni, S. 2014. "'The Lord is my shock absorber': A sociohistorical integrationist approach to mid-twentieth-century literacy practices in Ghana.' In Blackledge, A & Creese, A (eds.) *Heteroglossia as practice and pedagogy*. Amsterdam: Springer. pp.75-97.

Oh, Y., 2018. *Pop city: Korean popular culture and the selling of place*. Ithaca: Cornell University Press.

Park, J. 2022. 'Foreign language education at the nexus of neoliberalism and coloniality: subjectivity in South Korean discourses of education reform.' *Critical Inquiry in Language Studies*, 19(4), pp.336-354.

Wöhrle, P., 2021. 'Two Shades of Cringe: Problems in Attributing Painful Laughter.' *Humanities*, 10(3), p.99.

Yeo, R.C.L. 2018. *The Commodification of English in K-Pop: Globalisation and Multiple Markets*. MA Thesis. Singapore: National University of Singapore.

---

## Witchcraft and the Historical Paradox: Analysing Sentiment in Salem Witch Trial Protocols

Christa Schneider<sup>1,2</sup>, Christopher Strelluf<sup>3</sup>

<sup>1</sup>University of Bern, Bern, Switzerland. <sup>2</sup>University of Salzburg, Salzburg, Austria. <sup>3</sup>University of Warwick, Coventry, United Kingdom

This project explores the use of sentiment analysis in historical sociolinguistics to detect language variation caused by emotions. We apply sentiment analytic techniques to examine features of speech-like language in the Salem Witchcraft Papers corpus.

Sentiment analytic techniques collect numerical language users' ratings of emotional, psychological, cognitive, and semantic features of words. These ratings are then applied to words in a document to estimate overall emotional characteristics of language in the document. Sentiment analysis is commonly used in commercial applications, in corpus linguistics, and behavioral psychology, but its potential for sociolinguistic research has yet to be fully explored.

Historical texts pose particular challenges for sentiment analysis, because—in addition to better-studied patterns of language variation and change including the emergence or disappearance of words or changes in lexical semantics—emotional features associated with words may change over time. This creates an affective layer to Labov's "historical paradox," since we cannot know the extent to which people in the past "felt" differently about words.

In this project, we demonstrate possibilities for addressing these challenges to use sentiment analysis in historical sociolinguistics. We explore the Salem Witchcraft Papers (SWP)—a corpus of protocols recorded during the Salem Witch Hunt in 1692 and 1693. During that period, 200 individuals were accused of witchcraft and stood trial—with 30 being found guilty and 19 of these being executed. 140 protocols from these trials have so far been transcribed and made digitally available (<https://salem.lib.virginia.edu/home.html>).

To address problems of diachronic change in emotional features of language, we leverage the technical infrastructure provided by NLTK (<https://www.nltk.org>) and Python. Our preliminary results demonstrate that it is (a) possible to quantify and visualize emotions within the SWP corpus. Additionally, we have discovered that (b) a computed emotional value for a specific witch paper does not necessarily correlate with the verdict. Moreover, we suggest that (c) in cases where sentiment and verdict do not align, the detected higher diversity and quantity of collocations possibly point towards ongoing semantic change.

Thus, our approach allows us to quantify sentiment ratings as predictors of semantic change. In doing so, we demonstrate the potential of sentiment analysis as a tool for sociolinguistics generally and in historical sociolinguistics specifically. We also demonstrate a natural language processing solution to Labov's "historical paradox" in the context of lexical emotional and affective features. As such, our approach more broadly demonstrates the value of leveraging new computational technologies to develop novel solutions to the challenges of studying diachronic sociolinguistic change.

---

## Sissy that talk! A Venezuelan drag queen's Spanish rhotics

Michael Sciacca

Georgetown University, Washington, D.C., USA

This study addresses the lack of systematic documentation on the acoustic correlates interacting with gay and lesbian speaker social identities in Spanish. While over a decade has passed since Mack's (2010) research on the perception of sexual orientation in Puerto Rican Spanish, the need for comprehensive investigation remains. Previous literature reveals that men are more likely to delete consonants than women (Díaz-Campos & Killam, 2012), women articulate syllable-final (r) more than men in the Caribbean (Alfaraz, 2011; D'Introno et al., 1979; Prosper-Sanchez, 1995), and coda (r) retention is perceived as a prestigious variant (Díaz-Campos & Killam, 2012). However, Sciacca (2022) suggests that the previous literature incompletely analyzes the data. In Sciacca (2022), Venezuelan TikTokers' syllable-final (r) durations were measured, and the findings revealed that gay men patterned with heterosexual women, and gay women with heterosexual men. This calls for a nuanced analysis of gender contrasts considering sexual orientation and identity performance. As such, the present study investigates the coda rhotics produced by gay Venezuelan youtuber, La Divaza.

Data was collected from La Divaza's podcast titled "Radio Divaza," where she, in and out of drag, interviews Latin American celebrities. Three hours of audio were auto-transcribed and auto-aligned; 481 coda (r) segmentations were verified by the researcher according to the acoustic signals outlined by Bradley and Willis (2012). Data was then categorized by performance—drag vs. no drag—and interlocutor gender and sexual orientation to analyze the potential effects of accommodation. La Divaza interviewed one gay male, two heterosexual males, and two heterosexual females.

Given the challenges in measuring rhotics, this study aimed to identify the acoustic measure with the highest predictive power in characterizing taps, trills, and weakened rhotics. A random forest classification model assessed the accuracy of various measures—F1, F2, F3, F4, F3-F2 difference, duration, COG, standard deviation, skewness, and kurtosis—ranking duration as the most effective. Subsequently, a linear, mixed-effects model was run with duration as the dependent variable, word as the random effect, and independent variables including stress in word, manner of the preceding and following segments, preceding vowel duration, performance, interlocutor gender, and interlocutor sexual orientation. Preceding vowel duration was included as an indirect measure of speech rate, a known covariate of acoustic measures with a durational component (Balukas & Koops, 2015).

The results revealed a significant correlation between speech rate and rhotic length, as expected. However, unexpected interactional effects were observed between speech rate, interlocutor sexual orientation, and performance. La Divaza spoke significantly faster in drag, regardless of the interlocutor's social background. In drag, fast speech in interactions with heterosexual interlocutors resulted in decreased rhotic duration. However, in drag, fast speech produced the opposite effect when speaking with the gay interlocutor. Here, La Divaza's rhotics were even longer. These results suggest three main implications in La Divaza's Spanish: (1) duration effectively measures coda rhotics, (2) fast speech may be associated with her performance of femininity, and (3) longer rhotic duration may be associated with her performance of femininity in contexts where interlocutors share similar identities.

- Alfaraz, G. G. (2011). Cuban Spanish in the US context: Linguistic and social constraints on the variation of syllable final (r) among Cuban newcomers. *Sociolinguistic Studies*, 5(2), 291.
- Balukas, C., & Koops, C. (2015). Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*, 19(4), 423-443.
- Bradley, T. G., & Willis, E. W. (2012). Rhotic variation and contrast in Veracruz Mexican Spanish. *Estudios de fonética experimental*, 43-74.
- Díaz-Campos, M., & Killam, J. (2012). Assessing language attitudes through a matched-guise experiment: The case of consonantal deletion in Venezuelan Spanish. *Hispania*, 83-102.
- D'Introno, F., Rojas, N., & Sosa, J. M. (1979). Estudio sociolingüístico de las líquidas en posición final de sílaba y final de palabra en el español de Caracas. *Boletín de la Academia Puertorriqueña de la Lengua Española*, 7(2), 59-100.
- Mack, S. (2010). Perception and identity: Stereotypes of speech and sexual orientation in Puerto Rican Spanish. In *Selected proceedings of the 12th Hispanic linguistics symposium* (pp. 136-147). Somerville, MA: Cascadilla Proceedings Project.
- Prosper-Sanchez, G. D. (1995). Neutralización homofónica de líquidas a final de sílaba: aspectos sociolingüísticos en el español de Puerto Rico (Doctoral dissertation, University of Massachusetts Amherst).
- Sciaccia, M. A. (2022). Coda /r/ articulation in Venezuelan Spanish from TikTok: Implications for sexual orientation [Paper presentation]. *New Ways of Analyzing Variation*, Stanford University.

---

## "Oh sure." The final period in texting as an 'anti-exclamation' mark

Michael Shepherd, MJ Jaurena

California State University, Fresno, Fresno, CA, USA

Paralinguistic cues, including the constellation of phonetic features commonly known as 'tone of voice,' are central to our interpretation of speech (Burgoon and Hoobler 2002). Writing does not have tone of voice, but as texting evolved—into what McWhorter (2013) calls "fingered speech"—conventions developed that substitute other paralinguistic cues, including capitalization, letter repetitions, and punctuation (Kalman and Gergle 2014). Unpunctuated texts became the norm and final periods rare and pragmatically marked (Baron and Ling 2011). Now, when a text responding to an invitation ends with a period (vs. without punctuation), the respondent is rated as less sincere (Gunraj et al. 2016), as not wanting to engage in the activity as much, and as feeling more negatively about it (Houghton et al. 2018). Such findings support the widespread intuition that final periods in texting convey something unfavorable. However, exactly what they convey remains unclear.

To further explore the pragmatic significance of final periods in texting, we examined the phonetic characteristics of young people's readings of period-final (vs. unpunctuated) texts. Our participants were 50 native speakers of American English (27 female, 23 male), ages 18-29 ( $M = 21.3$ ,  $SD = 2.34$ ). Our methodology replicated and extended Houghton et al.'s (2018) study. We showed participants 32 short text exchanges, adapted from Houghton et al.'s stimuli, but asked them to read the texts aloud—conveying whatever 'tone of voice' they perceived in them—prior to rating them. Having participants read texts aloud parallels Heath's (2021) prosodic analysis of messages in ALL CAPS. If final periods convey abruptness, as Houghton et al. suggested, readings of period-final texts should be shorter in duration than readings of unpunctuated texts.

Consistent with Houghton et al.'s findings, when a text responding to an invitation, request, or offer ended with a period (vs. without punctuation), the respondent was rated as feeling significantly more negatively about the topic,  $t(29) = 3.89$ ,  $p < .001$ , Cohen's  $d = .55$ . Having replicated Houghton et al.'s rating results and effect size, we ran linear mixed-effects models (with participant and item as random factors) comparing the duration and pitch

characteristics of participants' readings of the period-final vs. unpunctuated stimuli. The model for duration showed that readings of period-final (vs. unpunctuated) texts were significantly longer,  $F(1, 734.00) = 6.51, p = .011$ . Additionally, the pitch analyses showed that readings of period-final (vs. unpunctuated) texts had significantly lower F0 (relative to each speaker's mean),  $F(1, 777.01) = 9.29, p = .002$ , a significantly narrower F0 range,  $F(1, 728.34) = 4.45, p = .035$ , and significantly less F0 variation (smaller standard deviation),  $F(1, 727.74) = 4.25, p = .039$ .

In light of the phonetic profile of period-final texts—longer duration (slower speech), lower pitch, narrower pitch range, and less pitch variation—we argue that they convey not abruptness (given their longer duration), but rather a lack of enthusiasm. In other words, the final period in texting has evolved into a sort of 'anti-exclamation,' conveying roughly the opposite of what an exclamation mark conveys.

Baron, Naomi S. and Rich Ling. 2011. Necessary smileys and useless periods: Redefining punctuation in electronically-mediated communication. *Visible Language* 45:45–67.

Burgoon, Judee K. and Gregory. D. Hoobler (2002). Nonverbal signals. In *Handbook of Interpersonal Communication*, 3rd edition, ed. M.L. Knapp and J.A. Daly, 240–299. Thousand Oaks, CA: Sage Publications.

Gunraj, Danielle N., April M. Drumm-Hewitt, Erica M. Dashow, Sri Siddhi N. Upadhyay and Celia M. Klin. 2016. Texting insincerely: The role of the period in text messaging. *Computers in Human Behavior* 55:1067–1075.

Heath, Maria. 2021. NO NEED TO YELL: A prosodic analysis of writing in all caps. U. Penn Working Papers in Linguistics 27.1: Proceedings of the 44th Annual Penn Linguistics Conference, ed. A. Kalomoiros and L. Paparounas, 67–76.

Houghton, Kenneth J., Sri Siddhi N. Upadhyay and Celia M. Klin. 2018. Punctuation in text messages may convey abruptness. *Period. Computers in Human Behavior* 80:112–121.

Kalman, Yoram M. and Darren Gergle. 2014. Letter repetitions in computer-mediated communication: A unique link between spoken and online language. *Computers in Human Behavior* 34:187–193.

McWhorter, John. 2013. Txtng is killing language. JK!!! Ted Talk. URL [https://www.ted.com/talks/john\\_mcwhorter\\_txtng\\_is\\_killing\\_language\\_jk](https://www.ted.com/talks/john_mcwhorter_txtng_is_killing_language_jk).

---

## Variation in /st/ cluster production of Western Andalusian Spanish

Carolyn Siegman

University of Hawai'i at Mānoa, Honolulu, USA

Lexical frequency effects have previously been found in reductive sound changes in progress. This study finds that while certain trends appear overall across a group of speakers, these trends do not manifest in the speech of some individual speakers. In Western Andalusian Spanish, /st/ clusters are currently undergoing a multi-phase sound change in which initially, /s/ was aspirated before the stop release ([ht]), then was retimed to follow the stop release ([th]), and more recently is shifting to affrication ([ts]) (Ruch 2012; Ruch & Harrington 2014). This study investigates lexical frequency effects in this sound change and finds differing trends across the group and individual levels of analysis.

Syllable-final /s/ aspiration in Andalusia has been known to correlate with social factors like age and socioeconomic/education level (Cronenberg et al. 2020; Vida-Castro 2022), and previous work examining other varieties of Spanish found that lexical frequency effects play a role in aspiration rate (File-Muriel 2009; Brown 2008). Lexical frequency effects have not previously been studied in the present phenomenon, and while this perspective has been applied to gradient changes in sound change, the phenomenon in Andalusian Spanish has several distinct stages that make it a unique application of this approach.

Recordings of eighteen speakers in Sevilla (nine male, nine female) from the PRESEEA corpus were used for this study. A total of 935 tokens of word-medial /st/ clusters were coded categorically as standard ([st]), pre-aspirated ([ht]), pre-and post-aspirated ([hth]), post-aspirated ([th]) or affricated ([ts]) using auditory analysis and spectrographic cues like presence of fricative noise, fricative positioning relative to the stop release, and VOT. This coding allowed for an analysis of tokens along the established sound change cline. Lexical frequency was calculated using the CREA corpus (Real Academia Español 1975-2004).

Analysis shows that while overall speakers produce /st/ clusters with post-aspiration more than any other variant, there is considerable variation in production depending on the speaker, even within the same age group and education level (Figure 1). Some speakers clearly favor the post-aspirated variant while others primarily use affrication or the standard variant. Thus, speakers are at different points along the cline, even within categories like age group and educational background. Analyzing the data along the cline in a cumulative link model (Table 1) shows overall there are significant differences between the highest and lowest frequency words, between age groups, and between some education levels, but lexical frequency effects are not present for individual speakers in the sample. The quantitative results demonstrate the importance of examining inter-speaker variation in order to understand the full scope of the sound change. Therefore, a qualitative analysis is also conducted to examine the patterns both individually by speaker and within education levels and age groups. These findings contribute to the understanding of the ongoing change in Andalusian Spanish as well as the complex nature of how a sound change

emerges within a population, such that while certain trends may appear on the group level, individual speakers may show vastly different production rates of each variant stage.

- Brown, E. K. (2008). A usage-based account of syllable- and word-final /s/ reduction in four dialects of Spanish. [Doctoral Thesis, University of New Mexico]. ProQuest Dissertations & Theses Global.
- Cronenberg, J., Gubian, M., Harrington, J., & Ruch, H. (2020). A dynamic model of the change from pre- to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 83, 101016.
- File-Muriel, R. J. (2009). The Role of lexical frequency in the weakening of syllable-final lexical /s/ in the Spanish of Barranquilla, Colombia. *Hispania*, 92(2), 348–360.
- PRESEEA (2014+). Corpus del Proyecto para el estudio sociolingüístico del español de España y de América. Alcalá de Henares: Universidad de Alcalá. [<http://preseea.linguas.net>].
- Real Academia Español. (1975-2004). Corpus de referencia del español actual. Database (CREA) [online] [Consulted April 2023].
- Ruch, H. (2012). Affrication of /st/ in Western Andalusian Spanish: Variation and change from a sociophonetic point of view. In S. Calamai, C. Celata, & L. Ciucci (Eds.), *Proceedings of “Sociophonetics, at the crossroads of speech variation, processing and communication”* (61–64). Pisa: Edizioni della Normale.
- Ruch, H. & Harrington, J. (2014). Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 45, 12-25.
- Vida-Castro, M. (2022). On competing indexicalities in southern Peninsular Spanish: A sociophonetic and perceptual analysis of affricate [ts] through time. *Language Variation and Change* 34, 137–163.

---

## Resisting the fall of democracy: The Linguistic landscapes of protest spaces in Israel, 2023

Elana Shohamy and Iair G. Or  
Tel Aviv University

The study is anchored in the situation that took place in Israel in the fall of 2022. The results of the election brought about fierce demonstrations all around Israel. In early 2023, shortly after the right-wing parties gained a majority in the parliament and were able to form a government, Israel has become an extremely turbulent space in which protests and polarization were at their peak. On January 4th, Minister of Justice Yariv Levin announced a radical judiciary reform that would weaken the justice system, giving the government full control of it. This was seen by many as an attempt to destroy the democratic system and institutions, and a wave of massive civil protests began, creating protest spaces that were both powerful and diverse. In the present paper we seek to expand the framework of LL in protest spaces and investigate what makes particular protest spaces unique and different from other protest spaces.

Government representatives tried to color the protestors as unpatriotic conspirers, even though the main bulk of protesters was mainly made of Ashkenazi, middle-class citizens who carried Israeli flags and mentioned the 1948 Israeli Declaration of Independence and its democratic foundations as their source of inspiration. The protest spaces were highly diverse, contained multiple voices and agendas, and were divided into areas with participants from different parts of society and different groups invoking multiple agendas – including, but not limited to the protection of democracy, human rights, and the fear of the establishment of a religious theocracy. In the signs, Israel was compared to a variety of countries including Hungary, Turkey, Morocco and Iran, and the agendas were very broad, claiming that much was at stake, including the end of the State of Israel itself. The research seeks to explore what is unique in these protest spaces, and the connection between the messages and the new meanings they carry. The study is based on participatory observation, pictures taken, media materials, social media and conversations with protestors. It reaches conclusions about the protests' development, mechanisms, and participants that made it unique.

The protests grew further around Israel, when this paper is written, it is the 27th demonstrations.

---

## Factors Conditioning Variable Assimilation in Korean Nasal-Liquid Sequences

Rok Sim<sup>1</sup>, Drew Crosby<sup>2</sup>, D. Eric Holt<sup>1</sup>

<sup>1</sup>University of South Carolina, Columbia, USA. <sup>2</sup>Korea University, Seoul, Korea, Republic of

This paper investigates variation in the directionality of assimilation in Korean nasal-liquid sequences. The direction of assimilation is complex (Kang 2002; Sohn 2006). Both (1) and (2) involve nasal-liquid sequences;

however, (1), a word composed of two bound morphemes, is realized as [ll] (i.e., regressive assimilation), while (2), a word consisting of a free morpheme and a bound morpheme, is realized as [nn] (i.e., progressive assimilation).

Interestingly, (2), which is expected to only allow progressive assimilation, allows assimilation in both directions as in (3), although the degree of acceptability between (3a) and (3b) differs; [ll] in (3b) is less acceptable than in (3a), perhaps linked with frequency-effects.

Additionally, variation in the directionality of assimilation is observed in English-loanword compounds as in (4). Canonical assimilation in loanword compounds is illustrated in the (i) forms (Sohn 2006), which shows that the nasal-liquid sequence shows progressive assimilation, surfacing as [nn], with regressive assimilation to [ll] in the (ii) forms. We again observe differential acceptability between (4a) and (4b), perhaps due to the syllable count of the first morpheme in the word.

We investigated factors conditioning the variation by performing a production task with seven native Korean speakers, grouped by age (over/under 35). Stimuli consisted of fifteen sentences containing nasal-liquid sequences divided into three groups based on the morphological structure of the target word: ‘bound morpheme + bound morpheme’ (three words), ‘free morpheme + bound morpheme’ (six words), and loanword compounds (six words). The realization was impressionistically determined to be either [nn] or [ll]. To examine social and morphological effects, we built two logistic regression models with realization ([nn]/[ll]) as the dependent variable and age (over/under 35), gender (men/women), and either frequency of the first morpheme (above/less than 50,000 hits in Google search), or syllable-count of the first morpheme (one or two syllables) as independent variables.

Results reveal that words with high-frequency initial-morphemes (e.g., Daehan, 3b) were more likely to surface as [nn] than words with low-frequency initial-morphemes (e.g., Nokeun, 3a) ( $p=0.05$ , Figure 1). Additionally, words with two-syllable morphemes (e.g., /taun/, 4b) favored [nn] assimilation over words with one-syllable morphemes (e.g., /wən/, 4a) ( $p=0.01$ ), and older participants exhibited more [nn] assimilation than younger participants ( $p<0.01$ , Figure 2).

To formally model this variation, we employ the constraint-based phonological model Optimality Theory (OT) (McCarthy & Prince 1995; Prince & Smolensky 2004). In OT, constraints are violable and the output (or ‘optimal’) form is the candidate that only violates lower-ranking constraints. Canonical OT offers no way to account for linguistic variation. However, to account for nasal-liquid variation we employ the OT offshoot Vestige Theory (Kostakis 2010). Vestige Theory posits that in accordance with language change constraints are demoted in the grammar of some speakers but a vestige constraint is left behind that flashes on-off according to factors such as social context, word-frequency, and syllable-count. By doing so, we are able to successfully model the variation in nasal-liquid using the machinery of formal phonology and the methodology of variationist sociolinguistics.

Kang, Hyunsook. 2002. On the optimality-theoretic analysis of Korean nasal-liquid alternations. *Journal of East Asian Linguistics* 11(1): 43-66.

Kostakis, Andrew. 2010. Vestige Theory: Sociolinguistic evidence for output-output constraints. *Lingua* 120: 2476-2496.

McCarthy, John J., and Alan Prince. 1995. Faithfulness and reduplicative identity. In Beckman, Walsh Dickey, and Urbanczyk (eds.), *Papers in optimality theory*, 249–384.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. John Wiley & Sons.

Sohn, Hyang-Sook. 2006. Morphological basis of sonorant assimilation in Korean. *Studies in Phonetics, Phonology and Morphology* 12(2): 321-338.

---

## The meaning and function of code-switching in a mixed language: Language contact and variation in Molise Croatian (Italy)

Lucija Šimičić<sup>1</sup>, Francesca Sammartino<sup>2</sup>

<sup>1</sup>University of Zadar, Zadar, Croatia. <sup>2</sup>University of Zagreb, Zagreb, Croatia

Molise Croatian or Na-našo is a language spoken by migrants from south Dalmatia (Croatia) who settled in the early 16th century in a hilly rural area off the Adriatic coast in today’s southern Italian region of Molise. Despite an almost five hundred-year long geographic and social isolation, due to which the Molise Croatian community has been treated as a language island in the linguistic literature (e.g. Marra, 2019), the community was in close contact with the neighboring Romance population (Babalini et al., 2005) and at least parts of the community were bilingual for a very long time according to some archival material (Rešetar, 1997 [1911]). An intense exposure to Romance varieties and a complete absence of contact with Slavic speakers until the 20th century led to the development of Na-našo as a high-contact variety between Štokavian Ikavian archaic variety of Croatian, south Italian dialects of Abruzzo and Molise, and more recently the standard Italian. Spoken by fewer than a thousand speakers with a limited intergenerational transmission in a community marked by exogamy and depopulation, it is nowadays a seriously endangered language, which nonetheless remains one of the crucial identity markers of the community.

The aim of the paper is to explore the meaning and relevance of the concept of code-switching among speakers in a community in which local linguistic practices are characterized by the simultaneous presence of Na-našo and

Italian, and where the two cannot always be easily separated. Departing from the idea that code-switching is a theoretically useful concept prone to (socio)linguistic analysis only when speakers themselves clearly distinguish the codes involved in their language behavior (Auer, 1998), we analyze the examples of linguistic practices which challenge earlier approaches to code-switching. The analysis is based on small corpora of spontaneous speech in Na-našo, including relevant parts of Pangloss collection and the podcast Čujemo se!, both available on-line. We argue that due to the pervasion of mixed forms at all levels of linguistic structure and an undergoing language shift in the community, the concept of code-switching cannot always be clearly differentiated from other instances of language-contact phenomena in mixed varieties such as Molise Croatian, and the concept of translanguaging (García & Wei, 2014) may be more applicable to account language behavior among the speakers of Na-našo.

Auer, Peter (Ed.) (1998). *Code-switching in conversation: language, interaction and identity*. Routledge.

Babalini, C. et al. (2005). The population history of the Croatian linguistic minority of Molise (southern Italy): A maternal view. *European Journal of Human Genetics*, 13(8), 902–912. <https://doi.org/10.1038/sj.ejhg.5201439>.

García, O., & Wei, L. (2014). *Translanguaging: Language, Bilingualism and Education*. Palgrave Macmillan.

Marra, A. (2019). Isole linguistiche: la comunità degli Slavi del Molise. In: Toso, F. (Ed.), *Lingue sotto il tetto d'Italia. Le minoranze alloglotte da Bolzano a Carloforte*. Treccani online. [https://www.treccani.it/magazine/lingua\\_italiana/articoli/scritto\\_e\\_parlato/Toso13.html](https://www.treccani.it/magazine/lingua_italiana/articoli/scritto_e_parlato/Toso13.html)

Rešetar, M. (1997). *Le colonie serbo-croate nell'Italia meridionale*. Amministrazione Provinciale.

---

### It's not the THOUGHT that counts for a LOT of Michigan English sound changes

Betsy Sneller, Suzanne Wagner, Yongqing Ye, Adam Barnhardt, Jack Rechsteiner, Caroline Zackerman  
Michigan State University, East Lansing, USA

Michigan English, as part of the Inland North, has long been characterized by the Northern Cities Shift (NCS) vowel system in white speakers (Labov et al., 2006; Labov 2010). In recent years, some aspects of the NCS have begun to change in Michigan (Wagner et al., 2016; Nesbitt et al., 2018; Nesbitt 2019). These changes include (1) an allophonic split in TRAP conditioned by nasality, where pre-oral TRAP (traditionally raised and diphthongal in an NCS system) is backed and lowered, and (2) the backing of LOT (traditionally a low front vowel in an NCS system) towards a merger with THOUGHT. These changes are consistent with the emerging “Elsewhere Dialect” found across North America (Becker, 2019, *inter alia*). One major hypothesis is that the Elsewhere Dialect emerges as a response to the LOT-THOUGHT merger.

Here, we present the first results from a large-scale remote recording project (“MI Diaries”), in which participants across the state of Michigan submit self-recorded “audio diaries”. We analyze 68 white speakers who were born and raised in Michigan, ranging in birth year from 1958 to 2015. Figure 1 shows the overall placement of the NCS vowels in normalized F1-F2 space, by speaker generation, hometown population density, and gender. The two changes mentioned above are present in our dataset as well. Figure 2 presents the emerging nasal conditioning of TRAP. Using a bootstrap with replacement method for comparing Pillai scores across speakers with different sample sizes (Stanley & Sneller, 2023), we find that the distance between pre-oral TRAP and pre-nasal HAND is increasing with birth year ( $p=0.03$ ), affirming the findings by Nesbitt (2019) that the canonical NCS raised continuous TRAP is undergoing an allophonic split for younger speakers.

Perhaps more remarkable in our data is the complete absence of a fronted LOT – one of the major hallmarks of the NCS – for any of our speakers; even the Boomer women from Detroit (the population most expected to exhibit a strong NCS pattern) exhibit a back-of-center LOT. We hypothesize that LOT has undergone lifespan change in these diarists, to reflect the changing community norms. This is supported by one comment from a participant (born 1958), who reports: “It’s not uncommon for me to say [...] tom[ɔ]rrow. Where’d that come from? It’s always been tom[a]rrow”. Although LOT is backed in all speakers, we still find younger speakers producing the most backed LOT. Using a bootstrap with replacement method for comparing Pillai scores, we find that the distance between LOT and THOUGHT is decreasing ( $p=0.006$ ) as birth year increases. Despite LOT’s retraction, there are only four speakers in the dataset whose LOT and THOUGHT are statistically merged (represented in red in Figure 2). Together, these findings suggest two major takeaways: (1) it is the phonetic movement of LOT (rather than its phonological merger with THOUGHT) that triggers the Elsewhere Shift, and (2) for LOT to back without also impacting the rest of the NCS suggests that the NCS is no longer a meaningful phonological subsystem in Michigan.

Becker, Kara (ed.). 2019. *The Low-Back-Merger Shift: Uniting the Canadian Vowel Shift, the California Vowel Shift, and short front vowel shifts across North America*. Publication of the American Dialect Society (PADS).

Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology, and Sound Change*. Berlin and New York: Mouton de Gruyter.

Labov, William, 2010. *Principles of Linguistic Change, volume 3: Cognitive and Cultural Factors*. Oxford, UK: Wiley-Blackwell.



- Nesbitt, Monica. 2019. Changing their minds: The impact of internal social change on local phonology. PhD Thesis, Michigan State University.
- Nesbitt, Monica, Suzanne Evans Wagner & Alexander Mason. 2019. A tale of two shifts: Movement toward the Low Back Merger Shift in Lansing, Michigan. In Kara Becker (ed.), *The Low Back Merger Shift: Uniting the Canadian Vowel Shift, the California Vowel Shift, and Short Front Vowel Shifts across North America*. Publications of the American Dialect Society 104. 144—165.
- Stanley, Joseph & Betsy Sneller. 2023. Sample size matters in calculating Pillai scores. *The Journal of the Acoustical Society of America* 153. 54—67.
- Wagner, Suzanne Evans, Alexander Mason, Monica Nesbitt, Erin Pevan & Matt Savage. 2016. Reversal and re-organization of the Northern Cities Shift in Michigan. *University of Pennsylvania Working Papers in Linguistics* 22(2). 171—179.

## Social variation in the use of *traer* and *tener* as verbs of possession in Mexican Spanish

Estefany Sosa

Indiana University Bloomington, Bloomington, US

Possession is a cross-cultural concept that can be difficult to define; furthermore, there is no one to one correspondence between form and meaning (Baron, Herslund & Sørensen 2001). Because possession is essentially the location of an object in space, possessive constructions are closely related to locative and existential constructions (Clark 1978); therefore, it is commonly found across the world's languages the use of location, accompaniment and movement to express possessive relationships (Stassen 2009). *Tener*, the prototypical verb of possession in Spanish, can alternate with various verbs of movement, such as *traer*, to establish possession. In a study of the verb *traer* in Mexican Spanish, Rábago Tánori & Alarcón Neve (2013), argue that the verb *traer* has three main stative functions (possession, affectation and attribution). The researchers determined that *traer* was typically found in constructions designating a physical or mental sensation susceptible to change and concrete possession, while *tener* was found in constructions with abstract meaning that are not susceptible to change. The aim of the present study is to analyze the expanded use of *traer* as a verb of possession in Mexican Spanish (see example (1)-(3)), while considering the sociolinguistic factors that condition its use and alternation with *tener*.

To conduct this pilot study, one thousand tokens taken from the Sociolinguistic Corpus of Mexico City were coded for presence/absence of a determiner, polarity, the content in the NP/DO, the type of possessive relationship, possessor/subject and possessee/object animacy, possessee/object type (e.g., object, body part) mood, person, and time/aspect; the sociolinguistic variables considered in this study were age, sex and socioeconomic level. A mixed-effects logistic regression was used to analyze the data and to determine the factors that predict the use of *traer* or *tener* as verbs of possession in Mexican Spanish.

Preliminary analysis of the data demonstrate that speakers favor the verb *traer* when describing physical possession (possessor and possessee share the same reference interval), while *tener* is favored when describing permanent possession (traditional notion of ownership) and abstract possession (intangible concepts). Overall, it can be concluded that *traer* shares the same temporal extension as the reference interval, while the temporal extension of *tener* extends beyond the reference interval. Furthermore, the use of *traer* as verb of possession is socially stratified as it is favored by working class speakers. A major contribution of this study is the inclusion of sociolinguistic variables in the statistical analysis which demonstrates that the use of *traer* and *tener* as verbs of possession in Mexican Spanish is not only semantically motivated, but also socially stratified.

### Examples

- (1) Yo traía/tenía el sobre del aguinaldo.  
I had the envelope with the bonus (money).
- (2) Traía/tenía el brazo enyesado.  
My arm was in a cast/I had my arm in a cast.
- (3) Traigo/tengo sangre coahuilense.  
I have coahuilense blood (in me).

- Bybee, J. L., Perkins, R. D., & Pagliuca, W. (1994). *The evolution of grammar : tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Baron, I., Herslund, M., & Sørensen, F. (2001). *Dimensions of possession*. Amsterdam: John Benjamins Pub. Co.
- Clark, Eve V. 1978. Locationals: existential, locative and possessive constructions. In Joseph H. Greenberg (ed.), *Universals of human language*, volume 4: Syntax. 85–126. Stanford, CA: Stanford University Press.
- Stassen, Leon. 2009. *Predicative possession*. Amsterdam: John Benjamins.
- Rábago Tánori, Á., & Alarcón Neve, L. J. (2013). Tipos De Relaciones Con Traer, en Construcciones Con Significado Estativo en El Español De México. *Forma y Función*, 26(1), 141–163.
- Rábago Tánori, Á. & Alarcón Neve, L. J. (2015). *Traer estativo en español de México: su estructura sintáctica y propiedades semánticas*. *Cuadernos de Lingüística Hispánica*, 26, 23–47.

---

## Regional Sound Change in a Minoritized Community: Sociophonetics of African American English in Albany, NY

George Stain

Dartmouth College, Hanover, USA

Phonological studies of African-American communities suggest that they participate in regional sound changes to varying degrees (Becker 2014; King 2018; Thomas and Yaeger-Dror 2010). Unfortunately, many such communities remain understudied. This sociophonetic study examines the vowel systems of the African-American community in Albany, NY. This community is located at the intersection of Inland North, Western New England, and Mid-Atlantic (Labov, Ash & Boberg 2006). Albany has not previously been studied from a variationist perspective, but work in the lower Hudson Valley (Dinkin 2009) has shown how these dialect regions interact. I examine (1) the Low Back Merger Shift (LBMS, Becker 2019) as a feature of New England English and (2) THOUGHT-backing (King 2018) and the short-a system (Labov 2007) as a feature of New York City. Using quantitative results and ethnographic analysis as a community member, I show the interactions between ethnic and regional sound changes in Albany and their different social meanings (class, age, and speech style).

### Methods:

This study is based on 36 interviews with African Americans from Albany during the summer of 2022, with birthyears 1941-2004. Interviews include a personal narrative and a wordlist. I processed the data using Montreal Forced Aligner/FAVE-Extract (DARLA) and Lobanov normalization, building best-fit mixed-effects models in Rbrul (Johnson 2009) with birthyear, socio-economic status, parent birthplace, occupation, education, and gender. I examine the status of the LBMS in Albany using the Short-Front-Vowel-Index (SFVI) (Becker, 2019) and Euclidean distance between LOT/THOUGHT vowels. Using novel methods, I examine the competition of NYC and LBMS split short-a systems using Euclidean distances between tense and lax manifestations of TRAP with different phonological conditions (Labov's 2007 short-a and the nasal split).

### Results:

- (a) The normalized Euclidean distance between LOT and THOUGHT (Table 1) and is decreasing as birthyear increases and socioeconomic status increases.
- (b) All the speakers' SFVI are between 1.55 and 2.87 (Fig 1) suggesting that many African American speakers are participating in the LBMS (Becker 2019 reports LBMS present in the SFVI range of 1.85-2.88.)
- (c) There is considerable spread in both Low-Back distance and SFVI, there is no correlation after accounting for age (Table 2). This may be due to interference between the African-American Shift and the LBMS.
- (d) Both LOT and THOUGHT are fronted in careful speech ( $p < 0.0001$ ,  $F_2 +0.128$  and  $+0.121$  in the wordlist, respectively)
- (e) Speakers are transitioning from a NYC short-a split to a LBMS nasal split (Fig 2,  $p < 0.0001$ , birthyear +1 => tense lax short-a distance  $-0.0172$ )

### Theoretical and empirical significance:

This study is the only sociophonetic study of the African American community in Albany and goes against the norm of sociolinguists first studying white communities and only then studying the speech of racial minorities (Alim, Rickford & Ball 2019). These data show that African-American Language scholars need not wait until white communities are studied to examine the speech of African-American communities. This study reveals differential diffusion of speech characteristics among African Americans and white speakers, showing that regionally marked features can be ethnically marked in other regional contexts.

- Becker, Kara. 2014. "Linguistic Repertoire and Ethnic Identity in New York City." *Language & Communication*, New perspectives on linguistic variation and ethnic identity in North America, 35 (March): 43–54. <https://doi.org/10.1016/j.langcom.2013.12.007>.
- . 2019. *The Low-Back-Merger Shift: Uniting the Canadian Vowel Shift, the California Vowel Shift, and Short Front Vowel Shifts Across North America*. Duke University Press.
- Johnson, Daniel Ezra. 2009. "Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis: Mixed-Effects Variable Rule Analysis." *Language and Linguistics Compass* 3 (1): 359–83. <https://doi.org/10.1111/j.1749-818X.2008.00108.x>.
- King, Sharese. 2018. "Exploring Social and Linguistic Diversity across African Americans from Rochester, New York." Ph.D., United States -- California: Stanford University. <http://www.proquest.com/docview/2458749852/abstract/44D8056078045CAPQ/1>.
- Labov, William. 2007. "Transmission and Diffusion." *Language* 83 (2): 344–87.
- Thomas, Erik R. 2007. "Phonological and Phonetic Characteristics of African American Vernacular English." *Language and Linguistics Compass* 1 (5): 450–75. <https://doi.org/10.1111/j.1749-818X.2007.00029.x>.
- Thomas, Erik R., and Malcah Yaeger-Dror. 2010. "AFRICAN AMERICAN ENGLISH SPEAKERS AND THEIR PARTICIPATION IN LOCAL SOUND CHANGES: A COMPARATIVE STUDY INTRODUCTION." *American Speech*, 1–20.

---

## Language Experience Impacts L2 English Scope Computation

LeeAnn Stover

The current study examines how the language experience of Mandarin-English bilinguals affects the computation of English doubly-quantified sentences such as (1), which are ambiguous in English (L2) and unambiguous in Mandarin (L1). Inverse scope readings (1b) are possible for native speakers and advanced learners of English [3, 4], but are argued to be unavailable in Mandarin [5].

(1) *A cat is chasing every mouse.*

a. SURFACE SCOPE  $\exists x [\text{cat}(x) \ \& \ \forall y [\text{mouse}(y) \rightarrow \text{chasing}(x,y)]]$

There is a single cat that is chasing every mouse.

b. INVERSE SCOPE  $\forall x [\text{mouse}(x) \rightarrow \exists y [\text{cat}(y) \ \& \ \text{chasing}(y,x)]]$

For every mouse, there is a (different) cat that is chasing each mouse.

The inverse scope (1b) is more costly to process than surface scope (1a), possibly due to a *Processing Scope Economy (PSE)* principle whereby individuals prefer to parse a construction with the fewest syntactic operations [6]. The PSE predicts that bilinguals should be able to compute a more complex configuration (inverse scope), but they will incur a greater processing cost. In contrast, the *Avoidance of Ambiguity (AA)* principle [1] predicts that Heritage Speakers (HSs), bilinguals with a language background experience involving an early shift in language dominance due to the societal majority language, disprefer ambiguous constructions and thus will not have inverse scope interpretations in their L2 English grammar [2].

This study probes 1) whether HSs pattern more in line with the well-documented *PSE* or the emergent *AA* strategy which is specific to the HS experience, and 2) if individual bilingual dominance differentially impacts the ability to compute inverse scope interpretations of doubly-quantified constructions in L2 English.

**Method.** Forty-three highly proficient bilingual participants completed a language background questionnaire and a forced picture choice task (which doesn't always include a Surface interpretation) using PC Ixex software, which remotely collected three measured variables of Response Time, Picture Selection, and Goodness Rating. The design of this study is 2x3, with Sentence Type (1, *a > every*) vs. (2, *every > a*) crossed with Picture Set (Surface+Inverse; Surface+Distractor; Inverse+Distractor, Fig. 1, with extra objects) (12 items).

### Results and Discussion.

Findings indicated that while bilinguals behaved predictably when a surface image was available, a different pattern emerged when forced to choose between images depicting an inverse interpretation and a non-corresponding distractor image. That is, HSs as a whole and individual highly English-dominant individuals selected the inverse image at a higher rate than Mandarin-dominant participants. Additionally, Mandarin-dominant participants showed a larger gap based on sentence type, presumably due to a larger influence of the L1 Mandarin. Bilinguals across a spectrum of language dominance select the Inverse image at a higher rate than a non-exhaustive Distractor, clearly indicating that HSs as a group and highly proficient L2 English bilinguals as a whole have an inverse interpretation available to them.

(2) *Every cat is chasing a mouse.*

- [1] Polinsky, M., & Scontras, G. (2019). Understanding heritage languages. *Bilingualism: Language and Cognition*, 1–17. <https://doi.org/10.1017/S1366728919000245>
- [2] Scontras, G., Polinsky, M., Tsai, C.-Y. E., & Mai, K. (2017). Cross-linguistic scope ambiguity: When two systems meet. *Glossa: A Journal of General Linguistics*, 2(1), 1- 28. <https://dx.doi.org/10.5334/gjgl.198>
- [3] May, R. C. (1978). *The grammar of quantification*. [PhD Thesis]. Massachusetts Institute of Technology.
- [4] Chu, C.-Y., Gabriele, A., & Minai, U. (2014). Acquisition of quantifier scope interpretation by Chinese-speaking learners of English. *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, 157– 168.
- [5] Huang, C. J. (1982). Move WH in a language without WH movement. *The Linguistic Review*, 1(4), 369–416. <https://doi.org/10.1515/tlir.1982.1.4.369>
- [6] Anderson, C. (2004). *The structure and real-time comprehension of quantifier scope ambiguity* [PhD Thesis]. Northwestern University Evanston, IL.
- [7] Zehr, J., & Schwartz, F. (2018). *PennController for Internet-Based Experiments (IBEX)*. <https://doi.org.17605/OSF.IO/MD832>

---

## Do women really use intensifiers more frequently than men? A cross-linguistic and cross-variety study

James Stratton

University of British Columbia, Vancouver, Canada

While intensifiers in English have been explored widely (e.g., Ito & Tagliamonte, 2003; Tagliamonte, 2008; Fuchs, 2017), the social conditioning of intensifiers in other languages has received considerably less attention. In *Language and Women's Place*, Lakoff (1975) claimed that women use intensifiers and hedges more frequently than men. Although several studies on English have found that women often gravitate towards intensification (e.g., Fuchs, 2017), work on German (Stratton, 2020) suggests that men and women may favor different types of intensifiers. Intensifiers that scale upwards from an assumed norm are called “amplifiers” (e.g., very, really) and intensifiers that scale downwards are called “downtoners” (e.g., a bit, kind of, somewhat). Stratton (2020) found that German women use amplifiers more frequently than men, but men use downtoners more frequently than women. To test the broad applicability of claims about the relationship between gender and intensification, both cross-linguistic and cross-dialectal data are required.

The present study explores this relationship using language data from eight speech communities across four languages: English (Liverpool, Yorkshire), German (Berlin, Hesse), Swabian (Stuttgart, Gmünd), and Norwegian (East Oslo, West Oslo). Data were collected from recordings of naturally occurring conversations and sociolinguistic interviews. Following previous research, the variable context was circumscribed to intensifiable adjectives and each intensifiable adjective was coded for the absence or occurrence of intensification. After defining the envelope of variation, over 10,000 intensifiable adjectives were examined.

Each adjective was also coded for intra- and extralinguistic factors: speech community, language, age, gender, education, syntactic position of the adjective, semantic classification of the adjective, data collection (e.g., sociolinguistic interview vs. corpus data). Data were subjected to a series of generalized linear regressions (Bates et al., 2015) and mixed effects logistic regressions in Rbrul (Johnson, 2009), with the speaker run as a random intercept.

Tentative results suggest that across all four languages women have a statistical tendency to use amplifiers more frequently than men, but men use downtoners more frequently than women. In line with the Dominance Approach (Lakoff, 1975), one interpretation of these findings is that women use amplifiers to make up for a power differential relative to men. In contrast, men may use downtoners to project masculinity, nonchalance, and indifference, traits that are often reinforced in Western societies. Men and women may therefore use different discourse strategies, such as scaling their speech upwards and downwards to convey different types of social meaning.

In addition to gender, other predictors were also at play. Speakers in more urban environments (e.g., Stuttgart, Berlin) exhibited higher rates of intensification, suggesting that individuals in urban areas feel the pressure to intensify more frequently than speakers in rural locations. Younger speakers also had higher rates of intensification than older speakers, suggesting that the need to intensify becomes less important as we age. Tentative comparisons of data from all four languages suggest that many women intensify more in Norwegian, German, and Swabian than in English, pointing towards the importance of exploring non-English varieties.

Fuchs, R. (2017). Do women (still) use more intensifiers than men? Recent change in the sociolinguistics of intensifiers in British English. *International Journal of Corpus Linguistics*, 22, 345–374.

Ito, R., & Tagliamonte, S. (2003). Well weird, right dodgy, very strange, really cool: Layering and recycling in English intensifiers. *Language in Society*, 32, 257–279.

Lakoff, R. (1975). *Language and woman's place*. Cambridge: Cambridge University Press.

Stratton, J. (2020). Adjective intensifiers in German. *Journal of Germanic Linguistics*, 32(2), 183-215.

Tagliamonte, S. (2008). So different and pretty cool! Recycling intensifiers in Canadian English. *English Language and Linguistics*, 12: Special issue on English intensifiers, ed. by Belén Méndez-Naya, 361–394.

---

## Prevelar Raising and Merger in California: A Sound Change In Progress

Riley Stray, Georgia Zellou

University of California, Davis, Davis, USA

Prevelar raising is a highly variable sound pattern that is characterized by the movement of TRAP and/or DRESS vowels before voiced velar /g/. Prevelar raising is a phonological sound pattern that is motivated by phonetic coarticulation; lifting of the tongue root in anticipation of the following velar consonant can result in perceived raising and fronting (Baker et al., 2008; Purnell, 2008; Zeller, 1997). Phonologized BAG-raising is most commonly found in Washington State, the Upper Midwest, and much of Canada, whereas LEG-raising is often distributed farther south (Freeman, 2021; Fridland & Kendall, 2017; Labov et al., 2005; Stanley, 2022). The highly variable

movement of these front vowels has resulted in some speakers being merged in production and/or perception in any combination of the VAGUE, LEG, and BAG vowels (Freeman, 2023; Sullivan, 2020; Zeller, 1997). Much recent work suggests that BAG-raising, in particular, is undergoing a reversal in the Pacific Northwest with younger speakers producing the phonologically shifted BAG less than middle-aged and older speakers (Becker et al., 2016; Freeman, 2021, 2023; McLarty et al., 2016; Stanley, 2018). The current study expands on work in the Pacific Northwest to explore the prevalence and distribution of prevelar raising in California. This study utilizes Pillai scores to evaluate overlap of vowel categories within individual talkers as the primary metric to examine California as a whole; subsequent linear regression models can evaluate differences between individual speakers' scores across regions. The data come from 29 typical-college-aged speakers, balanced for region of origin, who produced prevelar and non-prevelar vowels from a word list in a sound-attenuated booth. Results show generally high TRAP-BAG overlap for speakers from all regions, suggesting low rates of BAG-raising across all of California, corroborating other work in the American West (Becker et al., 2016; Freeman, 2014, 2023; Fridland & Kendall, 2017; Gunter et al., 2017; McLarty et al., 2016). However, LEG-raising and VAGUE-LEG merger showed much greater variation between speakers. A plurality of talkers maintained distinctions between VAGUE and LEG, supported by generally high Pillai scores, yet many talkers had intermediate Pillai scores with three speakers even showing a merger in production, as determined by Stanley & Sneller's (2023) formula for determining statistically significant merger in Pillai scores. Additionally, the study found evidence for differences between regions, with Northern Californians leading a separation of LEG from DRESS,  $p < 0.001$ , and a VAGUE-LEG merger,  $p = 0.0536$ , bolstered by the fact that all three of the merged talkers were Northern Californians. Overall, these results suggest an ongoing sound change that is most advanced in Northern California but indeed progressing in all three regions of California. Additionally, the finding that 10 of the 29 speakers have high phonetic overlap supports a model of gradual sound change. These 10 speakers might just produce partially overlapping vowels or perhaps the speakers shift in some words but not others; either way, the current study finds evidence favoring a model of sound change in which individual speakers are not dichotomous but somewhere on a spectrum.

- Baker, A., Mielke, J., & Archangeli, D. (2008). More Velar than /g/: Consonant Coarticulation as a Cause of Diphthongization. In C. B. Chang & H. J. Haynie (Eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics* (pp. 60–68). Cascadilla Proceedings Project.
- Becker, K., Aden, A., Best, K., & Jacobson, H. (2016). Variation in West Coast English: The Case of Oregon. *Pub. of the Am. Dialect Society*, 101(1), 107–134.
- Freeman, V. (2014). Bag, Beg, Bagel: Prevelar Raising and Merger in Pacific Northwest English. *University of Washington Working Papers in Linguistics*, 32.
- Freeman, V. (2021). Vague eggs and tags: Prevelar merger in Seattle. *Language Variation and Change*, 33(1), 57–80. <https://doi.org/10.1017/S0954394521000028>
- Freeman, V. (2023). Production and perception of prevelar merger: Two-dimensional comparisons using Pillai scores and confusion matrices. *Journal of Phonetics*, 97, 101213. <https://doi.org/10.1016/j.wocn.2023.101213>
- Fridland, V., & Kendall, T. (2017). Speech in the Silver State. *The Publication of the American Dialect Society*, 102(1), 139–164. <https://doi.org/10.1215/00031283-4295222>
- Gunter, K., Clayton, I., & Fridland, V. (2017). Pre-velar raising and categorization in Nevada English. *NorthWest Phonetics and Phonology Conference*, Vancouver, BC.
- Labov, W., Ash, S., & Boberg, C. (2005). *The Atlas of North American English: Phonetics, Phonology and Sound Change*. De Gruyter Mouton. <https://doi.org/doi:10.1515/9783110167467>
- McLarty, J., Kendall, T., & Farrington, C. (2016). 7. Investigating the Development of the Contemporary Oregonian English Vowel System. *The Publication of the American Dialect Society*, 101(1), 135–157. <https://doi.org/10.1215/00031283-3772934>
- Purnell, T. C. (2008). Prevelar Raising and Phonetic Conditioning: Role of Labial and Anterior Tongue Gestures. *American Speech*, 83(4), 373–402. <https://doi.org/10.1215/00031283-2008-028>
- Stanley, J. A. (2018). Changes in the Timber Industry as a Catastrophic Event: BAG-Raising in Cowlitz County, Washington. *University of Pennsylvania Working Papers in Linguistics*, 24(2). <https://repository.upenn.edu/pwpl/vol24/iss2/16>
- Stanley, J. A. (2022). Regional Patterns in Prevelar Raising. *American Speech*, 97(3), 374–411. <https://doi.org/10.1215/00031283-9308384>
- Stanley, J. A., & Sneller, B. (2023). Sample size matters in calculating Pillai scores. *The Journal of the Acoustical Society of America*, 153(1), 54–67. <https://doi.org/10.1121/10.0016757>
- Sullivan, L. (2020). The production-perception link in phonologically-conditioned pre-velar /æ/-raising [Generals Paper 1]. *University of Toronto*.
- Zeller, C. (1997). The Investigation of a Sound Change in Progress: /Ae/ to /e/ in Midwestern American English. *Journal of English Linguistics*, 25(2), 142–155.

---

## Monophthongal /ow/ among Nordic Americans in Puget Sound: Language Contact in the Development of Regional Dialect Features

Julia Swan, Shruti Piduri, Hannah Wells  
San José State University, San Jose, USA

When describing distinctive ethnolinguistic repertoires of English, sociolinguists often posit substrate influence. A feature may be introduced by speakers of a non-English community language and subsequently used by English dominant speakers with identity or stance motivations (Fought 2010). How exactly this happens has been difficult to trace. Ethnolinguistic features may derive from the community or heritage language but reallocated to other sociolinguistic meanings in the second generation (Gnevsheva 2020). Older and younger individuals with the same generational status deploy ethnolinguistic variation differently depending on timing of birth relative to salient cultural events in the community (Sharma & Sankaran 2011). Features associated with ethnolinguistic repertoires may evolve to become markers of regional identity or local stances (Labov 1963). This project documents the complexity of contact-induced language change, aligning with research that documents the influence of non-English languages on the development of regional dialects.

Monophthongal /ow/ is a notable feature of the English in Washington State, along with other parts of the Upper Midwest (Wassink 2015). The current acoustic analysis compares this feature among 30 first, second and third-generation Swedish and Norwegian immigrants to the Puget Sound area born in the 1920s to 1940s and 10 non-Scandinavian contemporaries. The interviews were collected as part of the Nordic American Voices and the Speaking of Seattle Oral History Projects by the Nordic Heritage Museum and the Museum of History and Industry in Seattle, WA. Analyzing a combined total of over 5,000 /ow/ tokens using a trajectory length measure and visualizations of formant trajectories, the paper describes the focusing and reallocation of monophthongal /ow/ among children and grandchildren of Nordic American immigrants.

Previously presented work on this project demonstrates that first generation immigrants exhibit more monophthongal /ow/ as indicated by significantly shorter trajectory length measures and trajectory shape differences. Subsequent generations of Norwegian and Swedish Americans display refunctionalization by following voicing environment; monophthongal /ow/ is restricted to the pre-voiceless environment while open-syllable tokens of /ow/ are relatively more diphthongal. A similar process of focusing and reallocation of dialect variants in contact led to a phonologized raising pattern in Fenland English (Britain and Trudgill 2003: 251). The current analysis extends this finding by exploring how the Scandinavian immigrants and their descendants used monophthongal /ow/ differently than their local contemporaries who were not of Scandinavian origin. The findings reveal that the contextual difference in trajectory length between phonological environments is larger for second-generation immigrants than for non-Scandinavian local contemporaries, showing a tendency to monophthongize in pre-voiceless environments. Notably, this pattern of pre-voiceless monophthongal /ow/ is an established feature of modern young adult talkers in Seattle (Swan 2016) suggesting the pattern may have originated or been reinforced by large-scale Scandinavian immigration in the early 20th century. The results shed light on the timing and origins of monophthongal /ow/ as a dialect feature of Seattle English and may provide insight into the feature's socioindexical meaning.

- Britain, David & Trudgill, Peter. (2003). Migration, new-dialect formation and sociolinguistic refunctionalisation: Reallocation as an outcome of dialect contact. *Transactions of the Philological Society*, 97, 245 - 256. 10.1111/1467-968X.00050.
- Fought, C. (2010). Ethnic Identity and Linguistic Contact. In *The Handbook of Language Contact*, R. Hickey (Ed.). <https://doi.org/10.1002/9781444318159.ch14>
- Gnevsheva, K. (2020). The role of style in the ethnolect: Style-shifting in the use of ethnolectal features in first- and second-generation speakers. *International Journal of Bilingualism*, 24(4), 861–880. <https://doi.org/10.1177/1367006920902520>
- William Labov (1963). The Social Motivation of a Sound Change, *WORD*, 19:3, 273-309, DOI: 10.1080/00437956.1963.11659799
- Sharma, D. & Sankaran, L. (2011). Cognitive and social forces in dialect shift: Gradual change in London Asian speech. *Language Variation and Change*. 23. 399 - 428. 10.1017/S0954394511000159.
- Swan, J. T. (2016). Language ideologies, border effects and dialectal variation: evidence from /æ/, /aʊ/ and /aɪ/ in Seattle, WA and Vancouver, B.C. Ph.D. Thesis, University of Chicago, August 2016.
- Wassink, A. (2015). Sociolinguistic Patterns in Seattle English. *Language Variation and Change*, 27(1), 31-58. doi:10.1017/S095439451400023

---

## The '1984' of linguistic change: A sociolinguistic shock point in the late 20th century

Sali Tagliamonte

University of Toronto, Toronto, Canada

In this paper, I synthesize across decades of research studies on language change in progress to offer insight into the language/society interface. The data come from spoken corpora of North American English from individuals born from the late 1800's to the early 2000's. The linguistic variables under consideration range from the choice of lexis variants, 3rd person nouns (guy), to adjectives (awesome), to tense/aspect (going to) to discourse-pragmatic (oh my God). Compare the variants in (1) to (2).

1. Oh my God, that guy is going to have an awesome time.
2. Golly, that fella will have a terrific time.

I model the incoming variants and the social broad characteristics of the individuals who use them, their date of birth and gender, key factors underlying the principles of linguistic change in progress (Labov 2001). The methodology employed is founded in comparative sociolinguistics (Poplack & Tagliamonte 2001) and employs triangulated statistical techniques in R (2007). Following Tagliamonte & Baayen (2012) and Gries (2018) conditional inference tree analysis to identify the temporal check points, random forests to substantiate the importance of the temporal junctures across variables and mixed effects modelling (GLMER) to corroborate statistical significance while controlling for interactions and the random effect of individual (e.g. Tagliamonte & Baayen).

The discovery is a repeating pattern in the trajectory of the 20th century among individuals born in the early 1980's. For every variable, there is a striking escalation, regardless of its stage of development (Tagliamonte & D'Arcy 2009:99). This shock point cannot be due to grammar internal factors because every variable exhibits this development and no one system has anything in common with all the rest. The driving force could only be external to the linguistic system. I appeal to the principle of interaction. As Labov (2001: 463) has argued "there must be a social force that activates the shift and drives the increment. The question is what is that force?"

The early 1980's were a particularly critical point on multiple counts, including technological advancement, globalization and media influence, economic and political change, language contact and the rise of digital natives (Prensky 2001). New media made face-to-face conversation possible regardless of location. Written language expanded exponentially into multiplex registers mixing features of spoken and written communication. The new modes of interaction disrupt traditional pathways of language diffusion because none of them requires mobility. Moreover, the social contexts of online worlds, games and their concomitant practices, styles and voices often involve interaction between people who have sparse social network connections. I will argue that the generation born in the 1980's were positioned at time when communication was expanding beyond geographic boundaries to global spaces leading not only to the uptake of innovative linguistic variants but also an unprecedented acceleration of the changes in progress at the time. Taken together these findings highlight the importance of studying language in social context and suggest that sociolinguistic research is key to understanding future developments in societies around the world.

Gries, S. T. (2018). On classification trees and random forests in corpus linguistics: Some words of caution and suggestions for improvement. *Corpus Linguistics and Linguistic Theory*. doi:<https://doi.org/10.1515/cllt-2018-0078>

Labov, W. (2001). Driving forces in linguistic change. Paper presented at the 2002 International Conference on Korean Linguistics, August 2, 2002, Seoul National University, South Korea.  
<http://www.ling.upenn.edu/~wlabov/Papers/DFLC.htm>.

Labov, W. (2001). *Principles of linguistic change: Volume 2: Social factors*. Malden and Oxford: Blackwell Publishers.

Prensky, M. (2001). *Digital Natives Digital Immigrants*. In *On the Horizon* (Vol. 9). MCB University Press.

Tagliamonte, S. A., & Baayen, R. H. (2012). Models, forests and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, 24(2), 135-178.

Tagliamonte, S. A., & D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language*, 85(1), 58-108.

---

## Pitch variability cues perceptions of Singlish: A perceptually-guided approach to sociophonetic variation

Yin Lin Tan<sup>1,2</sup>, Ting Lin<sup>1</sup>, Meghan Sumner<sup>1</sup>

<sup>1</sup>Stanford University, Stanford, USA. <sup>2</sup>National University of Singapore, Singapore, Singapore

Descriptions of Singlish, a colloquial variety of Singapore English, have identified its lexical, syntactic, and phonological differences compared to standard English. Although there is no clear consensus regarding what counts as Singlish to the average Singaporean, different models have been proposed to understand variation in how English is spoken in Singapore. Some posit a binary distinction between Singlish and standard Singapore English (Gupta

1994), while others locate them along a continuum (Alsagoff 2007). Calls have also been made for a variationist, feature-based approach (Leimgruber 2012), which requires linguists to identify relevant features and their social meanings. Using a perceptually-guided approach for prosodic features, we argue that pitch variability cues perceptions of Singlish and is important for an indexical account of Singapore English.

This study investigates what phonetic features cue perceptions of Singlish using a speeded forced-choice task. By probing participants' percepts of Singlish instead of a predetermined linguistic form, this study allows listeners, rather than linguists, to identify salient features. The task consisted of six blocks of 20 trials each. The stimuli comprised 40 natural-speech audio clips from ten Singaporean speakers obtained from podcasts. In each trial, participants (N=132) heard two randomly selected clips and had two seconds to select which clip sounded more Singlish before the next trial began. Each clip was one intonation phrase, 1.4-2.6 seconds long, syntactically and lexically similar to standard English, and controlled for semantic content. A follow-up questionnaire collected information about participants' sociolinguistic background and language attitudes. Participants' descriptions of more Singlish-sounding speakers focused on social attributes (e.g., rough, casual) and prosody (e.g., fast, monotonous, variety in intonation). Using these descriptions, we measured the articulation rate (syllables per second), pitch variance (standard deviation of mean semitone of each vowel), pitch Pairwise Variability Index (PVI; comparisons of adjacent vowels' maximum semitones), and durational PVI (comparisons of adjacent vowels' durations) of each clip.

A logistic mixed effects regression model showed that clips rated as more Singlish were associated with lower pitch variance ( $\beta=-0.18, p=0.023$ ) and higher pitch PVI ( $\beta=0.16, p=0.035$ ). Articulation rate was marginally positively correlated with Singlish ( $\beta=0.19, p=0.056$ ); durational PVI was not significant, contrary to previous work on Singapore English (Low et al. 2000). The correlation between Singlish and higher pitch PVI aligns with descriptions of adjacent tones in Singlish as stepping up or down, rather than "[gliding] more gradually from one pitch level to another" (Lim 2011: 277). These observations suggest that pitch variability is a primary prosodic cue to listeners' perceptions of Singlish. Crucially, Singlish is associated with more local pitch variability but less global pitch variability.

The findings indicate that pitch variability cues perceptions of Singlish and provide support for analyses of pitch variability in future variationist accounts of Singapore English. This study also emphasizes the utility of speeded forced-choice tasks in investigating sociophonetic variation. Future work can disentangle the finer-grained social meanings that pitch variability might index and examine how pitch variability interacts with other linguistic features.

Alsagoff, Lubna. 2007. Singlish: Negotiating capital, culture and identity. In Vinita Vaish, S. Gopinathan and Yongbing Liu (eds.), *Language, capital, culture*, 25-46. Rotterdam: Sense Publishers.

Gupta, Anthea Fraser. 1994. *The step-tongue: Children's English in Singapore*. Clevedon: Multilingual Matters.

Leimgruber, Jakob. 2012. Singapore English: An indexical approach. *World Englishes* 31(1). 1-19.

Lim, Lisa. 2011. Tone in Singlish: Substrate features from Sinitic and Malay. In Claire Lefebvre (ed.), *Creoles, their substrates, and language typology*, 271-288. John Benjamins Publishing.

Low, Ee Ling, Esther, & Nolan, Francis. 2000. Quantitative characterisations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech* 43(4). 377-401.

---

## Multiple Ways to Do Authenticity: A Case Study of a Diasporic Speaker

Marie Tano

Stanford University, Stanford, USA

Sociolinguistic investigations into the role of default, vernacular varieties in the conveyance of personal honesty have often ascribed such linguistic styles to a speaker's first learned variety (Sharma, 2018, 2021). The current study expands upon current understandings of biographical indexicality in order to consider how a single speaker indexes intimacy through the use of linguistic styles that do not solely reflect their L1.

My analysis focuses upon Jermaine, a multilingual immigrant with a complex racial identity that reflects his Nigerian and American upbringing. In this paper, I investigate his use of three features that are representative of three particular ethnicized varieties, African-American English (AAE), Nigerian English (NigE) and Nigerian Pidgin English (NigP), in order to understand how he uses language to reflect intersectional identification practices. In analyzing his use of copula absence (AAE feature), alveolar trills (NigP/NigE feature) and DH-stops (ubiquitous), all instances of rhotic phonemes (N = 213), words with (DH) (N=599), and copula environments (N=613) were extracted from five different self-recorded interactions with his Nigerian mom, American close friend and peers on an academic panel.

While realizations of copula (absence, contracted or full) were noted manually, the manner of articulation for each (DH) variable (fricative, deleted or stop) and rhotic segment (approximant, trill or deleted) was recorded, based upon spectrographic and waveform analysis. Figures 1-3 display the variability of features across speaking situations, which were revealed to have a significant effect on their realization by mixed-effects logistic regression



models. While this variation reflects a slight effect of audience, the variability within addressee contexts (DH in particular) also suggests interactional effects.

My discourse analyses of the specific speaking situations demonstrated Jermaine's deployment of the features at moments of ethnic authentication, personal honesty, and social alignment. Specifically, his specific recruitment of trills cross-situationally in linguistically-diverse but face-threatening moments, represents a reversion to his first-learn vernacular, NigE. At the same time, Jermaine's use of copula absence with features of NigE, reflects a natural shift towards AAE. As Jermaine appears to deploy his full repertoire at any moment, disregarding linguistic or perceived ethnic boundaries, his use of DH-stops further demonstrates the fluid nature of his racial identity.

While a traditional approach may presume that grounding an individual feature in an utterance of an ethnicized variety allows for one to index a particular ethnicity; I argue that Jermaine's full deployment of his repertoire reflects an intersectional personae, where he appears to reflect both his American and Nigerian heritage at any given moment. The current analysis posits that transnational speakers like Jermaine are able to completely deploy their broad linguistic repertoire in order to index authentic, pluralistic identities. In this way, Jermaine represents a common Black immigrant experience, where such individuals may not regard an African-American and African ethnic identity as mutually exclusive realities (Habecker, 2017).

Habecker, S. (2017). Becoming African Americans: African immigrant youth in the United States and hybrid assimilation. *Journal of Pan African Studies*, 10(1), 55.

Sharma, D. (2018). Style dominance: Attention, audience, and the 'real me'. *Language in Society*, 47(1), 1–31.

Sharma, D. (2021). Biographical indexicality: Personal history as a frame of reference for social meaning in variation. *Social Meaning and linguistic variation: Theorizing the third wave*, 243–264.

---

## Variable Equivalence for bilingual code-switching: Spanish-English clause combining

Rena Torres Cacoullos<sup>1</sup>, Dora LaCasse<sup>2</sup>

<sup>1</sup>Penn State University, University Park, USA. <sup>2</sup>University of Montana, Missoula, USA

Bilingual code-switching (CS)—going back and forth between languages—is generally agreed to be constrained by equivalence or congruence of some kind between the two languages (e.g. Muysken, 2000:27). As formalized in the Equivalence constraint, CS is permitted between analogous elements that are locally ordered the same way in the two languages (Poplack 1980:586; Sankoff 1998:46). Nevertheless, CS constraints are probabilistic: among permissible sites, bilinguals display preferences to switch at particular ones (cf. Kootstra 2015: 50–51). For example, in Spanish-English main-and-complement clauses, Spanish complementizer "que" is preferred over the corresponding English variant, "that". This asymmetry is not due to language dominance nor to a matrix language, however. The preference is accounted for by the Variable Equivalence hypothesis, which states that bilinguals favor CS with the equivalent option from one of the languages that is more frequent and predictable in their cumulative combined linguistic experience, considering both languages (Torres Cacoullos 2020).

We compare bilingual clause combining for complement versus adverbial (causal) clauses, to test Variable Equivalence. Spanish-English equivalence exists at the main-and-adverbial clause junction, inasmuch as the conjunctions are consistently present and placed the same way in the two languages. Equivalence is variable, however, with main-and-complement clauses, due to variation in the presence of English "that", unlike categorically-present Spanish "que". Over 2,000 tokens of clause combining were coded for conjunction, subordinate clause type and, for instances of CS between clauses (n=189), direction of CS, that is, going from a Spanish main verb to an English subordinate verb (1a-2a), or the reverse, from an English main to a Spanish subordinate (1b-2b). Choice between causal conjunctions "porque" and "(be)cause" is affected by CS direction (1). In contrast, the asymmetry in favor of the Spanish complementizer applies regardless of CS direction (2) (Figure 1, Table 1). Since Spanish "que" is invariable, in bilinguals' combined inventory, it is the more frequent option compared with English complementizer "that". Since "that" presence vs. absence is conditioned probabilistically (e.g., Shank et al. 2016: 202-213), "que" is also more predictable than "that". Accounting for bilinguals' CS preferences, the Variable Equivalence hypothesis brings linguistic variation and experience into CS constraints.

Kootstra, G. J. (2015). A psycholinguistic perspective on code-switching: Lexical, structural, and socio-interactive processes. In

G. Stell & K. Yakpo (Eds.), *Code-switching between structural and sociolinguistic perspectives*, 39-65. De Gruyter.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

Poplack, S. (1980). "Sometimes I'll start a sentence in Spanish y termino en español": toward a typology of code-switching. *Linguistics* 18, 581-618.

Sankoff, D. (1998). A formal production-based explanation of the facts of code-switching. *Bilingualism: Language and Cognition* 1(1), 39-50.

Shank, C., Plevoets, K., & Van Bogaert, J. (2016). A multifactorial analysis of that/ zero alternation: the diachronic development of the zero complementiser with think, guess and understand. In J. Yoon & S. T. Gries (Eds.), *Corpus-Based Approaches to Construction Grammar*, 201–240. John Benjamins.

## Determiner Drop in Vernacular English

Christina Tortora<sup>1</sup>, Carolina Fraga<sup>2</sup>, William Oliver<sup>3</sup>

<sup>1</sup>CUNY (CSI & The Grad Center), New York, USA. <sup>2</sup>CUNY (The Grad Center), New York, USA. <sup>3</sup>Stony Brook University, Stony Brook, USA

Our talk examines a phenomenon we label “determiner drop,” exhibited in the vernacular English sentences in (1):

- |                                  |                                      |
|----------------------------------|--------------------------------------|
| (1a) Doctor came to visit us.    | (2a) The doctor came to visit us.    |
| (1b) Guy’s never seen Star Wars. | (2b) The guy’s never seen Star Wars. |
| (1c) Air is too dry.             | (2c) The air is too dry.             |

The sentences in (1a-c) contain a matrix subject NP which has no overt determiner (D), despite the fact that they are truth-conditionally equivalent to the overt D sentences in (2a-c). This form of **determiner drop in subject position in declarative sentences** (henceforth det-drop) is a property of vernacular English more generally, and not restricted to any one variety. For instance, det-drop exists in African American English as discussed in Spears (2007), as well as in Appalachian English as discussed in Oliver (2022) in his analysis using the Audio-Aligned and Parsed Corpus of Appalachian English (AAPCAppE, Tortora et al. 2017). Our study uses the AAPCAppE to explore the syntactic distribution of det-drop sentences, and combines our corpus findings with judgment data to analyze their semantic and syntactic properties.

**Data.** Despite the truth-conditional equivalence of the sentences in (1) and (2), we find that there are several restrictions on det-drop sentences; here we discuss a few. First, they cannot be **embedded**, as illustrated in (3); this is corroborated by both a complete absence of embedded det-drop in the corpus and our own judgment data.

- (3) \*Maria said [doctor came to visit us].

Second, det-drop sentences are **evidential**. This is most apparent in universal and experiential perfects:

- (4a) The guy has never seen Star Wars.  
(4b) Guy has never seen Star Wars.

Sentence (4a) can be used simply to describe a property of the guy, without the speaker needing any direct evidence for its truth in order to assert it. In contrast, sentence (4b) has the further requirement that the guy in question is exhibiting some behavior evident to the speaker which leads the speaker to assert the proposition’s truth.

Third, det-drop sentences have a discourse function we characterize as “**narrative**” (Önnerfors 1996; Oliver 2022) or as involving “**no call on the addressee**” (Gärtner 2016; Portner et al. 2019). This is exhibited in det-drop’s robust use in narratives, which we argue require no interlocutor, by definition. It is also exhibited by their incompatibility with discourse markers which entail a common ground between speaker and addressee, like *well* (Jenkins 1993):

- (5a) Guy didn’t like the movie.  
(5b) \*Well, guy didn’t like the movie. (cf.: Well, the guy didn’t like the movie.)

Sentences like (5a) are only felicitous in contexts where there is no call on an addressee, a property we also see in exclamatives (*What a beautiful day!*) and certain types of commands (*No feeding the animals*), as well as in written subject omission (WSO: *Started raining this morning*; Haegeman 2019). Sentence (5b) is infelicitous because *well* suggests a common ground where the speaker’s claim is under discussion with an addressee (and in which the speaker either counters or anticipates the addressee’s possible contradiction).

Fourth, **det-drop in conditionals entails the absence of the conditional element *if***. To illustrate: first note the entirely independent fact that standard conditionals like (6a) exist alongside a vernacular counterpart in which the complementizer *if* is dropped (6b):

- (6a) If the plane arrives late, I’ll take a cab.  
(6b) The plane arrives late, I’ll take a cab.

Interestingly, in the context of vernacular *if*-drop as in (6b), the determiner can also be dropped, as in (6c):

- (6c) Plane arrives late, I’ll take a cab.

Crucially, however: det-drop is not possible in the presence of *if*:

(6d) \*If plane arrives late, I’ll take a cab. The data in (6) suggest that if the determiner is absent, then elements in the clausal spine above it must also be absent.

**Analysis.** We argue that all of these properties of det-drop sentences follow (a) if D is taken to be merged in the clausal spine above vP, à la Sportiche (2005) (or merged late, à la Thoms 2019), and (b) if det-drop structures involve truncation of the part of the higher functional field that otherwise includes the merged determiner. Our analysis allows us to unify det-drop in vernacular English with apparently unrelated phenomena cross-linguistically, such as: (i) V1 word order in narrative declaratives in German, which Önnerfors (1996) independently argues is the

result of truncation; (ii) det-drop in Hungarian, which Halm (2021) also independently argues is the result of truncation; (iii) the absence of discourse particles (such as *doch*) in German infinitival root clauses, which Gärtner (2016) attributes to truncation and to the absence of a “call on the addressee.” Our study thus confirms the value of quantitative and qualitative analyses of vernacular English structures.

- Gärtner, Hans. 2016. Root infinitivals and modal particles. An interim report. In J. Bayer & V. Struckmeier: *Discourse Particles: formal approaches to their syntax and semantics*, pp. 115- 143;
- Haegeman, Liliane. 2019. Register-Based Subject Omission in English and its Implication for the Syntax of Adjuncts. *French Journal of English Linguistics* 28;
- Halm, Tamás. 2021. Radically Truncated Clauses in Hungarian and Beyond: Evidence for the Fine Structure of the Minimal VP. *Syntax* 24:3;
- Jucker, Andreas. 1993. The discourse marker *well*. *Journal of Pragmatics* 19: 435-452;
- Oliver, William. 2022. *Covert Determiners in Appalachian English Narrative Declarative Sentences*. MA thesis, CUNY Graduate Center; Önerfors,
- Olaf. 1996. On narrative declarative V1-sentences in German. In Toril Swan & Olaf J. Westvik: *Modality in Germanic Languages*. pp. 293-319., De Gruyter;
- Portner, P., M. Pak, & R. Zanuttini. 2019. The speaker-addressee relation at the syntax-semantics interface. *Language* 95.1: 1-36;
- Spears, Arthur. 2007. “Bare Nouns in African American English (AAE).” In M. Baptista & J. Guéron: *Noun Phrases in Creole Languages*, pp. 421-434;
- Sportiche, Dominique. 2005. *Division of Labor between Merge and Move: Strict Locality of Selection and Apparent Reconstruction Paradoxes*. Ms, UCLA;
- Thoms, Gary. 2019. *Antireconstruction as layering*. In NELS 49, Cornell; Tortora, Christina, Beatrice Santorini, Frances Blanchette, & C.E.A. Diertani. 2017. *The Audio-Aligned and Parsed Corpus of Appalachian English (AAPCAppE)*, version 0.1. URL: [www.aapcappe.org.2](http://www.aapcappe.org.2)

---

## Better Late than Never: Bilingual Language Processing of Evidentiality in a Turkish-English Translation Task

Sumeyra Tosun<sup>1</sup>, Luna Filipovic<sup>2</sup>

<sup>1</sup>Medgar Evers College, Brooklyn, USA. <sup>2</sup>University of California, Davis, Davis, USA

Evidentiality is a linguistic property to encode the source of knowledge, indicating whether a narrated event was experienced firsthand or by hearsay or inference. Some languages, such as Turkish, mark evidentiality in the grammar; some other languages, such as English, encode it optionally. This study investigated whether the late bilingual speakers noticed this typological difference and whether learning Turkish or English late would influence their translation accuracy. Due to the interaction of multiple factors in bilingual processing, a multifactor model was needed to provide a theoretical backdrop for the study, and for that purpose, the model CASP (Complex Adaptive System Principles) for Bilingualism (Filipović & Hawkins, 2019) was employed. According to the model, bilingual speakers tend to 1) minimize the learning effort of all languages, 2) minimize processing effort, 3) maximize expressive power, 4) maximize efficiency in communication, and 5) maximize common ground in all languages. A total of 58 Turkish-English bilinguals (19 L2 English and 12 L2 Turkish, and 27 early bilinguals) were asked to translate simple declarative firsthand and non-firsthand sentences in either direction. Additionally, they were asked to make a judgment on how confident they were about whether the event in the statement actually occurred. The results demonstrated that AoA, the source of information and direction of the translation, influenced the speakers’ translation accuracy and perception of evidentials. Firsthand sentences were translated more accurately and judged as occurred more confidently by all bilingual groups. Non-firsthand sentences were translated less accurately than firsthand sentences. This source effect was more apparent when the translation was from Turkish to English. L2 Turkish and early bilinguals tend to skip the non-firsthand source of information when translating from Turkish to English more than L2 English bilinguals. When translating from English to Turkish, although all bilingual groups’ accuracy increased, L2 Turkish bilinguals more accurately translated non-firsthand sentences than L2 English and early bilinguals. Our results demonstrate that the typological differences in the languages of bilingual speakers influence their linguistic behaviors, resulting in differences in translated content and related judgments. The distinction between Turkish and English on how evidentiality is encoded in each language impacts Turkish-English bilinguals’ understanding and level of awareness about the source of information in past event narration. Our findings supported the predictions derived from CASP for Bilingualism model, namely, that in a bilingual model (dual language activation) all bilinguals would maximize common ground in CASP terms (i.e., gravitate towards a shared pattern in both languages) but in different ways due to the nature of the typological contrast. When translations are used in legal settings, translators and law enforcement officers should pay extra attention to the features of the language typology and the nature of bilingualism not to miss important information.

Filipović, L., & Hawkins, J. A. (2019). The Complex Adaptive System Principles model for bilingualism: Language interactions within and across bilingual minds. *International Journal of Bilingualism*, 23(6), 1223-1248.

---

## Revisiting the Concept of “Triggering” of Code-Switching

Sonya Trawick

Independent Scholar, Raleigh, USA

Work in psycholinguistics has explored local triggers for bilingual Code-Switching (CS), “the alternation [...] of stretches of one language with stretches of another” (Poplack, 2017:7) (1). For example, the Triggering Hypothesis (Clyne, 1967) claims that when bilinguals encounter a Trigger Word (TW), a word affiliated with both languages (e.g., cognates), the activation level of both languages are affected, and the level of even the language not being spoken is raised. Following TW, the two languages may become co-activated enough for a switch to that other language to occur. This would explain the switches in (1b) and (1d) occurring after the underlined words (Spanish-English cognates) in (1). Previous studies have claimed cross-linguistic support for this hypothesis (Broersma, 2009; Broersma, Carter, Donnelly, & Konopka, 2020; Broersma & De Bot, 2006; Broersma, Isurin, Bultena, & De Bot, 2009), but leave important gaps in the literature. The present study addresses the gaps and revisits previous conclusions with new findings.

Previous studies lack of work on data from the same community of speech, relevant to questions of CS as community norms dictate when and where it can be expected (e.g., Poplack, 1980). Additionally, they leave ambiguity as to the window of speech within which to look for triggering, describing the inspection of ‘clauses’ with finite verbs (e.g., Broersma et al., 2020), without explanation for clause boundaries or the treatment of other types of spontaneous utterances. The present study uses the New Mexico Spanish-English Bilingual (NMSEB) corpus (Torres Cacoullós & Travis, 2018: Chapters 2 & 3) of spontaneous speech from balanced bilinguals, and uses its prosodic transcription (Du Bois et al., 1993) to identify prosodic sentences as windows of speech for inspection, as seen in (1a-d). A further crucial gap is the ambiguity in how previous work distinguishes TW from the very CS that they supposedly trigger. This project identifies CS as multi-word stretches of unilingual speech, and candidate TW as cognate nouns surrounded by discourse in the same language, and noun Lone Other-Language Items (LOLIs), or single words from one language inserted into the running discourse of another (Poplack, 2017:6).

Comparing the presence vs. absence of TW and CS in sequential prosodic sentences, and running linear and logistic regression models that test for predictors previously claimed to provide evidence of triggering (e.g., interlocutor production of the previous window of speech, number of cognates in a window of speech, a speaker’s overall production of cognates), the present study does not reveal findings similar to those previously reported (Fig1). Instead, evidence from looking at the co-occurrence of CSes with other CSes, rather than with TW, appears more illuminating (Fig2). The NMSEB corpus represents a community of balanced bilinguals who CS often and for whom CS is not a salient behavior. A high resting level of co-activation, the prerequisite for habitual CS, may better explain the co-occurrence of CS and other items of bilingual status, ultimately supporting a case for correlation, rather than causal relationship proposed by the Triggering Hypothesis, as a more suited explanation for such communities.

- Broersma, M. (2009). Triggered codeswitching between cognate languages. *Bilingualism: Language and Cognition*, 12(4), 447-462. <https://10.1017/S1366728909990204>
- Broersma, M., Carter, D., Donnelly, K., & Konopka, A. (2020). Triggered codeswitching: Lexical processing and conversational dynamics. *Bilingualism: Language and Cognition*, 23(2), 295-308.
- Broersma, M., & De Bot, K. (2006). Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition*, 9(1), 1-13.
- Broersma, M., Isurin, M., Bultena, S., & de Bot, K. (2009). Triggered code switching: Evidence from Dutch-English and Russian-English bilinguals. In L. Isurin, D. Winford & K. de Bot (Eds.), *Crossdisciplinary approaches to code switching* (85-128). John Benjamins Publishers.
- Clyne, M. (1967). *Transference and triggering: Observations on the language assimilation of postwar German-speaking migrants in Australia*. Martinus Nijhoff.
- Clyne, M. (2003). Dynamics of transversion. In M. Clyne (Ed.), *Dynamics of language contact* (pp. 254-277). Cambridge University Press.
- Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription. In J. Edwards, & M. Lampert (Eds.), *Talking Data: Transcription and Coding in Discourse* (pp. 45-98). Lawrence Erlbaum Associates.
- Poplack, S. (1980). Sometimes i’ll start a sentence in spanish y termino en español: Toward a typology of code-switching. *Linguistics*, 18(7/8), 581-618.
- Poplack, S. (2017). *Borrowing: Loanwords in the speech community and in the grammar*. Oxford University Press.
- Torres Cacoullós, R., & Travis, C. E. (2018). *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge University Press.

---

## Can the variable use of English and Cantonese in a digital medium be predicted? Analyzing WhatsApp code-switching in Hong Kong computationally

Yuen Man Tsang, Wilkinson Daniel Wong Gonzales  
The Chinese University of Hong Kong, Hong Kong, Hong Kong

AI prediction tools like ChatGPT are increasingly being explored for language prediction (MacKenzie 2020; Szmrecsanyi et al 2019). However, existing language models have primarily been trained on dominant languages like English, neglecting lesser-known linguistic practices. These models typically focus solely on linguistic variables, which limits their understanding of unique linguistic phenomena and sociolinguistic nuances, and consequently, their performance.

This paper hopes to address this by developing a supervised predictive model that incorporates both social and linguistic factors to understand the variable use of languages in underrepresented multilingual practices. The study focuses on Cantonese-English code-switching in Hong Kong, specifically examining how linguistic and social factors influence language choice (English vs. Cantonese) in the context of Computer-Mediated Communication (CMC), particularly WhatsApp instant messaging. The goal is to identify robust predictors of bilingual variation in a digital setting. Then, utilizing these results, we evaluate how accurately a model informed by this analysis can forecast bilingual variation in the context of digital code-switching. For our analysis, we adopted a ‘bag-of-words’ approach to investigating language choice (Goldberg 2017:69; Zhang et al. 2010). A Bayesian regression analysis of roughly 329,087 Cantonese and English words from 55,000 utterances from 24 Hong Kong residents (linked to sociolinguistic data acquired from a survey) revealed some noteworthy findings:

1. Part-of-speech and stylistic context strongly condition bilingual variation regardless of the matrix language (probability of direction,  $pd=100\%$ ). Other factors such as sex (median= -0.06,  $pd=98\%$ ), year (median=-0.01,  $pd=92\%$ ), self-reported proficiency in English (median= -0.04,  $pd=91\%$ ), and attitudes towards code-switching as a marker of bilingual identity (median=0.04,  $pd=93\%$ ) have been found to also condition this variation, but to a noticeably lower extent.

2. Multilingual lexicon and resources can have dynamic meaning and functions, depending on their situated use. For example, switches from Cantonese to English signal ‘interpersonality’ whereas the maintenance of English in English-matrix clauses index ‘informationality’. Switches to Cantonese from English allow speakers to convey their intended meaning with greater nuance and emotional depth.

3. Individual factors or constraints have less of an impact than intralinguistic or extralinguistic factors (such as part-of-speech or style) when it comes to Cantonese-English mixing, suggesting a level of uniformity within the speech community.

4. Positive attitudes towards mixing and reported preference for frequent mixing do not necessarily correlate to (or perhaps, lead to) higher rates of mixing. Like the case of neighboring contact languages, language attitudes and verbal preferences do not have to align with actual practice.

The model was able to accurately predict the outcome (i.e., English, Cantonese lexicon) 85% of the time. The correlation between actual responses and predicted responses is significant ( $\rho=0.61$  [CI=0.59, 0.62],  $t=77.356$ ,  $p<0.0001$ ). This study stands out for using less-commonly used methods for analyzed variation in social media. Sociolinguistic variables, such as part-of-speech, style, and sentiment, are computationally inferred from the linguistic context instead of relying solely on participant reports. We hope to provide insights into another dimension of Cantonese-English code-switching and contribute to existing knowledge on the variable use of languages in Hong Kong and East Asia.

Goldberg, Yoav. 2017. *Neural network methods for Natural Language Processing*. (Ed.) Graeme Hirst. California, USA: Morgan and Claypool.

MacKenzie, Laurel. 2020. Comparing Constraints on Contraction Using Bayesian Regression Modeling. *Frontiers in Artificial Intelligence* 3.58. doi:10.3389/frai.2020.00058.

Szmrecsanyi, Benedikt.; Jason Grafmiller.; and Laura Rosseel. 2019. Variation-Based Distance and Similarity Modeling: A Case Study in World Englishes. *Frontiers in Artificial Intelligence* 2.23. doi:10.3389/frai.2019.00023.

Zhang, Yin.; Rong Jin.; and Zhi-Hua Zhou. 2010. Understanding bag-of-words model: A statistical framework. *International Journal of Machine Learning and Cybernetics* 1.43–52.

---

## AM/P~OM/P merger in Hong Kong vs. Toronto Cantonese: An under-documented homeland sound change in a heritage language context

Holman Tse

St. Catherine University, St. Paul, MN, USA

Hong Kong Cantonese has been described as having developed a dissimilatory merger in which /o/ becomes /ɐ/ in pre-labial contexts (/m/ and /p/). The /ɐ/ vs. /o/ distinction (henceforth AM/P vs. OM/P) is illustrated in the following minimal pair (with 35 indicating a mid-rising tone):

[kɛm35], ‘embroidered’, 錦

[kom35], ‘thus’, 噏

Acoustically based studies of this change, however, are lacking presumably because, as Bauer & Benedict (1997: 419–420) claim, the contrast disappeared by around the middle of the 20th century. Can this claim be confirmed through an acoustically based apparent time study? Furthermore, if Cantonese speakers born before the 1950s have settled in other cities such as Toronto, Canada, would heritage speakers there also participate in this change?

The current study specifically addresses the following:

Q1) How are AM/P vs. OM/P produced in terms of F1, F2, and F3?

Q2) How do Pillai Scores (Nycz & Hall-Lew 2015) vary based on Year of Birth?

Q2) Are there differences based on City (Toronto vs. Hong Kong) or on Generational Group (Gen1 vs. Gen2)?

The data comes from sociolinguistic interviews from the Heritage Language Variation and Change (HLVC) Corpus (Nagy 2011). The 38 participants analyzed include eight from Hong Kong and 30 from Toronto (including 15 Gen1 and 15 Gen2). Year of Birth ranged from 1922 to 1998. For each recording, vowel formant measurements were collected using a Praat script. All vowels were normalized and scaled to Hertz values using the Nearey technique (Thomas & Kendall 2007) based on a vowel space of 12 monophthongs. Historic vowel class information was based on Eitel (1877).

Mixed effects modeling of all AM/P (n = 889) and OM/P (n = 816) tokens (with Speaker and Word as random effects; Vowel Category as a fixed effect; and midpoint F1, F2, or F3 as the dependent variable) show that OM/P is significantly raised (lower F1,  $p < 0.001$ ), significantly retracted (lower F2,  $p < 0.001$ ), and significantly more rounded (lower F3,  $p < 0.01$ ) than AM/P. Pillai Scores ranged from 0.04 to 0.86 (with 0 indicating most merged and 1 the least merged). While a Pearson correlation test shows lower Pillai Scores with more recent Years of Birth ( $r(36) = -0.361$ ,  $p < 0.05$ ), the change appears to be ongoing, contrary to Bauer & Benedict’s (1997) claim. In fact, the highest Pillai Score comes from a Toronto speaker born in 1969. Nevertheless, neither City nor Generational Group, came out as significant in any model.

To conclude, through acoustic analysis of Homeland speakers, this study shows that a change that was previously described as complete is, in fact, still ongoing. This study, thus, contributes to the description of Hong Kong Cantonese vowel variation, which has been under-researched compared to consonantal and tonal variables (cf. Matthews & Yip 2011: 73–76; Fung & Lee 2019). This study also contributes to a growing number of studies addressing the extent to which heritage speakers participate in Homeland sound change (ex: Thepboriruk 2015; Kang & Nagy 2016; Cheng 2019).

Bauer, Robert S. & Paul K. Benedict. 1997. *Modern Cantonese Phonology* (Trends in Linguistics 102). Berlin: Mouton de Gruyter.

Cheng, Andrew. 2019. VOT merger and f0 contrast in Heritage Korean in California. *University of Pennsylvania Working Papers in Linguistics* 25(1). <https://repository.upenn.edu/pwpl/vol25/iss1/9>.

Eitel, Ernest John. 1877. *A Chinese dictionary in the Cantonese dialect*. Trübner and Company.

Fung, Roxana SY & Chris KC Lee. 2019. Cantonese vowel merger-in-progress. In *Proceedings of the 19th International Congress of Phonetic Sciences, Australia 2019*, 3225–3229.

Kang, Yoonjung & Naomi Nagy. 2016. VOT merger in Heritage Korean in Toronto. *Language Variation and Change*. Cambridge University Press 28(2). 249–272. <https://doi.org/10.1017/S095439451600003X>.

Matthews, Stephen & Virginia Yip. 2011. *Cantonese: A Comprehensive Grammar*. 2nd edn. Routledge.

Nagy, Naomi. 2011. A Multilingual Corpus to Explore Variation in Language Contact Situations. *Rassegna Italiana di Linguistica Applicata* 43(1/2). 65–84.

Nycz, Jennifer & Lauren Hall-Lew. 2015. Best practices in measuring vowel merger. *Proceedings of Meetings on Acoustics* 20(1). 060008. <https://doi.org/10.1121/1.4894063>.

Thepboriruk, Kanjana. 2015. *Thai in Diaspora: Language and Identity in Los Angeles, California*. University of Hawai’i at Manoa Ph.D. Dissertation.

Thomas, Erik & Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite. <http://lingtools.uoregon.edu/norm/norm1.php>.

---

## The decline of rhoticity in Lancashire, Northern England: data from ultrasound and sociolinguistic interviews

Danielle Turton, Robert Lennon  
Lancaster University, United Kingdom

The towns of East Lancashire in Northern England are amongst the few remaining rhotic areas in the country. i.e., speakers retain [RL1] the /r/ in words like *car* and *bird*. This is notable because rhoticity has almost disappeared from Anglo-English entirely. In addition, it is heavily stigmatised, which has led many to predict rapid loss of rhoticity in England (e.g., Trudgill 2000; Britain 2002).

We present new data from ultrasound tongue imaging, taken from the speech of 28 speakers from Blackburn, a traditionally rhotic area of Lancashire. Speakers are divided equally into two age groups (below 35, above 50) and by binary gender. This data is complemented by sociolinguistic interview data with a further 28 speakers. Auditory analysis confirms previous accounts of a decline in apparent time. That said, speakers almost always indicate a difference in minimal pair tests between pairs like *spar* and *spa*, even the youngest generation. An acoustic analysis [RL2] (minimum F3 and F3-F2 difference) indicates that rather than simply a categorical deletion, /r/ is weaker for younger speakers higher F3 acoustically). /r/ is also weaker in women, demonstrating a gradient female-led change in apparent time. Ultrasound data complements this by demonstrating that different vocalic contexts produce different kinds of /r/s for some speakers (bunched vs. retroflex), as do different positional contexts. We argue that some of these realisations may be more susceptible to weakening from both an articulatory and perceptual perspective.

From the perspective of sound change, we consider two paths to the loss of rhoticity found in our dataset. That is, Blackburn speakers show two strategies towards /r/ loss in the sociolinguistic interviews: i) a categorical absence of /r/: these speakers had no trace of an /r/ acoustically; Or ii) a gradient gestural or acoustic loss: a residue of /r/ remained and the acoustic spectrum of pairs such as *spa* and *spar* looked consistently different in the sociolinguistic interviews.

We discuss these effects with reference to sound change in progress, stigmatisation of non-standard variants and whether the process is above or below the level of conscious awareness, arguing that the additional articulatory evidence affords us a window into hitherto unexplained sound change.

- Barras, William Simon. 2011. Sociophonology of rhoticity and r-sandhi in East Lancashire English. Doctoral Dissertation, The University of Edinburgh.
- Blaxter, Tam, Kate Beeching, Richard Coates, James Murphy, and Emily Robinson. 2019. Each person does it their way: Rhoticity variation and the community grammar. *Language Variation and Change* 31:91–117.
- Britain, David J. 2002. Phoenix from the ashes?: The death, contact and birth of dialects in England. *Essex Research Reports in Linguistics* 41:42–73.
- Labov, William. 2001. *Principles of linguistic change: social factors*. Oxford: Blackwell.
- Leemann, Adrian, Marie-José Kolly, and David Britain. 2018. The English Dialects app: The creation of a crowdsourced dialect corpus. *Ampersand* 5:1–17.
- Trudgill, Peter. 2000. *Sociolinguistics: An introduction to language and society*. Penguin UK.

---

## Stability in the face of contact: Evidence from Toronto Heritage Tagalog /u/

Pocholo Umbal  
University of Toronto, Toronto, Canada

Heritage languages provide an additional testing ground for investigating the complex interplay of internal and external factors on variation and change in language contact situations. This study presents a comparative variationist analysis of Tagalog /u/, comparing the F2 values among two generations of heritage speakers in Toronto (where Tagalog is the sixth largest heritage language), and two groups of age-matched homeland speakers in Manila, the hometown of the Toronto speakers' families.

Existing studies of Tagalog vowels suggest that the F2 of /u/ is not undergoing change (Gonzalez 1970; Delos Reyes et al. 2009; Intlekofer 2012). By contrast, the F2 of /uw/ in Canadian English (the majority language of the heritage speakers) is increasing in apparent time, indicating presence of on-going /uw/-fronting (Hoffman 2016). Previous work on heritage speakers in other languages have shown that variability in /u/ may be driven by dominant-language transfer (Tse 2022), language-internal factors (Cheng 2021), or identity marking (Cummings Ruiz 2019). What type of variability do Heritage Tagalog speakers exhibit; and which processes might account for the patterns?

1,617 tokens of /u/ were extracted from a forced-aligned corpus of Tagalog spontaneous speech, comprising 16 heritage speakers in Toronto, stratified by immigrant generation (GEN1 vs. GEN2) and gender (female vs. male); and 12 homeland speakers in Manila, stratified by age (Older vs. Younger) and gender. Tokens were Lobanov-

normalized and then coded for relevant linguistic (i.e., preceding sound class, vowel duration), social (generation, age for homeland speakers, and gender) and contact-based (ethnic orientation scores for identity and language use/preference for heritage speakers) factors.

Analyses using mixed-effects linear regression show no evidence of on-going change in the homeland variety, and likewise, in the heritage variety. Heritage GEN1 females show more fronted [u] compared to GEN1 males, but this is attributed to speech rate (resulting in an overall reduction of the GEN1 females' vowel space area). For all groups (i.e., Homeland, GEN1 and GEN2 heritage speakers), the F2 of [u] is conditioned by the preceding sound: coronals have a favouring effect, resulting in more fronted [u], while noncoronals have a disfavouring effect. Crucially, the magnitude of this effect is smaller than that reported for English /uw/. Finally, more Filipino-oriented heritage speakers approximate the preceding class constraint of homeland speakers more closely than more Canadian-oriented speakers.

These findings are consistent with stable variation and absence of contact-induced transfer from English in both overall fronting and the magnitude of the preceding class effect (cf. Flemming 2011). Variable patterns in the heritage variety reflect inherent variation, conditioned by coarticulation, and the universal effects of speech rate. Further, heritage speakers may be using linguistic constraints (as opposed to mean values; cf. Hoffman & Walker 2010) as resources for ethnic identity construction. Taken together, this study demonstrates that phonetic convergence is not a necessary (or sole) outcome in language contact situations (Poplack & Levey 2010) and provides a nuanced account of how language-internal, social, and contact-based factors may conspire to give rise to variation (and change) among heritage speakers.

- Cheng, Andrew. 2021. Maintenance of phonetic and phonological distance in the English and Korean back vowel contours of heritage bilinguals. *Journal of Phonetics* 89: 101–109.
- Cummings Ruiz, Laura D. 2019. North midland /u/-fronting and its effect on heritage speakers of Spanish. In *Proceedings of the 19th International Congress of Phonetic Sciences*, ed. Sasha Calhoun, Paola Escudero, Marija Tabain, and Paul Warren, 1099–1103. Canberra: Australian Speech Science and Technology Association.
- Delos Reyes, Jevalene, Paul Julian Santiago, Danielle Tadena, and Louward Allen Zubiri. 2009. Acoustic characteristics of the Filipino vowel space. *Proceedings of the 6th National Natural Language Processing Research Symposium* 7–11.
- Flemming, Edward. 2011. The grammar of coarticulation. In *La coarticulation: Des indices à la représentation*, ed. Mohamed Embarki and Christelle Dodane, 189–211. Paris: L'Harmattan.
- Gonzalez, Andrew. 1970. Acoustic correlates of accent, rhythm, and intonation in Tagalog. *Phonetica* 22: 11–44.
- Hoffman, Michol F. 2016. 'Back to front': The role of ethnicity in back vowel fronting in Toronto English. Paper presented at *New Ways of Analyzing Variation* 45, Simon Fraser University.
- Hoffman, Michol F., and James A. Walker. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22: 37–67.
- Intlekofer, Darlene G. 2012. Tagalog /u/-lowering: An instrumental study of spontaneous speech. *Generals Paper*, CUNY Graduate Center.
- Poplack, Shana and Stephen Levey. 2010. Contact-induced grammatical change: A cautionary tale. In *Language and space: An international handbook of linguistic variation*, ed. Peter Auer and Jürgen Erich Schmidt, 301–419. Berlin: Mouton de Gruyter.
- Tse, Holman. 2022. What can Cantonese heritage speakers tell us about age of acquisition, linguistic dominance, and sociophonetic variation? In *Variation in second and heritage languages: Crosslinguistic perspectives*, ed. Robert Bayley, Dennis R. Preston, and Xiaoshi Li, 97–126. Amsterdam: John Benjamins.

---

## Variation in the acceptability of singular they in Singapore English

Alamelu Venkatachalam, Rebecca Starr  
National University of Singapore, Singapore, Singapore

Despite longstanding prescriptivist critiques, the English third-person pronoun they has been used for centuries to reference singular antecedents (Bodine 1975). While singular they has primarily been used to reference generic and epicene antecedents in the past (e.g., “Every citizen should cast their vote”), recent studies of ‘Inner Circle English’ varieties (Kachru 2006), such as American English, have observed an increase among younger speakers in the acceptability of singular they when referencing specific and definite antecedents, as in, “The teacher has lost their pencil.” (Conrod 2019; Moulton et al. 2020). The rise in acceptability of singular they has been linked in public discourse with the increasing visibility of non-binary identities and broader shifts toward gender-neutral language in the Global North. Scholars have yet to examine how singular they functions in Outer Circle varieties such as Singapore English, which has been proposed to be simultaneously shifting towards internal linguistic norms but also adopting certain innovations due to exposure to non-local varieties (Starr 2019). The present study explores the linguistic, social, and attitudinal factors that influence the acceptability of singular they in Singapore English.

This study draws on data from a survey of 102 speakers of Singapore English in which respondents rated 54 sentence stimuli for grammaticality on a five-point scale. These stimuli varied by pronoun (she, he, they),



definiteness of antecedent (definite versus generic), gender prototypicality of referent (gendered versus non-gendered personal names, e.g., Claire versus Frem), and case (nominative, accusative, genitive). Results were analyzed via linear mixed-effects modelling using the lmerTest package in R.

Significant main effects were identified in the data for gender and religiosity of respondent, with male and religious respondents significantly less accepting of singular they. Consistent with prior findings in Inner Circle varieties, younger respondents (ages 20 to 24) rated singular they as more acceptable than older speakers (ages 25 to 65). Age interacted significantly with both gender and religiosity; older participants exhibited greater differences between male versus female and religious versus non-religious respondents. While younger respondents showed greater acceptance than older respondents for definite instances of they, this effect was not significant in the model. Respondents' levels of linguistic prescriptivism were also not found to be significant predictors of singular they rating.

These findings indicate that singular they is experiencing a rise in acceptability in Singapore English similar to its trajectory in Inner Circle varieties. The data also suggest that ideological stances associated with particular social groups, rather than prescriptivist attitudes, constrain the acceptance of singular they in Singapore. The uptake of singular they in this variety underscores that the influence of external norms on Outer Circle Englishes should not be uniformly characterized as the adoption of prestigious features by speakers who wish to speak a more 'standard' variety.

- Bodine, Ann. 1975. Androcentrism in prescriptive grammar: Singular 'they', sex-indefinite 'he', and 'he or she'. *Language in Society* 4(2). 129-146. doi:10.1017/S0047404500004607.
- Conrod, Kirby. 2019. Pronouns raising and emerging. Seattle, WA: University of Washington dissertation.
- Kachru, Braj B. 2006. The English language in the Outer Circle. In Kingsley Bolton & Braj B. Kachru (eds.), *World Englishes*, 241-255. New York: Routledge.
- Moulton, Keir, Chung-hye Han, Trevor Block, Holly Gendron & Sander Nederveen. 2020. Singular they in context. *Glossa: a journal of general linguistics* 5(1):122 1–13. <https://doi.org/10.5334/gjgl.1012>.
- Starr, Rebecca L. 2019. Cross-dialectal awareness and use of the BATH-TRAP distinction in Singapore: Investigating the effects of overseas travel and media consumption. *Journal of English Linguistics* 47(1). 55-88. doi: 10.1177/0075424218819740.

---

## Negative marking as a sociolinguistic variable among Palestinian speakers in Beirut

Nina VAN KAMPEN  
INALCO, Paris, France

This study examines variation in the realization of verbal negation in Arabic spoken by Palestinians in Beirut, presenting a case of dialect contact resulting from the creation of the state of Israel in 1948 and the subsequent influx of Palestinian refugees to Lebanon.

Palestinian Arabic is reported to have three possible realizations of verbal negation: the preverbal, bipartite and postverbal construction (Bergsträsser 1915, Palva 1984, Lucas 2010). As for Beirut Arabic, it is reported to know the preverbal construction only (Naim 2011).

The present study will examine the distribution of preverbal, bipartite and postverbal negative constructions, looking at both internal (linguistic) and external (social and spatial) factors. The analysis is based on data collected by the researcher between 2018 and 2022. The informants are Palestinians of different ages, genders and social and geographical backgrounds living in Beirut, both in and outside of the refugee camps.

The results show that the bipartite and postverbal constructions are used rather marginally, while the preverbal form - coinciding with the construction used in Beirut Arabic - is prevalent. Looking at the bipartite and postverbal constructions, it is observed that these forms marked as Palestinian Arabic are mostly used by the older generations, but that they are also sometimes used by younger male speakers living in the refugee camps. This suggests that processes of convergence to and divergence from Beirut Arabic happen simultaneously. While, as a result of long term dialect contact between Beirut Arabic and Palestinian Arabic, broader processes of levelling seem to be taking place (Trudgill 1986) leading to a loss of certain Palestinian Arabic features, it can be argued that these same Palestinian Arabic features are used by certain speakers as a socially meaningful resource indexing Palestinianness.

- Bergsträsser, G. 1915. *Sprachatlas von Syrien und Palästina*, Zeitschrift des Deutschen Palästina-Vereins 38: 169-222.
- Lucas, C. 2010. "Negative -š in Palestinian (and Cairene) Arabic: present and possible past." *Brill's Annual of Afro-Asiatic Languages and Linguistics* 2: 165–201.
- Naim, S. 2011. "Beirut Arabic", In *Encyclopedia of Arabic Language and Linguistics*, edited by Kees Versteegh, Vol. 1, 274-86. Amsterdam: Brill.
- Palva, H. 1984. "A General Classification for the Arabic Dialects Spoken in Palestine and Transjordan", In *Studia Orientalia Electronica* 55, p. 357-376.
- Trudgill, P. 1986. *Dialects in contact*. Oxford etc: Blackwell.

---

# Quantifying grammaticalization via translation: The position of Lo Cunto de li Cunti's Neapolitan on the Romance grammaticalization cline

Riley VanMeter

The University of California, Berkeley, Berkeley, USA

There is a developing practice of positioning features of languages or even entire languages/dialects with respect to one another on a cline of grammaticalization. For example, Poplack et al. (2018) finds that the indicative/subjunctive mood alternation in Spanish is less grammaticalized than that of Italian or French. Similarly, Miola (2017) finds that Piedmontese varieties align with, or slightly surpass, French on the Romance grammaticalization continuum (RGC). However, these previous studies all diverge in their methodologies with Poplack et al. (2018) using frequency-based measures and Miola (2017) using both paradigm and frequency. Semantic-based analyses have also been explored (Burgo, 2010). The present study addresses methodological approaches within applied grammaticalization theory with the goal of bridging the gap between translation studies and contrastive/comparative linguistics to advocate for their joint applications in cross-linguistic variationist research.

The combination of the two subfields has been explored by Teich (1999), Ramón García (2021), and Shang (2022). However, almost all previous studies have focused on the implications that corpus-based contrastive linguistics can have in the field of translation studies, i.e., a corpus-based contrastive approach that is applied to translation studies. The present study does exactly the opposite, exploring the repercussions and applications that translations can have within quantitative language change analyses.

In particular, the present research outlines the process wherein the parameters of paradigmaticity and paradigmatic variability (and therefore obligatorification) (Lehmann, 2015) can be used to identify differing levels of grammaticalization through the analysis of parallel translation(s). This is done by identifying parallel semantic expressions and morphosyntactic structures and quantifying both frequency and size of the paradigm between languages. Use of translation bypasses methodological concerns presented by Szmrecsanyi (2016), including potential 'environmental' content changes. In an application of the methodology, the understudied Neapolitan language, specifically that of Giambattista Basile's 17th century *Lo Cunto de li Cunti*, is positioned along the RGC (Lamiroy & De Mulder, 2012), with reference to its own late 20th century Italian translation (Basile, 2007/1986). Contemporary spoken Neapolitan, neo-standard Italian (Zholudeva, 2020), and their mother language Latin are also positioned along the cline, contextualizing 17th century Neapolitan within the greater continuum. Around 500 pages of text are used to analyze the morphosyntactic features of mood alternation, perfect auxiliaries, and possessive adjective position (Fig. 1). Comparative positioning of modern Neapolitan and neo-standard Italian relied on prior descriptive analyses (cf. Berruto, 2017; Ledgeway, 2009). It is ultimately found that contemporary spoken Neapolitan is overall the most grammaticalized, but the literary Neapolitan of the source text is still more grammaticalized than spoken neo-standard Italian and much more so than the literary Italian of the text's translation (Fig. 2).

This methodology's implications in synchronic variationist studies challenge historically diachronic views of grammaticalization. Within the Neapolitan context, this study sheds light on sociolinguistic mechanisms and motivations behind the unidirectional grammaticalization process, specifically in differing language standardization practices, with more intervention seemingly slowing grammaticalization. This study also introduces future questions to consider such as how the translator's idiolect, language history, and bilingualism influence grammaticalization.

Basile, G. (2007). *Lo cunto de li cunti* (M. Rak, Ed.; 9. edition). Garzanti. (Original work published 1986)

Berruto, G. (2017). What is changing in Italian today? Phenomena of restandardization in syntax and morphology: an overview.

In M. Cerruti, C. Crocco, & S. Marzo (Eds.), *Towards a New Standard* (pp. 31–60). De Gruyter.

<https://doi.org/10.1515/9781614518839-002>

Burgo, C. (2010). A Case of Grammaticalization in the Use of the Perfect for the Preterite in Bilbao Spanish. *Studies in Hispanic and Lusophone Linguistics*, 3(2), 301–328. <https://doi.org/10.1515/shll-2010-1077>

Lamiroy, B., & De Mulder, W. (2012). Degrees of grammaticalization across languages. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Grammaticalization* (1st ed., pp. 302–317). Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780199586783.013.0024>

Ledgeway, A. N. (2009). *Grammatica diacronica del napoletano*. Walter de Gruyter – Max Niemeyer Verlag.

<https://doi.org/10.1515/9783484971288>

Lehmann, C. (2015). *Thoughts on grammaticalization* (3rd edition). Language Science Press.

Miola, E. (2017). The position of Piedmontese on the Romance grammaticalization cline. *Folia Linguistica*, 51(1), 133–167.

<https://doi.org/10.1515/flin-2017-0004>

Poplack, S., Cacoullos, R. T., Dion, N., Berlinck, R. de A., Digesto, S., Lacasse, D., & Steuck, J. (2018). Variation and grammaticalization in Romance: A cross-linguistic study of the subjunctive. In W. Ayres-Bennett & J. Carruthers (Eds.), *Manual of Romance Sociolinguistics* (pp. 217–252). De Gruyter. <https://doi.org/10.1515/9783110365955-009>

- Ramón García, N. (2021). Contrastive Linguistics and Translation Studies Interconnected: The Corpus-based Approach. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 1, 393–406. <https://doi.org/10.52034/lanstts.v1i.27>
- Shang, X. (2022). When Contrastive Analysis meets Translation Studies: A historical perspective. *Target. International Journal of Translation Studies*, 186–214. <https://doi.org/10.1075/target.20027.sha>
- Szmrecsanyi, B. (2016). About text frequencies in historical linguistics: Disentangling environmental and grammatical change. *Corpus Linguistics and Linguistic Theory*, 12(1), 153–171. <https://doi.org/10.1515/cllt-2015-0068>
- Teich, E. (2003). Cross-Linguistic Variation in System and Text: A Methodology for the Investigation of Translations and Comparable Texts. *DE GRUYTER*. <https://doi.org/10.1515/9783110896541>
- Zholudeva, L. I. (2020). The Concept of Italiano Neostandard. *Bulletin of Kemerovo State University*, 22(2), 499–506. <https://doi.org/10.21603/2078-8975-2020-22-2-499-506>

---

## Exploring the interplay of literacy and socioeconomic status in language processing: A field-based study

Jessica Vélez-Avilés<sup>1</sup>, Estilita M. Cassiani-Obeso<sup>2</sup>, Paola (Giuli) Dussias<sup>1</sup>

<sup>1</sup>Penn State University, USA. <sup>2</sup>Northwestern University, USA

Sociolinguistic factors such as socioeconomic status (SES) and literacy level can influence some of the cognitive processes studied in psycholinguistics, including language use and processing. Following Labov (1972, 2012) on the primacy of the speech community, this study merges lab and field research. Previous studies suggest that processing advantages are modulated by markers of language skills such as literacy (Favier et al., 2021; Kukona et al., 2016; Mishra et al., 2012) and that high-literacy individuals anticipate target information sooner relative to lower-literacy individuals. However, past research has not consistently controlled for SES, an important confounding variable related to an array of outcomes across the lifespan, including lower educational achievement and poor language development (Fernald et al., 2013; Hackman et al., 2010; Morgan, 2009). Thus, the advantage reported for higher-literacy individuals may be due to SES-related disparities and not to literacy skill per se. Here we explored the role of literacy on language processing during spoken language comprehension with Spanish-speaking adults with varying levels of literacy who belong to the same SES, capitalizing on robust findings that comprehenders use grammatical gender information encoded in pronominal modifiers to facilitate the processing of upcoming nouns (Lew-Williams & Fernald, 2007).

Participants were Spanish speakers from San Basilio de Palenque, a small Afro-Colombian village that offers a unique opportunity to keep SES constant while varying participants' literacy level. The village faces many economic difficulties that impede social mobility and sources of employment are limited, especially those that could provide better living conditions. As a result, Palenqueros are from low SES and live in what would be considered substandard living conditions (e.g., intermittent access to electricity, no running water).

In a visual world eyetracking study, high (n=21; female=9, male=12) and low (n=20; female=13, male=7) literacy speakers heard instructions in Spanish (*Encuentra laFEM/elMASC...* 'Find the') that named one of two objects displayed on a computer screen. Target objects in the spoken instructions were preceded by a feminine or a masculine article that agreed in gender with the two objects in the visual scene (same-gender trials; e.g., *pelotaFEM* 'ball' displayed alongside *galletaFEM* 'cookie') or with only one of the two objects (different-gender trials; e.g., *pelotaFEM* 'ball' displayed with *zapatoMASC* 'shoe'). The data were analyzed by comparing the proportion of fixations to target objects in each condition.

The results (Figure 1) show that high- and low-literacy speakers looked sooner at the target object on different-gender trials than on same-gender trials, replicating results from previous studies with Spanish speakers elsewhere. Crucially, both literacy groups looked at the upcoming target object before target word onset, suggesting that higher- but also lower-literacy individuals were able to anticipate this information. Our findings are discussed in light of studies reporting differences between high- and low-literacy speakers, highlighting the importance of extending findings to lesser-studied and diverse communities to assess the validity of theories of language comprehension, which are largely based on findings from a small sample of the world's languages and on college-aged and educated participants (Arnett, 2008; Henrich et al., 2010). (499)

- Arnett, J. J. (2008). The neglected 95%. Why American psychology needs to become less American. *American Psychologist*, 63(7), 601-614. <https://doi.org/10.1037/0003-066x.63.7.602>
- Favier, S., Meyer, A. S., & Huettig, F. (2021). Literacy can enhance syntactic prediction in spoken language processing. *Journal of Experimental Psychology: General*, 150(10), 2167–2174. <https://doi.apa.org/doi/10.1037/xge0001042>
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16(2), 234-248. <https://dx.doi.org/10.1111%2Fdesc.12019>
- Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, 11(9). <https://doi.org/10.1038/nrn2897>

- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61-135. <https://doi.org/10.1017/S0140525X0999152X>
- Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., Pugh, K. R., Shankweiler, D. P., & Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica (Amst.)*, 171, 72-84. <https://doi.org/10.1016/j.actpsy.2016.09.009>
- Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.
- Labov, W. (2012). What is to be learned. The community as the focus of social cognition. *Review of Cognitive Linguistics*, 10(2), 265-293.
- Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18(3), 193-198. <https://doi.org/10.1111/j.1467-9280.2007.01871.x>
- Mishra, R. K., Singh, N., Pandey, A., & Huetting, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability – Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5(1), 1-10. <https://doi.org/10.16910/jemr.5.1.3>
- Morgan, P. (2009). Risk factors for learning-related behavior problems at 24 months of age: Population based estimates. *Journal of Abnormal Child Psychology*, 37(7), 401-413. <https://dx.doi.org/10.1007/s10802-008-9279-8>

---

## Region > Age > Gender: On the role of geographical and social factors on (morpho-)syntactic variation in the German dialects of Austria

Philip C. Vergeiner, Lars Bülow  
University of Munich, Munich, Germany

Traditionally, variationist sociolinguistics has put less emphasis on the investigation of (morpho-)syntactic variation compared to variation on other linguistic levels (e.g., phonetics and phonology). This can be attributed to methodological considerations but also to the widespread assumption that morphosyntactic variables – or, in general, features at a deeper level of linguistic structure – exhibit less social stratification (e.g., Labov, 2001: 28–29; Hinskens, 1998: 160). (Morpho-)Syntactic features are often regarded as abstract and less salient, limiting their potential for social indexicalization (Cornips, 2022: 167). Furthermore, the relative infrequency of (morpho-)syntactic features may hinder their association with social groups or situations (Eckert, 2018: 190). However, further research is needed to investigate these claims, especially for languages other than English, such as German.

To address this desideratum, the present paper investigates the social variation of (morpho-)syntactic features in the Bavarian and Alemannic dialects of Austria, with a special focus on the influence of age and gender. In addition, the study aims to compare the significance of social factors with geographical factors which have been the primary focus of previous research on Austrian dialects. The paper addresses two research questions: (RQ1) What is the overall effect of the social variables age and gender on (morpho-)syntactic variation in the Bavarian and Alemannic dialects of Austria, and how strong is their effect compared to geographical factors? (RQ2) Are there differences in the social structuring of individual variables, and if so, what explanations can account for these differences – is there, for instance, a correlation with the salience or frequency of a variable?

To answer these questions, dialect data (21,272 tokens) from 163 dialect speakers from 40 locations across Austria are analyzed. Using a multivariate approach, the paper examines 18 (morpho-)syntactic variables which can be assigned to different areas of (morpho-)syntax (see table 1). The statistical analyses are based on a semi-parametric (resampling-based) MANOVA design (Friedrich et al. 2019).

Regarding RQ1, our results indicate that on a global level the factor age has a significant effect while the effect for gender is not significant. In comparison to both social variables, however, regional variation is found to be much more important. Regarding RQ2, our study shows that the variables differ in how far social factors – in particular age – are relevant. While almost all (morpho-)syntactic features display significant regional differences, there are significant age-related differences for eight variables (see figure 1) and gender-related differences only for one variable (see figure 2). Follow-up tests reveal that in particular variables at the intersection of morphology and syntax are more prone to variation than purely syntactic variables. Conversely, salience or frequency of usage do not appear to explain the social variability of the individual features.

- Cornips, L. (2022). The predictability of social stratification of syntactic variants. In T. K. Christensen & T. J. Jensen (Eds.), *Explanations in sociosyntactic variation* (pp. 144–170). Cambridge: Cambridge University Press.
- Eckert, P. (2018). *Meaning and linguistic variation. The third wave in sociolinguistics*. Cambridge: Cambridge University Press.
- Friedrich, S., Konietschke, F., & Pauly, P. (2019). Resampling-based analysis of multivariate data and repeated measures designs with the R package MANOVA.RM. *The R Journal* 11(2), 380–400.
- Hinskens, F. (1998). Variation studies in dialectology and three types of sound change. *Sociolinguistica*, 12, 155–193.
- Labov, W. (2001). *Principles of linguistic change, vol. 2: Social factors*. Philadelphia: John Benjamins.

---

## Making a Splash :The Saliency of Liquid Variation among Boston Spanish speakers

Lee-Ann Marie Vidal Covas  
Boston University, Boston, USA

In sociolinguistics, studying linguistic variables has been crucial for understanding language variation. While previous research focused on individual variables, recent advancements highlight the importance of exploring the relationships between multiple variables within a speech community [4, 3, 2]. This paper presents preliminary findings from a study on liquid variation in Spanish dialectal variation. It examines its relationship with five other variables: Filled Pauses, Subject Personal Pronouns, SPP placement, general subject placement, and coda /s/.

The study investigates the covariation and coherence of these variables and considers the influence of saliency on their usage. Building upon previous research, the study emphasizes the inclusion of multiple salient variables in sociolinguistic studies. [1] & [2] found covariation among low-saliency features while the salient feature, coda /s/, remained independent. This study expands the investigation to include liquid variation.

Liquid variation and coda /s/ are Labovian stereotypes, representing highly socially aware features with systematic variation, distinct variants, phonetic differences, and a role in maintaining phonological contrast. Liquid variation, including laterals and rhotics, is the primary focus, exploring its distribution and intra-linguistic factors influencing its usage. The study also aims to determine whether there is covariation between liquid variation and the other variables.

This study examines whether non-salient and salient features exhibit different patterns. The prediction is that low saliency features show contact-induced structural convergence, while salient features are treated differently by speakers, with their indexical potential amplified by saliency.

My objectives include examining distribution patterns and intra-linguistic factors of liquid variation among Spanish speakers in the Greater Boston Area, as well as investigating its potential covariation with the other five variables. This study contributes to understanding variable linguistic coherence within Spanish dialectal variation, particularly in the context of dialectal and linguistic contact.

For the liquid variation, the preliminary results in Figure 1 demonstrate a consistent tendency for dialectal persistence, where child arrivals (Donaldo & Priscila) closely mirror the linguistic characteristics of the adult arrivals (David & Pascal). This is consistent with [2]'s findings for coda /s/ in that the structural pressure of convergence with English, where systematic liquid neutralization is not a generalized feature of the phonology, doesn't hold for salient features. In other words, these salient features demonstrate a tendency towards persistence rather than convergence in the context of dialectal and linguistic contact driven by their saliency. These results emphasize the importance of considering saliency and covariation in sociolinguistic research to advance our understanding of variable linguistic coherence in dialectal and linguistic contact contexts.

- [1] D. Erker. "Contact, Co-Variation, and Sociolinguistic Saliency: What Mister Rogers Knows about Language Change". In: University of Pennsylvania Working Papers in Linguistics (Selected Papers from NWAV45) 23.2 (2017).
- [2] D. Erker. "How social saliency can illuminate the outcomes of linguistic contact". In: *The Coherence of Linguistic Communities*. Ed. by K. V. Beaman and G. R. Guy. 1st. New York: Routledge, 2022. Chap. 8.
- [3] D. Erker and R. Otheguy. "Contact and coherence: Dialectal leveling and structural convergence in NYC Spanish". English. In: *Lingua* 172-173 (2016).
- [4] G. R. Guy and F. Hinskens. "Linguistic coherence: Systems, repertoires and speech communities". In: *Lingua* 172-173 (2016).

---

## Quantitative perspectives on regional variation from the map-based accent-recognition task: the cases of Dutch and Frisian

Cesko Voeten<sup>1,2,3,4</sup>, Anne-France Pinget<sup>5,4</sup>

<sup>1</sup>University of Pennsylvania, Department of Linguistics, Philadelphia, PA, USA. <sup>2</sup>University of Pennsylvania, Department of Biology, Philadelphia, PA, USA. <sup>3</sup>University of Amsterdam, Amsterdam, Netherlands. <sup>4</sup>Fryske Akademy, Leeuwarden, Netherlands. <sup>5</sup>Utrecht University, Utrecht, Netherlands

We present a perceptual-dialectological study of regional variation in Dutch and Frisian. Contrasting with the usually production-based, objective data used in traditional dialectology (e.g. Levenshtein distances; [1]), perceptual dialectology ([2]) is a complementary approach that elicits perception-based, subjective associations between language and geography. We demonstrate that a perceptual-dialectological approach can provide a rich quantitative perspective on regional variation, with relatively little effort. This makes it a suitable option for under-resourced languages, such as Frisian.

We applied a map-based accent-recognition task ([3]) to Netherlandic Dutch and Frisian. Whereas Dutch is well-studied, Frisian is an understudied minority language spoken in the province Friesland in the Netherlands. There is a

moderate body of work on Frisian dialectology (see [4]), but no known insight about how variation in Frisian is perceived and used by language users. Including Dutch and Frisian in a single piece of work allowed us to assess the validity of the task by examining whether its results in Dutch match established dialectological results. The Frisian version of the task was deployed after this confirmation of the task's validity was successfully obtained.

Both versions of the task were constructed by taking forty 20-second fragments of regionally-accented spontaneous or interview speech from a geographically representative selection of twenty cities and towns (in the Frisian experiment: within Friesland; in the Dutch experiment: from all other provinces). These were used to create two online accent-recognition experiments, where participants listened to the provided fragment and indicated in a Google-Maps interface where they believed the speaker to be from. We recruited 1,578 (Dutch) and 1,848 (Frisian) participants, who all performed ten trials of their respective experiment. In both experiments, listeners' own regional origins (reported in a pre-experiment questionnaire) were highly varied, making the results geographically representative.

We present three quantitative perspectives on the two datasets. Using multinomial logistic regression, we show how listeners associated the input fragments with geographical locations. Using the Dutch data, we subsequently show that distances in these accent placements significantly correlated (Mantel-test  $r = .50$ ,  $p < .001$ ) with Levenshtein distances by [5], which supports the task's (convergent) validity. In addition, we use cluster analyses to show listeners' geographical partitioning of perceptual isoglosses (Figure 1). For Dutch, the cluster analysis shows that listeners largely placed the speech fragments along the (Google-Maps-provided) province boundaries, but with additional dialectologically-motivated internal subdivisions. In Frisian, the clusters reproduce the east-west division between the two biggest dialects (Clay/Wood Frisian), but do not differentiate them from the third dialect (Southwest-Corner Frisian). Instead, they propose one resp. two subdivisions in Wood and Clay Frisian. The Clay split has been theorized ([6]), and the theoretical boundary coincides with the cluster boundary. However, the other subdivisions are novel. We ascribe these divergences to the data sources: traditional Frisian dialectology is based on lexicographic data, while we studied regional accents. We end with prospective implications for Dutch/Frisian dialectology and perceptual dialectology of regional accents.

- [1] Heeringa, W. J. (2004). Measuring dialect pronunciation differences using Levenshtein distance. [Doctoral dissertation, University of Groningen].
- [2] Preston, D. R. (1989). Perceptual dialectology: Nonlinguists' views of areal linguistics. *Foris*.
- [3] Pinget, A.-F. & Voeten, C. C. (in press). Social factors in accent recognition: a large-scale study in perceptual dialectology. *Journal of Linguistic Geography*.
- [4] Landsbergen, F., Tiberius, C., & Dernison, R. (2014). Taalportaal: an Online Grammar of Dutch and Frisian. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), 2206–2210.
- [5] Heeringa, W., & Nerbonne, J. (2006). De analyse van taalvariatie in het Nederlandse dialectgebied: Methoden en resultaten op basis van lexicon en uitspraak. *Nederlandse Taalkunde*, 3, 218–257.
- [6] Duijff, P. (2008). Towards Standard Frisian in the Friesch Woordenboek. In: Mooijaart, M. & van der Wal, M. (Eds.). *Yesterday's Words: Contemporary, Current and Future Lexicography* (pp. 53–66). Cambridge Scholars Publishing.

---

## Macro and meso approaches to queer linguistics: variation in non-binary pronouns in Dutch

Hielke Vriesendorp

Utrecht University, Utrecht, Netherlands. Fryske Akademy, Leeuwarden, Netherlands

How do queer speakers use linguistic variation to position themselves in the social landscape? Queer speakers face stereotypes and harmful generalisations, which make making large decontextualised claims about them potentially discriminatory, as well as inaccurate. Furthermore, developments in intersectional approaches to social identity have shown that static larger identity labels are unwieldy tools for analysis, especially in a queer context. Because of this, most work in queer linguistics (in line with wider contemporary socio-linguistic work) has sought to analyse language use in highly contextualised case studies with explicitly limited generalisability (Jones, 2021). This leaves a gap in knowledge on what linguistic behaviours are common.

The current paper maps the usage of third person singular pronouns to refer to non-binary referents in Dutch, through a survey of largely inclusivity-oriented speakers of Dutch (N=702). Participants were asked to fill out the blanks in a Dutch conversational story about non-binary singer Sam Smith (who uses they/them exclusively in English) and to judge the acceptability of a range of pronominal options. The participants most often combined two main pronominal systems *die/diens* and *hen/hun* to near-match grammatical constraints in Dutch. This suggests those forms are the emerging standard, but that they are still quite grammatically marked, which participants aim to minimise by strategically picking the 'closest' grammatical match.

Participants also exhibited large demographic variation. It was found that some identity-based macro categories were predictive of linguistic behaviour: gender was a significant predictor of variant choice. Non-binary participants were more likely to use *hen/hun* than people who were not non-binary, and women and men both were less likely to

use emerging standard. Cisness, age, heterosexuality, and being a person of colour all made it less likely that a participant used one of the emerging standard forms.

At the same time cultural alignment influenced pronoun choice: position on a scale from conservative to progressive and on a scale from right wing to left wing significantly predicted pronoun usage, as did knowledge about Sam Smith's non-binary gender: progressives, left-wing participants and participants who were aware used the emerging standard forms more.

Finally, the amount of time spent in mostly cisgender heterosexual spaces and the amount of time spent in social spaces with a mix of LGBT+ people of different identities predicted variant choice: more time in these spaces meant lower usage of the emerging standard forms. Orientation towards those two groups may indicate participants are more concerned with mainstream acceptability, leading to the use of fewer marked forms and more alternative strategies.

I argue that the data show that it is possible to conduct queer linguistic research in a way that maps larger, generalisable patterns. By asking questions beyond macro-identity labels, and introducing questions about cultural and political alignment nuance is added to broad statements on how large groups speak compared to one another, allowing for an analysis of social positioning for queer speakers. This combination may allow for generalisation without essentialisation.

Jones, Lucy. 2021. Queer linguistics and identity: The past decade. *Journal of Language and Sexuality* 10(1). 13–24.

---

### Sociolinguistic research projects as brands

Suzanne Wagner, Betsy Sneller, Jack Rechsteiner

Michigan State University, East Lansing, USA

We make the case that branding is critical for the success of many sociolinguistic projects, especially (but not exclusively) those that involve longitudinal relationships between the researcher and the community. Scholars may be resistant to the concept of branding, because their institutions appear to spend money on corporate image rather than on research or teaching support (Mathews 2021; Philpott, Dooley, O'Reilly & Upton 2011). Yet as sociolinguists we are unusually well placed to understand the most critical essence of branding, namely that a brand is an indexical field (Eckert 2008). Every brand, from Coca-Cola to IBM, indexes a constellation of “intangible attributes” (Sammut-Bonici 2015) like fun or reliable, that contribute to brand identity. Since sociolinguists are familiar with how social indices work, we already have the skillset to be highly effective brand creators. As such, we can use limited resources efficiently to ensure that branding helps us to recruit and retain participants and project personnel; to facilitate fundraising; and to engage and serve the public. This paper offers an overview of relevant public sector brand theory (Boenigk & Becker 2016) with practical examples from four recent sociolinguistic projects: *MI Diaries* (Sneller, Wagner & Ye 2022), *Accent Bias in Britain* (Levon, Sharma, Watt & Perry 2020), *Manchester Voices* (Drummond, Dan, Tasker & Ryan 2022) and *Our Dialects* (MacKenzie 2018).

Many sociolinguists already employ branding, even if they do not realize it. For example, fieldworkers are already akin to personal brands, projecting the indexical values of trustworthiness and approachability that are needed to recruit participants and community partners. But personal brands are insufficient for bigger projects and for those involving remote data collection. The university's brand can also be leveraged, but it typically does not index approachability, which is critical for eliciting vernacular language. We explain how sociolinguistic research projects can take a more intentional, customized approach to develop a high level of brand equity (i.e., value based on brand image, brand trust, and brand awareness, Sammut-Bonnicci 2015) that transcends reliance on personal or institutional brands. We describe the utility of a mission, vision, and values exercise (University of Minnesota 2023), and suggest ways to take the next steps beyond it. Throughout, we emphasize ways in which sociolinguistic research project brands can be built with minimal institutional resources while also offering vocational professional development opportunities training to students. In sum, we propose that sociolinguistic projects can benefit from brand management practices to meet their research goals.

Boenigk, Silke & Anniko. Becker. 2016. Toward the importance of nonprofit brand equity. *Nonprofit Management and Leadership* 27(2). 181–198.

Drummond, Rob, Holly Dann, Sarah Tasker & Sadie Durkacz Ryan. 2022. The Manchester Voices Accent Van: taking sociolinguistic data collection on the road. *Linguistics Vanguard* 8(1). 263–277.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics*. 12(4). 453–476.

Levon, Erez, Devyani Sharma, Dominic Watt & Christina Perry. 2020. *Accent Bias in Britain: Attitudes to accents in Britain and implications for fair access*. <https://accentbiasbritain.org/wp-content/uploads/2020/03/Accent-Bias-Britain-Report-2020.pdf>, retrieved April 10, 2023.

MacKenzie, Laurel. 2018. Dialect surveys and student-led data collection. In Christine Mallinson, Becky Childs & Gerard Van Herk (eds.), *Data Collection in Sociolinguistics*, 151–153. 2nd edn. New York, NY: Routledge.

Mathews, José. 2021. Corporate branding in higher education. *IUP Journal of Brand Management*. 18(3). 7–28.

- Philpott, Kevin, Lawrence Dooley, Caroline O'Reilly & Gary Lupton. 2011. The entrepreneurial university: Examining the underlying academic tensions. *Technovation* 31(4). 161–170.
- Sammut-Bonnici, Tanya. 2015. Brand and branding. In Cooper, C. (ed), *Wiley Encyclopedia of Management, Volume 12: Strategic Management*. 3rd edition. 1–3.
- Sneller, Betsy, Suzanne Evans Wagner & Yongqing Ye. 2022. MI Diaries: ethical and practical challenges. *Linguistics Vanguard*. De Gruyter Mouton 8(s3). 307–319.
- University of Minnesota. 2023. The roles of mission, vision, and values. In University of Minnesota (ed.), *Principles of Management*, <https://open.lib.umn.edu/principlesmanagement/chapter/4-3-the-roles-of-mission-vision-and-values/>, retrieved May 30, 2023.

## “In my real voice?”

### The role of elicitation task on the speech of American and British transatlantic migrants

Abby Walker, Andres Lopez, Sara Carter, Casey Branson  
Virginia Tech, Blacksburg, USA

It is well-established that speakers use features differently in wordlist vs. interview speech, with two main findings: speakers typically use more standardized variants in wordlist speech (Labov 2006) and, in cases of change, conservative speakers can use more innovative variants in wordlist speech (Labov et al. 1991). Work in second dialect acquisition (SDA) has also found that the speech production task can be an important factor in the degree to which a speaker uses new dialect (D2) features, but the direction of the effect differs across studies: Johnson & Nycz (2015) find that North American domestic migrants use more D2 features in conversational vs. read speech, while Lin (2023) finds that Chinese domestic migrants use more D2 features in the wordlist. In the current study we compare wordlist and conversational speech of English speaking, international migrants.

Data from this study comes from recordings made of 21 US-born migrants living in the UK (recorded in London), and 19 English migrants residing in the US (recorded in Ohio), as well as controls for each group (30 US non-migrants and 23 English non-migrants). While we expect to have data from all speakers by NAWAV, our current analysis is based on a subset of 60 speakers (15 from each group). These participants completed a naming task at a computer, where they read words aloud that appeared in isolation on the screen, and then they engaged in an interview, which ranged from 10-90 minutes in length.

The wordlist was designed such that speakers produced a number of word-internal intervocalic /t/, which is typically flapped in US but not UK English, and words with (historically) non-prevocalic /r/, which would be realized in most US varieties, but not in most English varieties. In previous analyses of the full dataset, we see evidence of SDA in UK speakers on rhoticity, and in US speakers on intervocalic /t/. Using /t/ duration and F3 Proportion (dividing a speaker's rhotic F3 values by their average non-rhotic F3 value, see Hagiwara 1995), we see the same trends hold in the current subset of the data. However, when we look at these two features in the interview data, there is no evidence of differences in /t/ or /r/ between migrant and non-migrant groups from either American or English speakers.

These results confirm previous work finding that speech style mediates the amount of D2 features we will see in migrant speakers, highlighting both methodological concerns around how we elicit speech, but also theoretical questions about what it means for someone to acquire a second dialect. However, while we find that we observe more D2 features in wordlist speech (see also Lin 2023), Johnson & Nycz (2015) found more D2 features in conversational speech. Assuming that wordlist speech does represent some sort of “phonic intention” (Labov 2006: 152), these cross-study differences highlight how context- and speaker-specific these intentions can be, reflecting not only standard language ideology and dialectal-allegiance, but also what the speaker understands the function of a wordlist task to be (cf. Gafter 2016).

- Gafter, Roey J. (2016). What's a stigmatized variant doing in the word list? Authenticity in reading styles and the Hebrew pharyngeals. *Journal of Sociolinguistics* 20: 31–58.
- Johnson, Daniel E. & Jennifer Nycz (2015). Partial Mergers and Near-Distinctions: Stylistic Layering in Dialect Acquisition. *Penn Working Papers in Linguistics: Selected Papers from New Ways of Analyzing Variation (NWAY)* 43: 21(2).
- Labov, William, Mark Karan & Corey Miller. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3:33–74.
- Labov, William. (2006). *The Social Stratification of English in New York City*. Cambridge: Cambridge University Press.
- Lin, Yuhua (2023). Towards a comprehensive model of style-shifting: Evidence from sibilant variation in Mandarin. *Language & Communication* 89: 23–41.



---

## On the morphophonological productivity of /v/ glide insertion following underlying final /u/ in dative case suffixation of modern Kannada nouns

Grace Walters

Rice University, Houston, USA

This paper explores morphophonological productivity loss in the Dravidian language Kannada, investigating the extent of its /v/ insertion rule's productivity and how sociolinguistic variables affect the probability of its use. This paper contributes to the dearth of scholarship examining Kannada's phonetics and phonology, as well as that of the Dravidian languages at large. As Kannada is a diglossic language long influenced by the Indo-Aryan family, this work broadens understandings of the sociolinguistic area of India, examining how language contact and bilingualism affect change.

Except for some loans, Kannada words end in vowels. Nouns ending in /u/ are placed into two categories: words with underlying final /u/ and words with epenthetic final /u/. There is no evidence that Kannada's epenthetic and underlying /u/s differ articulatorily, acoustically, or perceptually. However, suffixation triggers salient differences in nouns of these two classes; epenthetic /u/-final nouns drop /u/ and take the suffix, while underlying /u/-final nouns insert /v/ before the suffix (see Fig. 1). The classes are also distinguishable by phonotactic cues from root syllable structure and etymology, but whether speakers know this is unclear.

Using a wug test, this paper investigates the productivity of Kannada's /v/ glide insertion rule, and thus, its underlying-epenthetic /u/ distinction. 10 Kannada-English bilinguals translated sentences from English to Kannada that prompted speakers to place target nouns into the dative case via suffixation. The first sentence block included real Kannada target nouns, while the second featured novel target nouns mimicking the phonotactics of underlying and epenthetic /u/- final words. This allows us to explore whether Kannada speakers distinguish between words ending in underlying and enunciative /u/ when inflecting, and if so, whether these categories are lexically specified or based on phonotactic cues. Additionally, this paper probes the potential relationship between the existence of an underlying-epenthetic /u/ distinction and social variables like speaker age, location, formal education in Kannada, and so on.

Results suggest that Kannada's /v/ glide insertion rule in underlying /u/-ending nouns is lexically specified. In real Kannada nouns, speakers consistently differentiated between the /u/s, but did not always insert the /v/ glide, suggesting that /v/ glide insertion and the epenthetic- underlying /u/ distinction are separate phenomena. In nonce Kannada nouns, speakers did not distinguish between the two /u/s, evidencing a lack of awareness of phonotactic cues that separate the two classes. Furthermore, no single morphophonological pattern replaces /v/ glide insertion in nonce nouns; rather, speaker-specific strategies exist – how a speaker treats epenthetic /u/ nouns, they will treat underlying /u/ nouns. Recent immersion in a Kannada- speaking environment is the only social variable shown to affect productivity in real and novel nouns; however, immersion increases likelihood of /v/ glide insertion in both noun classes in a case of overextension.

This work enriches understandings of morphophonological productivity, as it affirms studies establishing productivity as probabilistic rather than categorical (Jun 2021), and matches extant paradigms of productivity loss as a gradient effect often occurring first in colloquial spoken varieties (Jo 2020), with immersion in formal language increasing the likelihood of rule overextension.

Jo, Jinyoung. "Loss of productivity in phonological processes: the case of Korean vowel harmony." M.A. Thesis, University of California at Los Angeles, 2020.

Jun, Jongho. "Morphophonological gradience in Korean n-insertion." *Glossa: a journal of general linguistics* 6.1 (2021).

---

## Preaching to the Choir: What all linguists need to know about defining and operationalizing ethnicity and race in research

Alicia Wassink<sup>1</sup>, Robert Squizzero<sup>1</sup>, Martin Horst<sup>1</sup>, Alex Panicacci<sup>1,2</sup>, Emily Bender<sup>1</sup>

<sup>1</sup>University of Washington, Seattle, USA. <sup>2</sup>Queen Mary University of London, London, United Kingdom

“All linguistic research has the potential to reproduce or challenge racial notions.” - Linguistic Society of America Statement on Race (2019)

Linguistics has been criticized for its undertheorized application of the constructs race and ethnicity. Sociolinguists, sociocultural linguists, linguistic anthropologists, and raciolinguists are in a strong position to influence the field of linguistics regarding how linguists (mis)use these concepts, how we might show better respect for communities, and sustain more ethical research practices. Important critiques and essays include (Spears 2014 and 2020, Charity Hudley 2017, Charity Hudley et al. 2020, García 2020, and Lanehart 2023). In this talk, we define the complex concepts race and ethnicity and explain their relevance for linguistic research. We report the results of a survey of 62 linguistics methods textbooks, which found that scant attention is paid to appropriately collecting and

using race and ethnicity data in training linguistic researchers. (Most relevant advice came from texts in sociolinguistics.) We discuss study design considerations, building on ethnographic and sociolinguistic methodologies, and offer specific recommendations for the solicitation and use of race and ethnicity data with a focus on avoiding social harm, providing appropriate levels of generalization for theory building, and improving descriptive accuracy, especially for undersampled and minoritized groups. For subfields that do not historically attend to variation in their languages of study, we recommend practices geared toward better representing variation in the worlds' languages. We consider issues germane to collecting self-identified ethnicity and race information in a range of study types spanning multiple subfields of linguistics: computational and corpus-based linguistics, formal linguistics, quantitative and experimental linguistics (including sociolinguistics), and qualitative linguistics. We speak to ethical considerations, including the importance of using locally-constructed labels, analyst positionality, and respect for communities. Our goal is to provide researchers with tools to reflect on their own study design, reflect on their own responsibility to participants and communities, and design study prompts that allow more nuanced representation of race or ethnicity.

We are an interdisciplinary team of students and faculty, working in sociolinguistics, applied linguistics, syntax, computational linguistics, areal linguistics, social psychology and sociophonetics.

- Charity Hudley, Anne H. 2017. Language and racialization. *The Oxford Handbook of Language and Society*. Oxford, UK: Oxford Handbooks.
- Charity Hudley, Anne H.; Christine Mallinson; and Mary Bucholtz. 2020. Toward racial justice in linguistics: Interdisciplinary insights into theorizing race in the discipline and diversifying the profession. *Language*. 96(4).200–235.
- García, Justin. D. 2020. Race and Ethnicity, *An Open Invitation to Cultural Anthropology*, 2nd edn ed. by Nina Brown, Thomans McIlwraith, and Laura Tubelle De González. Perspectives: 444–455. American Anthropological Association. Online: <http://perspectives.americananthro.org/>.
- Lanehart, S. 2023. *Language in African American Communities*. (Routledge guides to Linguistics). London: Routledge.
- Linguistic Society of America. 2019. Statement on Race. Online: <https://www.linguisticsociety.org/content/lisa-statement-race>.
- Spears, A. 2020. Racism, Colorism, and Language within their macro contexts. In *The Oxford Handbook of Language and Race* (H. S. Alim, A. Reyes, & P. Kroskrity, Eds.). New York: Oxford.
- Spears, A. 2014. Thinking about Diversity, Languages and Dialects in the U.S. ed. by Marianna diPaolo and Arthur K. Spears. New York: Routledge.

---

## En la zona donde ello[h] [b]i[v]ían: dialectal and language contact outcomes on limited spirantization and labiodentalization of /b/ in the production of US Salvadorans

Deborah Waughtal  
Ohio State University, Columbus, USA

Linguistic research shows that US Salvadorans seem to alter the way they speak in public, in specific ways. Their usage of socially salient or stigmatized features declines—they produce lower rates of coda /s/ weakening, and voseo is often not used in public (Lipski 1988; Aaron and Hernandez 2007; Raymond 2012; Woods & Shin 2016). But, we don't know much about speakers' patterns of non-stigmatized features like /b/-[b], which could fly under the radar for their lack of salience; or even prestigious ones like /b/-[v]. This sociophonetic study explores Spanish /b/—in particular, limited spirantization and labiodentalization—in the production of 17 US Salvadorans who participate in sociolinguistic interviews drawn from the Boston Spanish Corpus. An overarching question in this research is whether a dialect-specific variant undergoes leveling as a result of dialectal contact.

Dialects that exhibit limited spirantization (Campos-Astorkiza 2018), including Central American ones (Carrasco et. al 2012), differ from canonical descriptions of spirantization in Spanish where voiced stops weaken to approximants after all non-continuant sounds. Speakers with limited spirantization typically produce approximants after vowels, and voiced stops in other phonetic contexts, as in *mi vaso*, 'my glass' [mi. 'βa.so], and, *el vaso*, 'the glass' [el. 'ba.so]. The other feature analyzed, labiodental realizations of /b/, as in *el vaso*, 'the glass' [el. 'va.so], is understood as a contact outcome in some US Spanish dialects. Perception research seems to present a [b]-[v] asymmetry in that US-born Spanish speakers perceive [v] as marker of status and confidence in women's speech (Chappell 2018). Production studies suggest that labiodental realizations are constrained by task type, gender, length of time in the US (Stevens 2000), and writing proficiency levels (Trovato 2017).

The current study uses a variationist approach to investigate the distribution of Spanish /b/, a feature that figures less prominently in the literature, in the speech of speakers from El Salvador residing in the American Northeast, in a locale where more people of non-Mexican origin reside than in any other state in the US (Pew Research Center, Hispanic Trends WEB). Acoustic and auditory methods are used to determine which of a set of linguistic and social factors conditions (i) manner and place of articulation variation in /b/ and (ii) the degree of constriction of /b/ realizations, measured along a continuum of constriction degree using consonant-vowel intensity ratios. The factors included are age, gender, level of education, time in the US, urban-rural upbringing, degree of language contact and

bilingualism, phonetic context, word position, stress, and grapheme. Preliminary results suggest that limited spirantization persists; and, in some conversations, both *v* and *b* are labiodentalized, as in *hablo* 'I speak' ['a.vlo] when little contact with English speakers is indicated, suggesting that labiodentals may be an endemic, rather than a contact feature. Additionally, preceding sound, syllable stress, word position, and orthography shape speakers' production to varying degrees. The variants of /b/ under investigation, unequal in terms of social salience, present a new area for research from a language and dialectal contact perspective.

- Aaron, Jessi Elana, and Jose Esteban Hernández. "Quantitative Evidence for Contact-Induced Accommodation: Shifts in /s/ Reduction Patterns in Salvadoran Spanish in Houston." *Spanish in Contact*, John Benjamin, 2007: 327-341.
- Campos-Astorkiza, R. (2018). Consonants. In K. Geeslin (Ed.), *The Cambridge Handbook of Spanish Linguistics* (Cambridge Handbooks in Language and Linguistics, pp. 165-189). Cambridge: Cambridge University Press.
- Carrasco, Patricio, José Ignacio Hualde & Miquel Simonet. 2012. Dialectal differences in Spanish voiced obstruent allophony: Costa Rican Spanish vs Peninsular Spanish. *Phonetica* 69(3).149–179.
- Fernandez, J. (1982). The allophones of /b d g/ in Costa Rican Spanish. *Orbis*, 31. 121 - 146.
- Hernandez, Jose Esteban. (2002). Accommodation in a Dialect Contact Situation. *Revista de Filología y Lingüística*, 28(2), 93-110.
- Lipski, J. (1988). Central American Spanish in the United States: Some remarks on the Salvadoran community. *Aztlán*, 17. 91–123.
- Raymond, C. W. (2012). Generational divisions: Dialect divergence in a Los Angeles-Salvadoran household. *Hispanic Research Journal*, 13(4), 297–316.
- Stevens, John. (2000). On the Labiodental Pronunciation of Spanish /b/ among Teachers of Spanish as a Second Language. *Hispania*. 83. 139.
- Trovato, Adriano. (2017). Chapter 16. Presence of the voiced labiodental fricative segment [v] in Texas Spanish: Selected papers from the 45th Linguistic Symposium on Romance Languages (LSRL), Campinas, Brazil.
- Woods, Michael & Shin, Naomi. (2015). "Fijáte...sabes que le digo yo." Salvadoran voseo and tuteo in Oregon. 10.1075/ihl.10.15woo. In M. Moyna & S. Rivera Mills (Eds.), *Forms of address in the Spanish of the Americas*, John Benjamins Publishing Company.

---

## Why are -ing and -in' processed differently? Exploring a processing asymmetry between canonical and noncanonical sociolinguistic variants

Yosiane White

Utrecht University, Utrecht, Netherlands

Sociolinguistic work on variable ING (thinking~thinkin') uses conversational speech as a window onto the locus of ING variation and the use of the variants -ing and -in' in speech [1,2]. This study uses experimental priming [3,4] to complement this literature by asking how the variants of ING are mentally represented (e.g. are -ing and -in' represented as distinct pieces in the mental lexicon), and how their (non)canonicity impacts spoken word recognition.

261 participants completed two continuous auditory lexical decision experiments. The critical stimuli were disyllabic progressive verbs (frequency controlled). The critical conditions (Table 1) asked whether working and workin' prime thinking and thinkin' by comparing their response times (RTs) to a baseline (e.g. jiggle-thinking). Participants heard 120 critical stimuli using a 400-600ms inter-stimulus interval, and 396 fillers (50% nonwords), and responded whether they heard a real word of spoken American English or a nonword. RTs (measured from sound-file onset to response) were analysed using LMERS. Fixed effects included PrimeCondition, TrialNumber, Prime/TargetFrequency, and PrimeRT. Random intercepts for Participant and Target were included (random slopes did not improve the model fit).

Exp1 shows significant priming for all critical conditions (Table 1 & Figure 1 - left panel). The finding that working primes thinking and workin' primes thinkin' suggests that -ing and -in' are both mentally represented as pieces distinct from the whole words they occur in, since they elicit priming when the prime and target stems are unrelated (i.e. work vs. think). We further detect a processing asymmetry: -ing and -in' primes yield equal facilitation for -ing targets, but same-variant primes facilitate -in' targets better than cross-variant primes (Figure 1). We call this the -in' boost.

A possible explanation for the -in' boost is that the unexpectedness of noncanonical variants causes a processing speedup (the unexpectedness account)[5]. Such a boost should be short-lived[6]. Exp2 therefore uses the same method as Exp1, but introduces an intervening word (e.g. blue) between half the primes and targets, to a) replicate the -in' boost in the no-intervener condition, and b) test whether the asymmetry is attenuated in the 1-intervener condition.

The no-intervener condition replicates the -in' boost (Table 1). Crucially, the 1-intervener condition eliminates the asymmetry and gives equal priming from both prime types to both target types (Figure 1 - right panel). This result supports the idea that shared surface form between sequential -in' primes and targets is violating listeners'

expectations causing an -in' boost (if this were exclusively a shared surface form effect we would expect a parallel ing-ing boost).

However, an alternative and well-supported explanation is that -ing and -in' are mentally represented differently, such that the -in' boost is caused by a shared phonological process that changes an underlying -ing to a noncanonical -in' (the representational account). Evidence in favour of the representational account comes from follow-up studies that manipulate variant proportions and model talker expectations. Upcoming studies manipulate model talker accent and include a variant frequency predictor to further disentangle the unexpectedness account from the representational account.

[1] Tamminga (2016). Persistence in phonological and morphological variation. *Language, Variation & Change*, 28(3).

[2] Houston (1985). *Continuity and change in English morphology: The variable (ING)*. University of Pennsylvania.

[3] Cutler (2012). *Native listening: Language experience and the recognition of spoken words*. MIT Press

[4] Goodwin Davies & Embick (2020). The representation of plural inflectional affixes in English: Evidence from priming in an auditory lexical decision task. *Language, Cognition & Neuroscience*, 35(3).

[5] Bernolet & Hartsuiker (2010). Does verb bias modulate syntactic priming?. *Cognition*, 114(3), 455-461.

[6] Bacovcin, Goodwin Davies, Wilder & Embick (2019) Intervention in two rhyme priming tasks.

---

## Social identity and self-perception in the language classroom: a practical approach

Ioana Wicker  
Graduate Center/CUNY

Recent research demonstrates a powerful connection between positive student outcomes and teaching activities that connect to students' identities, and sense of purpose and belonging. (Krumm, 2010) The process of acquiring a new language has shown to leave an imprint on a person's identity, identity roles and perception of self. (Carter, 2021) This study will examine how L3 + learners chose to visually express how they perceive their person, role and social identities, and how their perception changes over time, as a result of acquiring a new language. The study will test subjects over a period of one year, during which 4 tasks will be administered.

---

## Multiracial identities as locally constructed styles

Emma Kogure Wilkinson  
Northwestern University, Evanston, USA

An ongoing shift in the approach to understanding racialized identity has been to capture its linguistic construction in varied local contexts (e.g., Wong & Hall-Lew, 2014). Studies pursuing this line of inquiry have focused on monoracial individuals and have not always considered the potential for the modulation of local racialized identity construction based on an interlocutors' identity. Historically, work on the linguistic construction of racialized identities has largely overlooked multiracial identities (c.f. Holliday, 2019). Recent shifts toward capturing the viewpoints and experiences of multiracial individuals note the range in multiracial self-identification (e.g., Holliday, 2016). How might co-construction of local racialized identity occur in multiracial individuals, and how might this vary by interlocutor? This study seeks to understand localized, racialized identity construction, acknowledging the potentially flexible co-construction of racialized identity depends on interlocutors' identities.

Given the potential flexibility in mixed-race racialized self-identification (Rockquemore & Brunisma, 2008), there is opportunity for varying co-construction of racialized identities with interlocutors of different racialized identities. Mixed white and Asian Chicagoans can construct their identities in a multitude of ways, modulated by numerous social factors and localized in specific places and communities (Rockquemore & Brunisma, 2008). The highly segregated and predominantly white, Black, and Hispanic nature of Chicago (US Census Bureau, 2020) creates a unique social setting where neighborhoods are culturally and demographically distinct. Two emblematic features of the Northern Cities Vowel Shift (NCVS), a salient chain shift in Chicago, are a raised, fronted TRAP and a fronted LOT. D'Onofrio and Benheim (2020) found reversal of the NCVS correlates with orientation away from a locally constructed 'white Chicagoan' persona. Style shifting offers another perspective on analyzing linguistic variation, further framing and allowing for the observation of situational identity construction modulated by the interlocutors' identities (e.g., Hall-Lew & Boyd, 2017). This demonstrates the need for research on style shifting of racialized identity construction among mixed-race individuals. I propose a study to observe how two mixed Asian and white English speakers in the Chicagoland area situationally construct their racialized identities in (i) a sociolinguistic interview and in (ii) three self-recorded conversations.

The interview will center questions on participants' relationships with the Chicagoland area and their home neighborhood, their sense of belonging to these places, and how they contextually experience their racialized

identity. Participants will then self-record conversations with three different social groups, noting their interlocutors' identities and the nature of their relationships. I predict participants' variable multiracial identity and orientation to a Chicagoan identity (Wong & Hall-Lew 2014), and their identities' proximity to 'white Chicagoan' (D'Onofrio & Benheim, 2020), will be reflected in TRAP and LOT formants closer to or further from NCVS depending on conversational context (Hall-Lew & Boyd, 2017). Comparison between each participants' vowel proximity to Northern Cities-shifted vowels as a function of their individual place-based and racialized identities to 'white Chicagoan' and their different interlocutors will enhance our understanding of local racialized identity in that it is situationally co-constructed.

D'Onofrio, A., & Benheim, J. (2020). Contextualizing reversal: Local dynamics of the Northern Cities Shift in a Chicago community. *Journal of Sociolinguistics*, 24(4), 469–491. <https://doi.org/10.1111/josl.12398>

Hall-Lew, L. and Boyd, Z. (2017). Phonetic Variation and Self-Recorded Data. University of Pennsylvania Working Papers in Linguistics: Vol. 23: Iss. 2, Article 11. <https://repository.upenn.edu/pwpl/vol23/iss2/11>

Holliday, N. R. (2016). Identity Performance among Black/Biracial Men through Intonation: Examining Pitch Accents and Peak Delay. *University of Pennsylvania Working Papers in Linguistics*, 22(2). <https://repository.upenn.edu/pwpl/vol22/iss2/9>

Holliday, N. R. (2019). Multiracial identity and racial complexity in sociolinguistic variation. *Language and Linguistics Compass*, 13(8). <https://doi.org/10.1111/lnc3.12345>

Rockquemore, K., & Brunson, D. L. (2008). *Beyond Black: Biracial Identity in America*. Rowman & Littlefield Publishers.

U.S. Census Bureau. (2020). U.S. Census Bureau quickfacts: Chicago City, Illinois. United States Census Bureau. <https://www.census.gov/quickfacts/chicagocityillinois>

Wong, A. W., & Hall-Lew, L. (2014). Regional variability and ethnic identity: Chinese Americans in New York City and San Francisco. *Language & Communication*, 35(1), 27–42. <https://doi.org/10.1016/j.langcom.2013.11.003>

---

## Taking center stage: Measuring variation and change across the center and periphery of lexical fields

Kaleigh Woolford

University of Toronto, Toronto, Canada

To account for nuanced differences in meaning between lexical alternatives, recent variationist studies have invoked Sankoff et al.'s (1978) notion of semantic, or lexical, fields (i.e., meanings and their lexical choice sets) to circumscribe the context of lexical variables (Tagliamonte and Brooke 2014, Tagliamonte and Pabst 2020). Cognitive studies, however, offer a number of theoretical predictions regarding synchronic and diachronic characteristics of lexical fields that have yet to be addressed through systematic quantitative analysis. This project launch poster will present a novel method of analyzing lexical variation and change via contemporary computational tools to account for two key theoretical predictions: first, some members of a lexical field are more representative (i.e., more prototypical) of that field than others; and second, this variability creates internal structure such that lexical fields comprise a core of central members that are the best, or most prototypical, examples of that field, and a periphery of other, less prototypical members (Geeraerts 1989, 2010). Drawing on these insights, this project launch poster will consider the following questions:

- I. Is the synchronic structure of lexical fields predictable? Tagliamonte and Pabst (2020) hypothesize that one meaning may only be encoded by up to three main forms, suggesting that lexical fields may have predictable restrictions regarding their central vs. peripheral members.**
- II. How do diachronic trajectories of change interact with the structure of a lexical field? Tagliamonte and Brooke (2014: 33) find that “while synonyms are resplendent within semantic fields [...] [fields] appear to be dominated by one form and then another”. Is this a common process of peripheral members becoming more central? If not, how else do innovative members take hold?**
- III. Do changes at an individual level (i.e., the shift into or development of a new meaning) affect the lexical field more broadly? Existing diachronic studies have shown that a single lexeme undergoing change in its membership status (i.e., from a central to peripheral member, or vice versa) can result in global reorganization within a field (e.g., Begley 2018, Molina 2003).**

To address question **I**, I will perform a synchronic study of adjectival field structure based on data from the Corpus of Contemporary American English (Davies 2008–). Following methods from Leeuwenberg et al. (2016), I will use computational, vector-space representations of meaning (*word2vec*, Mikolov et al. 2013) as a data-driven approach of extracting variants from a broad sample of adjectival fields (e.g., the field of STRANGENESS or ANGER), as well as to measure each member's degree of centrality (Xu et al. 2020). To address questions **II** and **III**, I will then perform a diachronic comparison of these fields using the Corpus of Historical American English

(Davies 2010), where I will also employ variationist methods to compare changes in a variant's probabilistic use against shifts in their level of centrality over time. Together, these methods will reveal how processes of change (e.g., lexical replacement) interact with the structured nature of lexical fields, allowing us to probe deeper into the mechanisms motivating lexical variation and change.

- Begley, M. (2018). The Middle English lexical field of INSANITY: Semantic change and conceptual metaphor. Unpublished PhD thesis. University of Manchester.
- Davies, M. (2008-). The Corpus of Contemporary American English (COCA). Available online at <https://www.english-corpora.org/coca/>.
- Davies, M. (2010). The Corpus of Historical American English (COHA). Available online at <https://www.english-corpora.org/coha/>.
- Geeraerts, D. (1989). Prospects and problems of prototype theory. *Linguistics* 27: 587–612.
- Geeraerts, D. (2010). *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Leeuwenberg, A., Vela, M., Dehdari, J., & van Genabith, J. (2016). A minimally supervised approach for synonym extraction with word embeddings. *The Prague Bulletin of Mathematical Linguistics* 105(1): 111–142.
- Molina, C. 2003. Semantic interface in the entrenchment of loanwords and the estrangement of cognates. *Studia Anglica Resoviensia* 2: 85–94.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Proceedings of the Conference and Workshop on Neural Information Processing Systems (NIPS13)*: 3111–3119.
- Sankoff, D., Thibault, P. & Bérubé, H. (1978). Semantic field variability. In D. Sankoff (ed.), *Linguistic Variation: Models and Methods*. 23–43. New York: Academic Press.
- Tagliamonte, S. A., & Brooke, J. (2014). A weird (language) tale: Variation and change in the adjectives of strangeness. *American Speech* 89(1): 4–41.
- Tagliamonte, S. A. & Pabst, K. (2020). A cool comparison: Adjectives of positive evaluation in Toronto, Canada and York, England. *Journal of English Linguistics* 48(1): 3–30.
- Xu, A., Stellar, J. & Xu, Y. (2020). Prototype theory and emotion semantic change. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*: 730–336.

---

## The floating bubble: Linguistic innovations of the highly mobile Third-Culture Kids (TCKs)

Mechelle Wu

University of Toronto, Toronto, Canada

With globalization on the rise, Third-Culture Kids (TCKs), those who live internationally mobile lifestyles during their developmental years (Useem & Downie, 1976), are also increasing, with O'Bryan claiming that an estimated 230 million people could identify themselves as one as of 2022. Despite the growing TCK population, there is limited linguistics research on them, with existing studies on TCKs being recent (see Starr et al., 2019; Starr et al., 2017; Vowell, 2022) and centered around perception and acquisition differences instead of the features that make them a distinct group.

The present pilot study explores the variety of English spoken by this non-geographically-rooted, border-transcending speech community that functions in a “TCK bubble.” Operationalized as individuals who have lived in upwards of three host countries between the ages of 5 and 18 for at least 4-6 years (Tan et al., 2021, pp. 90, 95), the three TCKs under investigation are multilingual L1 English speakers aged 17-18, with different places of birth and residence. Each participated in a modified virtual sociolinguistic interview (Hall-Lew & Boyd, 2020; Labov, 1984), coined the 360 Task.

Despite their varied backgrounds, the similarly diverse linguistic input and overlapping space in society TCKs occupy result in shared linguistic innovations, which also call for the re-evaluation of established boundaries of sociolinguistic concepts. One such finding is the higher degree of intraspeaker variation to index the TCK identity, in contrast with the common notion that the favoring of variants indexes membership/affiliation. Specifically, with the three speakers, there was the clear alternation of place names. For example, two speakers used both cafeteria and caf to mean “place in a school where lunch is eaten.” The third speaker used homeroom, advisory, and S-L-A to mean “place where a subset of students gathers with the same teacher, usually for attendance or administrative purposes.” Notably, this usage of multiple variants reveals that online, simulated settings also produce location effects (Hay et al., 2017). Other manifestations of increased and marked intraspeaker variation were also present, such as “synonym doubling” (e.g., Our year has- is, like, the most cliquy. Or the most segregated). Furthermore, the data collected hints that the line between code-switching/translanguaging and style-shifting may be blurred (e.g., She was uninvited from a party from one of the bule-s. = faithful Indonesian phonology, unfaithful Indonesian morphology/English morphology). Unique to TCKs, these linguistic practices are products of prolonged and repeated language contact, the multilingualism of its members, and the celebration of linguistic diversity embedded in their international/intercultural school curriculum (Bell-Villada et al., 2011; Van Reken et al., 2017).

This qualitative analysis, aided by a novel data collection method optimized for TCKs, is especially relevant now because this rapidly growing population often ends up immigrating to metropolises, positioning them as leading brokers of change (Labov, 2003, p. 9). As a result, the examination of TCKs has significant theoretical and empirical implications within the study of language variation and change that highlight the complexity and richness of contemporary society.

- Bell-Villada, G. H., Sichel, N., Eidse, F., & Orr, E. N. (2011). *Writing out of limbo: International childhoods, global nomads and third culture kids*. Cambridge Scholars Pub.
- Hall-Lew, L., & Boyd, Z. (2020). Sociophonetic perspectives on stylistic diversity in speech research. *Linguistics Vanguard*, 6(s1). <https://doi.org/10.1515/lingvan-2018-0063>
- Hay, J., Podlubny, R., Drager, K., & McAuliffe, M. (2017). Car-talk: Location-specific speech production and perception. *Journal of Phonetics*, 65, 94–109. <https://doi.org/10.1016/j.wocn.2017.06.005>
- Labov, W. (1984). *Field Methods of the Project on Linguistic Change and Variation*. In J. Baugh & J. Sherzer (Eds.), *Language in use: Readings in sociolinguistics* (pp. 28–66). essay, Prentice-Hall.
- Labov, W. (2006). *Some Sociolinguistic Principles*. In C. B. Paulston & G. R. Tucker (Eds.), *Sociolinguistics: The essential readings*. essay, Blackwell Publishing.
- Pollock, D. C., E., V. R. R., & Pollock, M. V. (2017). *Third culture kids: Growing up among worlds*. Brealey Publishing.
- Starr, R. L. (2019). Attitudes and exposure as predictors of -t/d deletion among local and expatriate children in Singapore. *Language Variation and Change*, 31(3), 251–274. <https://doi.org/10.1017/s095439451900022x>
- Starr, R. L., Theng, A. J., Wong, K. M., Tong, N. J., Ibrahim, N. A., Chua, A. M., Yong, C. H., Loke, F. W., Dominic, H., Fernandez, K. J., & Peh, M. T. (2017). Third culture kids in the Outer Circle: The development of sociolinguistic knowledge among local and expatriate children in Singapore. *Language in Society*, 46(4), 507–546. <https://doi.org/10.1017/s0047404517000380>
- Tan, E. C., Wang, K. T., & Cottrell, A. B. (2021). A systematic review of third culture kids empirical research. *International Journal of Intercultural Relations*, 82, 81–98. <https://doi.org/10.1016/j.ijintrel.2021.03.002>
- Useem, J., & Useem, R. (1967). The interfaces of a binational third culture: A study of the American community in India. *Journal of Social Issues*, 23(1), 130–143. <https://doi.org/10.1111/j.1540-4560.1967.tb00567.x>
- Useem, R., & Downie, R. (1976). *Third Culture Kids*. *Today's Education*, 103–105.
- Vowell, B. (2022). *Dialect change in “returning” third culture kids* (dissertation). Te Herenga Waka—Victoria University of Wellington. <https://doi.org/10.26686/wgtn.19524670>

---

## Language shift and morphosyntactic change in Tgdaya Seediq

Yi-Chi Wu

Harvard University, Cambridge, USA

This project aims to describe ongoing morphosyntactic changes in Tgdaya Seediq—an Austronesian language of Taiwan—in the context of language shift due to contact with the dominant Mandarin speaking population of the island. In our fieldwork, we have observed that the morphologically rich voice and case system in Seediq has reduced and shifted in semantic function, alongside a switch from VSO to SVO word order that is also found in Mandarin. This project seeks to 1) describe the morphosyntactic variation across different generations of speakers, 2) explore its implications for morphosyntactic theory and language shift, and 3) understand the ideologies within the Tgdaya Seediq community towards their linguistic landscape.

In comparison to literature on Tgdaya Seediq over ten years ago (Holmer 1996, Lin 2005)—and on other Seediq varieties (Oiwa 2014, Song 2015)—the language spoken today appears to differ in several notable ways. First, morphological voice alternations have either been reanalyzed as aspectual distinctions in the case of patient voice (PV) and locative voice (LV) (see example 1), or have lost their productivity and been fossilized in the case of instrumental voice (IV). Second, nominal arguments are no longer consistently case marked with nominative or genitive particles, with the nominative particle acquiring a focus-like interpretation instead (see example 2). Other significant changes include shift of adverbs from verbs to adjuncts in terms of voice and inflectional morphology, and increased use of paraphrastic constructions for tense and aspectual information.

The goal of this project is thus to document how these changes are emerging in Tgdaya Seediq, intentionally including members who are not fully fluent. Through semi-structured narratives and conversations, we aim to investigate which speakers are participating in these changes and in what contexts, and hope to compare these data to existing archival materials (NTU 2020, ILRDF 2022). We will conduct sociolinguistic interviews to understand the extent of metalinguistic awareness about language variation in the community, and what attitudes members hold towards preserving certain aspects of their language which can sometimes be at odds with efforts to make reclamation more sustainable. Based off the observed changes in Tgdaya Seediq, the central research question of the project is this: in small and under-resourced shifting communities, how much of language change can be attributed to grammatical influence of the contact language, as opposed to the inherent progression and characteristics of language change and loss? In other words, are the observed changes expected in terms of structural similarities to

Mandarin, and are they found elsewhere in other indigenous communities of Taiwan of similar and/or different circumstances?

Furthermore, since the nature of Austronesian voice systems still remains controversial (Richards 2000, Pearson 2005, Chen 2017), we wish to bring in new evidence by assessing the likelihood of certain morphosyntactic changes in Tgdaya Seediq given the different theoretical approaches to the system. We hypothesize that Austronesian voice systems reflect information structural configurations, due to its potential to shift in semantic function to represent properties such as aspect, topic, and focus.

Chen, Victoria. 2017. A reexamination of the Philippine-type voice system and its implications for Austronesian primary-level subgrouping. Doctoral dissertation, University of Hawaii at Manoa.

Holmer, Arthur. 1996. A parametric grammar of Seediq. Lund University Press.

Indigenous Languages Research and Development Foundation (ILRDF). 2022. Indigenous Languages Online Dictionary and Resources. <https://web.klokah.tw/>

Lin, Hsiu-hsu. 2005. The grammaticalization of tense/aspect auxiliaries in Seediq. *Concentric* 31(2): 111-132.

National Taiwan University (NTU). 2020. Corpus of Formosan Languages. <https://corpus.linguistics.ntu.edu.tw/>

Oiwa, Mayumi. 2014. Adverbial verb constructions in Truku Seediq. Working Papers in Linguistics: University of Hawai'i at Mānoa 45(1).

Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language & Linguistic Theory* 23(2): 381-457.

Richards, Norvin. 2000. Another look at Tagalog subjects. *Formal issues in Austronesian linguistics*: 105-116.

Song, Li-mei. 2015. 賽德克語語法概論 [Seediq Language Grammar]. Taiwan: 原住民族委員會 [Council of Indigenous Peoples].

---

### Three-way variance in scope judgments in Mandarin

Hongchen Wu

Georgia Institute of Technology, Atlanta, USA

As widely noted, and experimentally attested to (Huang 1982, Aoun and Li 1993 and subsequent literature), Mandarin simple transitives (1) are unambiguous. However, according to an empirical observation made by Wu et al. (2018), examples like (2), although a simple transitive, can be read ambiguously wrt to scope under proper contexts.

(1) *yi-ge-xuesheng du-le mei-ben-shu.*

one-CL-student read-ASP every-CL-book

'A student read every book.' ( $\exists > \forall$ ;  $*\forall > \exists$ )

(1) (*Anzhao falü guiding*) *san-ming jingcha kanshou meige chukou.*

(As law demand) three-CL police.officer guard every exit

'As required by law, three policemen guard every exit.' ( $3 > \forall$ ;  $\forall > 3$ )

Following Kuroda 1972'sthetic-categorical distinction where categorical predications attribute properties to particular individuals whilethetic predictions assert general regularities or the existence of events, Larson and Wu (2018) argues that, understood thetically, the inverse scope reading of (2) is even pragmatically preferred and this interpretation is enhanced by a parenthetical like *anzhao falü guiding*. However, Lee (2019) proposes an alternative account of inverse scope in (3), suggesting that it results from the fact that "The object is really a locative, so the greater tendency for inverse scope is a result of the role of the thematic hierarchy, with locatives taking priority over the theme in scope assignment".

No studies have done a large-scale experiment on whether Mandarin simple transitives like (2) allow inverse scope reading and whether a thematic role is the decisive factor for the inverse scope reading, to the best of my knowledge. Therefore, to test the empirical observation on the scope ambiguity of Mandarin matrix transitives like (2), we conducted an untimed, offline experiment and will report the design and results of this experiment here.

The task used in this experiment is a multiple-choice task, a revised version of the "How many" task used by Anderson (2004). The contexts given for each target sentence did not favor a particular scope reading. Each target



sentence was followed by a comprehension question with four possible answers (Answer A corresponds to surface scope reading; Answer B and C and D: inverse scope reading), as in (3).<sup>1</sup>

(3) *Context*: The person in charge of the stadium security asked the Assistant to the Chief at the Police Station about the specific arrangements of the security working for the stadium. The bold and black sentence below is the reply of the Assistant to the Chief.

*Target sentence*: Four policemen guard every exit.

*Comprehension question and available choices*: Assume that there is a total of five exits in the stadium. How many policemen are needed to undertake the security tasks of guarding the stadium? A. 4 B. 4 to 20 C.  $5 * 4 = 20$  D. more than 20

To minimize the potential influence that some contexts may bias towards a specific reading (for example, one may think that the context regarding policemen guarding exits as in (3), given the pragmatics that more policemen generally mean better safety), we added more non-locative objects and diversified the scenarios that were used for the contexts in the stimuli. The contexts involve policemen guarding exits, nurses taking care of ICUs, volunteers assisting with exhibition stands, professors reviewing dissertations, notaries supervising matches, students decorating classrooms, kids recording songs, and guards surveilling arrested officials.

Two factors were controlled for in the stimuli: the types of clauses (matrix transitives, embedded clauses of *shuo* ‘say’) and the presence of *hui* (future tense ‘will’); 4 (=2\*2) conditions in total. 8 sets of 4 sentences (one sentence with *hui* and one sentence without *hui* in each set) were created as the target sentences. 32 target sentences were randomized with 112 fillers and distributed across 8 lists in a Latin Square Design. Each participant was presented with 4 target sentences (one sentence for each condition) intermingled with 14 fillers. Among the 14 fillers each participant would read, there are four “dummy” fillers to check if participants pay attention, each of them associated with four possible choices but only a single correct answer.

The experiment was conducted through Qualtrics. 134 native Mandarin speakers participated in this experiment. 89 out of them completed the survey and only their data were included in the results (Table 1). The data show a three-way variance: for each condition, a group of participants chose only surface scope reading (i.e., choosing Answer\_Type A only), another group of participants chose only inverse scope reading (i.e., choosing from a subset of Answer\_Type B, C, D only), and another group of participants chose ambiguity (both surface and inverse scope reading).

About half of the participants considered an inverse scope reading available for simple transitives, empirically confirming Wu et al. (2018)’s observation. Participants considered an inverse scope reading available for stimuli where the object is not locative, arguing against the thematic account proposed by Lee (2019). The three-way variance in the scope judgments suggests that participants may project different structures when constructing scope interpretations. It, in fact, supports Larson and Wu (2018)’s proposal which employs thematic- categorical distinction to account for quantifier scope in Mandarin. The notions ‘thematic’ vs. ‘categorical’ in the Kuroda/Brentano theory differs from other alternative accounts for Mandarin quantifier scope in appealing to a distinction in speaker judgment, and thus something that is fundamentally “internalist” in character.

Anderson, C.. 2004. The Structure and Real-Time Comprehension of Quantifier Scope Ambiguity. Ph.D. thesis. Northwestern University.

Larson, R., and H. Wu. 2018. Quantifier Scope and Topicality in Mandarin. Presentations at IACL-26. University of Wisconsin-Madison. (May 6, 2018).

Lee, T.H. 2019. The reality of scope isomorphism in Chinese: Some methodological issue in the investigation of quantifier scope. Presented at the International Symposium on Formal Approaches to Meaning in Chinese. Beijing Language and Culture University, June 22-23, 2019

Kuroda, S.-Y.. 1972. The categorical and the thematic judgment. *Foundations of Language* 9(2): 153-185.

Wu et al. 2018. Rethinking quantifier scope in Mandarin. In Proceedings of NELS 48, vol. 3, 257-263. Amherst: University of Massachusetts, GLSA.

---

<sup>1</sup> Here we only present the English translation of the Mandarin version. A multiple-choice task, instead of a forced choice task, was used to encourage participants to choose all possible readings they can get, not just a preferred scope reading. But, we are aware that offering four possible answers and the concrete number calculations (as given in the C option in (3)) might induce a cognitive burden and draw extra attention to the C option. We will conduct a following-up experiment with only two possible answers given (as the exactly same settings in Anderson (2004)) in the future, to see whether the setting of the four possible answers indeed would make an influence on the results.

---

## “The bureaucrat”, “the neighborhood grandma”, and “the childish girl” in Beijing: Uncovering the social meaning of stylistic timing variation

Robert Xu

Stanford University, Stanford, USA

Speakers vary durations in speech to create perceptual timing patterns fit for situational use. However, few sociolinguistic studies have captured such variation beyond phonetic descriptions of lengthening and rhythm, and connect these patterns to their expressiveness in interactions. This study examines how stylistic performances of character types in Beijing Mandarin vary in the temporal domain, and how such variation takes on social meanings.

Character types (Agha 2003) are abstractions of salient and performative social images of personhood enregistered with linguistic features. They establish the most conventional associations between variation forms and social meanings for personae in a community. This study highlights three such types in Beijing - Bureaucrat, Neighborhood Grandma, and Childish Girl. I captured these types in an interactive game with 63 speakers (19M/44F, Age 19-63) in triads, where they each took turns performing character types without scripts. For a meta-discursive analysis, the participants also joined a focus group discussing and evaluating each type in the local discourse.

The performances were transcribed, forced-aligned, and acoustically analyzed for durational measurements. The results show that all character types are packaged with unique timing profiles that vary in speech rate, rhythm, juncture characteristics, and lengthening. Furthermore, these timing profiles are associated with conversational strategies that typify the character types in their interactional context, evidenced in the meta-discursive analysis.

For example, the Bureaucrat uses a significantly slow speech rate, long pauses within utterances, and long and variable utterance duration. This timing profile coincides with conversational organizers such as floor-holding filled pauses. The specific pacing and meticulous articulation enable the Bureaucrat to sound important and official, utilizing the speaking time as capital for power. The Grandma uses a fast speech rate sprinkled with syllable contractions, a feature that indexes “smooth” Beijingness (Zhang 2007). Moreover, speakers use rush through (Schegloff 1982) and abrupt join (Local & Walker 2004) at junctures via accelerated speech rate, as well as using address terms and interjections as modular pivots (Walker 2007) – all to construct a nosy persona gossiping around the neighborhood. The Childish Girl uses highly variable syllable durations and extra lengthened utterance finals. These features co-occur with the prevalent use of turn-initial discourse markers and sentence final particles. Together these timing features and conversational organizers portray a hyper-feminine persona that uses childlike stubbornness to attain attention and gain favors.

Based on these results, I argue that the timing profiles are essential to the expressiveness of these styles. Character types are intrinsically dialogical as they utilize conversational organizers and strategies such as discourse markers and salient timing variation at what would be transition-relevance places in interactions. I propose that what makes the surface timing variation socially meaningful is the underlying conversational strategies that motivate them. By displaying the conversational strategies as various perceptible timing signatures, the character types become interactional frames (Goffman 1974) that preserve the specific relationships between the speaker and the addressee(s). Methodologically, sociolinguists can reach a better understanding of temporal variation by connecting quantitative sociophonetic descriptions to the conversation analysis toolkit.

Agha, Asif. 2003. *The Social Life of Cultural Value*. Language and Communication. Vol. 23.

Goffman, Erving. 1974. *Frame Analysis: An essay on the organization of experience*. Boston: Northeastern University Press.

Local, John, & Gareth Walker. 2004. Abrupt-joints as a resource for the production of multi-unit, multi-action turns. *Journal of Pragmatics*, 36 (8), 1375–1403.

Schegloff, Emanuel. 1982. Discourse as an interactional achievement: Some uses of ‘uh huh’ and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk*. Washington, DC: Georgetown University Press.

Walker, Gareth. 2007. On the design and use of pivots in everyday English conversation. *Journal of Pragmatics*, 39 (12), 2217 – 2243.

Zhang, Qing. 2017. *Language and Social Change in China: Undoing Commonness through Cosmopolitan Mandarin*. Routledge.

---

## Nasalization Change over Time in Michigan English

Yongqing Ye

Michigan State University, East Lansing, USA

This paper examines articulatory variation for a subphonemic representation, nasalization, in Michigan English. Research on sound change has primarily focused on changes to the phonetic implementation of individual phonemes, or to more phonological properties like allophonic conditioning. More research on community-level change to non-segmental features is needed to deepen our understanding of how such features change over time. Michigan English has been claimed to have the presence of nasal peaks and anti-formants in non-nasal environments

(Plichta 2004), but this has not been tested in real or apparent time. Prior work on Philadelphia English (Zellou and Tamminga 2014), however, confirmed that coarticulatory vowel nasalization can be targeted for change. The present study therefore tests whether degree and duration of vowel nasalization across different age groups in two distinct phonological contexts (pre-oral [CVC] and pre-nasal [CVN]) in Michigan English exhibit change in apparent time.

We analyzed self-recorded audio dairies in the state of Michigan (Sneller et al. 2022). A total of 22 hours of recordings from 54 native speakers of Michigan English (aged 18-74, 20 male, 32 female and 2 other) was analyzed. The degree of nasalization was measured with the acoustic value A1-P0 (the amplitude of the first harmonic peak F1 minus the amplitude of the lowest nasal peak) (Chen 1997) in four target vowels – thought, lot, face and goat. The formant extraction was automated using a modified Praat script designed to measure nasality in acoustic data (Styler 2017). Measurements were taken at ten timepoints during the vowel in three different phonological environments, the baseline [SVS], pre-oral [CVC] and pre-nasal [CVN] contexts. The duration of nasalization was calculated as the significant divergent timepoint between the nasal trajectories in pre-nasal and pre-oral contexts and the baseline. The results were analyzed in R (R Core Team 2021) with linear regression models.

The results confirm Plichta's (2004) observation of nasality even in pre-oral contexts and show that as birth year increases, the degree of nasalization decreases (Figure 1) but the variability in the degree of vowel nasalization increases (Figure 2). This pattern is present across all contexts for all four vowels. The results for the duration of nasalization find that birth year does not have an effect on the duration of nasalization across all vowels and contexts in comparison with the baseline. Further examination shows participants, unsurprisingly, fully nasalize in pre-nasal context but exhibit a bimodal distribution in pre-oral context – they are either fully nasalizing their pre-oral vowels or not nasalizing at all.

To summarize, younger speakers of Michigan English exhibit less nasality, in the degree of nasalization but not the duration, than older speakers. Additionally, young people's production of nasalization has a higher degree of variability. This could be an indication that nasalization in Michigan English could be undergoing generational change, as increased interspeaker variability is expected for the early stages of certain types of phonological change (Fruehwald 2013; Sneller 2018). This paper underscores that language change in subphonemic representations may be fruitfully studied in apparent time.

Chen, Marilyn Y (1997). "Acoustic correlates of English and French nasalized vowels". In: *The Journal of the Acoustical Society of America* 102.4, pp. 2360–2370.

Fruehwald, Josef (2013). *Phonological involvement in phonetic change*. PhD thesis. University of Pennsylvania.

Plichta, Bartłomiej (2004). *Interdisciplinary perspectives on the Northern Cities chain shift*. PhD thesis. Michigan State University.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Sneller, Betsy (2018). *Mechanisms of phonological change*. PhD thesis. University of Pennsylvania.

Sneller, Betsy, Suzanne Evans Wagner, and Yongqing Ye (2022). "MI Diaries: ethical and practical challenges". In: *Linguistics Vanguard* 8.s3, pp. 307–319.

Styler, Will (2017). *Nasality Automeasure Script Package*.

[https://github.com/stylerw/styler\\_praat\\_scripts/tree/master/nasality\\_automeasure](https://github.com/stylerw/styler_praat_scripts/tree/master/nasality_automeasure).

Zellou, Georgia and Meredith Tamminga (2014). "Nasal coarticulation changes over time in Philadelphia English". In: *Journal of Phonetics* 47, pp. 18–35.

---

## Modeling sociophonetic variation in L2 varieties of English: considerations of proficiency and norm orientation

Ping Hei Yeung

Georgetown University, USA

English, a global lingua franca, is spoken by millions of people worldwide as an L2. In Outer Circle countries like Nigeria and the Philippines, it is caused by British and American colonization, whereas in Expanding Circle countries like the Netherlands and Japan, it is the result of globalization of English but without direct colonial influence (Kachru, 1986). Despite the ubiquity of L2 English varieties, there is yet to be a unified model that explores their sociophonetic variation, especially one that incorporates theories of both second language acquisition and variationist sociolinguistics. Thus, I propose a new model for analyzing variation in L2 English varieties that considers both the developmental factor of language proficiency and the sociolinguistic factor of norm orientation.

As predicted by models of L2 phonology acquisition (Best & Tyler, 2007; Flege, 1995), L2 speakers are influenced by their L1 phonology, causing a transfer of L1 phonological features into their L2. If they take the norm of L1 English speakers (e.g., British English) as their target of acquisition, proficiency should mediate inter-speaker variation, with greater convergence to L1 varieties with increasing proficiency. While the phonology of L2 speakers is almost never identical to that of L1 speakers due to developmental constraints (Flege et al., 1999; Piske et al.,

2001), the speech production of speakers with higher proficiency may be closer to L1 norms than speakers with lower proficiency due to lesser degree of L1 transfer.

Then, as demonstrated by Schneider's Dynamic Model (2003, 2007), L2 English-speaking communities may undergo endonormative stabilization in which a local standard of English replaces L1 varieties to be the norm. In other words, L2 English speakers do not always treat L1 varieties as their goal of acquisition. Depending on how fully developed this local norm is, the relation between proficiency and similarity to L1 norms may differ. In societies where the L1 norm persists, there would be a clearer trend of convergence to L1 varieties with increasing proficiency. In societies where a local norm exists however, one may observe less convergence, and deviance from L1 norms may persist among speakers with higher proficiency as L1 varieties are not their goal of acquisition.

To test the validity of this model, I will examine vowel production of L2 English speakers in societies with (e.g., the Philippines (Borlongan, 2016)) and without (e.g., the Netherlands (Edwards, 2014)) a local norm of English. Speakers of different proficiency levels will participate in a sociolinguistic interview and an English proficiency test. To minimize bias against non-L1 varieties of English, a new assessment method will be developed and employed based on the guidelines of Canagarajah (2006). Formant measurements of all monophthongs and diphthongs will be extracted from the interviews, and generalized additive mixed models will be used to test the effects of proficiency and norm orientation on the degree of similarity of formant trajectories to Received Pronunciation and General American English, the two L1 models commonly used in English education worldwide (Dziubalska-Kołaczyk & Przedlacka, 2008).

- Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege* (pp. 13–34). <https://doi.org/10.1075/llt.17.07bes>
- Borlongan, A. M. (2016). Relocating Philippine English in Schneider's dynamic model. *Asian Englishes*, 18(3), 232–241. <https://doi.org/10.1080/13488678.2016.1223067>
- Canagarajah, S. (2006). Changing Communicative Needs, Revised Assessment Objectives: Testing English as an International Language. *Language Assessment Quarterly*, 3(3), 229–242. [https://doi.org/10.1207/s15434311laq0303\\_1](https://doi.org/10.1207/s15434311laq0303_1)
- Dziubalska-Kołaczyk, K., & Przedlacka, J. (Eds.). (2008). *English Pronunciation Models: A Changing Scene* (2nd ed.). Peter Lang.
- Edwards, A. (2014). *English in the Netherlands: Functions, forms and attitudes* [Ph.D. Dissertation]. University of Cambridge.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, 41(1), 78–104. <https://doi.org/10.1006/jmla.1999.2638>
- Kachru, B. B. (1986). *The Alchemy of English: The Spread, Functions, and Models of Non-native Englishes*. University of Illinois Press.
- Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29(2), 191–215. <https://doi.org/10.1006/jpho.2001.0134>
- Schneider, E. W. (2003). The Dynamics of New Englishes: From Identity Construction to Dialect Birth. *Language*, 79(2), 233–281. <https://doi.org/10.1353/lan.2003.0136>
- Schneider, E. W. (2007). *Postcolonial English: Varieties around the World* (Illustrated edition). Cambridge University Press.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition* (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315370279>

---

## Social Perception and Categorization of Southern Mandarin Accents

Irene Yi, Grace Wong  
Stanford University, Palo Alto, USA

Geographic proximity and frequency-of-exposure affect how listeners categorize different regional accents (Clopper & Pisoni 2004ab, 2007). For example, someone living in a given region breaks talkers of that region up into a high number of specific categories (e.g. “New England”, “North Midland”, “South Midland”, etc.). Frequency/proximity-based accounts also suggest that someone living outside the region would break those same talkers up into fewer, broader categories (“West”). Over the last decade, research has accumulated to suggest that listeners make associations between how people speak and where they live (c.f., Wade, 2022). We extend this line of thinking to accent categorization and investigate the effects of social enregisterment (Agha 2003) on perceptual dialectology. We hypothesize that when listeners hold strong associations between accent and place, these fine-grained categories emerge even for those living outside the region. Mandarin accents spoken in China provide an opportunity to test this hypothesis, as many features of Southern accents of Mandarin have been enregistered in styles and character types through metalinguistic awareness and semiotic salience (e.g. Zhang 2017; Gao & Forrest 2023).

We extend Clopper and Pisoni's (2007) free classification task to 10 Mandarin accents (5 critical Southern accents, 5 distractor Northern accents). This methodology, in which participants create groups from audio stimuli clips with no predefined labels, minimizes imposed geographic structure onto listeners. This probes listeners' organic categorizations of accents. Figure 1 shows the full list of Mandarin accents in the stimuli; Figure 2 shows sample task interfaces. Online metalinguistic commentary (forums, news, pedagogy, etc.) suggests that Southern accents of Mandarin are recognized as more socially distinct than Northern accents, and we hypothesize that this enregisterment asymmetry may be due to China's sociohistorical forces and changing sociolinguistic landscape (cf. Zhang 2017). We expect our 5 critical Southern accents (3-5 speakers each) to be socially enregistered across listeners, with all listeners grouping the voices into a high number of specific categories. As a control, our 5 distractor Northern accents – which are just as phonetically distinct from one another as Southern accents are – are expected to be less enregistered across listener populations, with (Southern) listeners forming a few, broad categories. If classification is driven only by direct-exposure frequency and geographical proximity alone, we expect Northern listeners to make fewer groupings for Southern accents compared to Southern listeners.

We have recorded speakers of 8 accents, and are currently programming the experiments. We aim for data collection of the project's first phase, with only Southern listeners, to be completed by early September. The second phase will extend to Northern listeners, teasing apart how Southern accents are socially enregistered even when (Northern) listeners are not geographically proximal. The proposed analysis includes additive clustering and multidimensional scaling to measure how many groups are created, perceptual similarity of groupings, and which dimensions groups were formed on.

Through this study, we can better understand enregisterment of Mandarin accents in China as a result of social salience, and expand our understanding of social factors underlying dialect perception beyond geographical proximity and exposure frequency.

- Agha, Asif. 2003. The social life of cultural value, *Language & Communication*, Volume 23, Issues 3–4, 2003, Pages 231-273, [https://doi.org/10.1016/S0271-5309\(03\)00012-0](https://doi.org/10.1016/S0271-5309(03)00012-0).
- Clopper, Cynthia., & Pisoni, David. 2004a. Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change*, 16(1), 31-48. doi:10.1017/S0954394504161036
- Clopper, Cynthia. G., & Pisoni, David. B. 2004b. Effects of Talker Variability on Perceptual Learning of Dialects. *Language and Speech*, 47(3), 207–238. <https://doi.org/10.1177/00238309040470030101>
- Clopper, Cynthia. G., & Pisoni, David. B. 2007. Free classification of regional dialects of American English, *Journal of Phonetics*, Volume 35, Issue 3, 421-438, <https://doi.org/10.1016/j.wocn.2006.06.001>.
- Daidone, Danielle, Lidster, Ryan, & Kruger, Franziska. 2023. Free classification as a method for investigating the perception of nonnative sounds. *Studies in Second Language Acquisition* (2023), 1–27. doi:10.1017/S0272263123000050
- Gao, Feier, & Forrest, Jon. Jon Forrest. 2023. Indexical meaning of Mandarin full tone in the construction of femininity: Evidence from social perceptual data. *Language & Communication*, Volume 91, 1-20, <https://doi.org/10.1016/j.langcom.2023.04.001>.
- Wade, Lacey. 2022. Experimental evidence for expectation-driven linguistic convergence: Supplementary material. *Language*, vol. 98 no. 1. doi:10.1353/lan.2021.0097.
- Zhang, Qing. 2017. *Language and social change in China: Undoing commonness through Cosmopolitan Mandarin*. Taylor and Francis. <https://doi.org/10.4324/9781315886251>

---

## Social attribute ratings of Mandarin varieties in different countries

Irene Yi, Grace Wong  
Stanford University, Palo Alto, USA

**Introduction:** Language attitudes are often tied to one's identity with, and orientation towards, one's country. Such attitudes are also reflected in how one makes social judgments about varieties spoken in different countries. This study investigates Singaporeans, Chinese people, and Taiwanese people's social attribute ratings of different countries' varieties of Mandarin. This experiment will replicate previous studies done with Singaporean listeners (Cavallaro et al. 2018, Chong & Tan 2013, a.o.), and crucially extend to more listener populations, i.e. Chinese and Taiwanese listeners, to probe cross-listener-population social meanings. Additionally, the present study distinguishes Beijing Mandarin from Standard Mandarin. These two Chinese varieties are socially and phonetically distinct from one another but often treated as the same variety in perception studies, or at least described ambiguously as to which variety is being represented. Another goal of this study, then, is to understand differences in social meaning between these two varieties within and across listener populations.

**Hypothesis:** How do listeners socially rate different varieties of Mandarin, depending on their background? Our specific hypotheses are:

For Singaporean participants, following previous studies, we hypothesize they will rate one or both Chinese varieties of Mandarin (Beijing/Standard) as high in status, while rating Singaporean Mandarin as higher in

solidarity. Singaporean participants will rate Taiwanese Mandarin as slightly lower than Beijing Mandarin, at least in status traits.

For Chinese participants, extrapolating from the literature, we hypothesize they will rate Standard Mandarin, and perhaps Beijing Mandarin, above Singaporean Mandarin on status and solidarity. However, it is unclear how they will rate Taiwanese Mandarin in comparison to Standard and Beijing Mandarin, given changing social valuations of these varieties (cf. Zhang 2017).

For Taiwanese participants, we hypothesize they will rate Taiwanese Mandarin higher than all other accents in both status and solidarity.

**Methods:** We will conduct a verbal guise experiment where participants listen to clips and complete a social attribute rating task, with scores measured on a Likert scale. The social attributes are drawn from previous literature: kind, honest, humorous, trustworthy, confident, intelligent, educated, and wealthy; “status” vs. “solidarity” traits will be determined through a Principle Component Analysis. Critical trials include clips of Singaporean Mandarin, Beijing Mandarin, Standard (Chinese) Mandarin, and Taiwanese Mandarin recorded speech (reading passage). Filler clips will be other Mandarin varieties/accents of the same sentences. The primary variable of interest (reflected in sampling of participants) is participants’ country of origin. Secondary variables of interest include age, language background, and whether they have moved and/or lived abroad (and where).

Exploratory variables are attribute ratings for filler clips, representing other Mandarin accents we are curious about (in social evaluation), but have been previously unattested in the literature to this end. We will collect data for one listener group (country) in June, with the other two listener groups in following months. These experiments will help us better understand social perceptions of many Mandarin varieties in a systematic way (within/across listener populations), with implications for how socially enregistered these varieties are, to different listener populations (Agha 2003).

Agha, Asif. 2003. The social life of cultural value, *Language & Communication*, Volume 23, Issues 3–4, 2003, Pages 231-273, [https://doi.org/10.1016/S0271-5309\(03\)00012-0](https://doi.org/10.1016/S0271-5309(03)00012-0).

Cavallaro, F., Seilhamer, M. F., Ho, Y. Y., & Ng, B. C. 2018. Attitudes to Mandarin Chinese varieties in Singapore. *Journal of Asian Pacific Communication*, 28(2), 195-225. doi:10.1075/japc.00010.cav

Chong, Rachael & Tan, Ying-Ying. 2013. Attitudes toward accents of Mandarin in Singapore. *Chinese Language & Discourse*. 4. 10.1075/cld.4.1.04cho.

Zhang, Qing. 2017. *Language and social change in China: Undoing commonness through Cosmopolitan Mandarin*. Taylor and Francis. <https://doi.org/10.4324/9781315886251>

---

## Examining rhoticity patterns in L2 English acquisition: A comparative study of Chinese speakers from northern and southern regions of China

Zichuan Yu

National University of Singapore, Singapore, Singapore

Post-vocalic rhoticity is a well-documented linguistic phenomenon observed in various language varieties, including American English and Chinese dialects (Ladefoged and Maddieson, 1990). Previous research has examined the manifestation of post-vocalic /r/ in Chinese and its regional variations. Specifically, northern varieties are mostly rhotic, while southern varieties lack post-vocalic /r/. Despite extensive exploration, there is a need for further investigation that focuses on the cross-linguistic nature of this phenomenon. This study examines the production patterns of English post-vocalic /r/ in Chinese speakers from the northern and southern regions of China using Generalized Additive Mixed Models (GAMMs; Wood, 2006), shedding light on how their native dialects influence the acquisition of this sound in their second language.

Participants of the study were ten postgraduate students from the National University of Singapore, all Chinese citizens born in mainland China. Evenly divided into northern and southern cohorts, the participants represented linguistically diverse backgrounds, including speakers of Beijing Mandarin (north), Hokkien and Cantonese (both south). Data collection involved audio recordings of reading tasks in English and Chinese, as well as questionnaire responses. The readings targeted specific wordlists, passages and (near)-minimal pairs designed to elicit rhotic tokens. F3 values were extracted at every 10% interval of the course of each token, resulting in a dataset of 9031 English and 1430 Chinese data points. After being normalized, the data were then coded and analysed using GAMMs, which provided a robust framework for capturing temporal variations in F3 trajectories and were well-suited for comparing the northern and southern groups.

Findings of this study can be summarized in three layers. First, in analysing Chinese data, a notable distinction emerged between the two regional groups. The southern participants displayed an earlier initiation of rhoticity, indicated by significantly lower F3 values during the initial 0-7% of the time course. Regarding the English data, although a holistic examination of all English tokens showed no significant difference in the two groups, it was found that the vocalic environments preceding the rhotic sounds had a significant impact on rhoticity patterns across

regions. Particularly, a significant interaction effect was observed between region and the NORTH vowel, suggested by pronounced differences in F3 values. Lastly, in comparing the Chinese and English data, an important distinction was observed for the southern participants. They exhibited a significant difference in F3 values between English post-vocalic /r/ and Chinese post-vocalic /r/. However, for the northern group, no significant difference was found across languages. This finding suggests that the influence of native dialects on the production of English post-vocalic /r/ varies between the southern and northern groups, highlighting distinct transfer effects.

This preliminary investigation on post-vocalic /r/ in Chinese speakers provides valuable insights into the nuanced transfer effects from native dialects in second language acquisition. This study also highlights the significant role of vocalic environments in the production of rhoticity and underscores the utility of GAMMs for similar research endeavours. Overall, these findings contribute to our understanding of cross-linguistic phonetic variations and language transfer.

Ladefoged, P., & Maddieson, I. (1990). Vowels of the world's languages. *Journal of Phonetics*, 18(2), 93–122.  
[https://doi.org/10.1016/S0095-4470\(19\)30396-1](https://doi.org/10.1016/S0095-4470(19)30396-1)

Wood, S. (2006). *Generalized additive models: an introduction with R*. Boca Raton: CRC Press.

---

## Artificial accents: assessing phonological variation in voice cloning software

Naitian Zhou

UC Berkeley, Berkeley, USA

Voice cloning technology generates speech audio that resembles a speaker's voice by learning from a small reference audio sample. It is already seeing use in the real world. One popular use-case is generating voice-overs for content creators, but voice cloning has also been used to preserve voices for “users at risk of losing their ability to speak” (as in Apple's newly announced Personal Voice program, which targets people with ALS) and recreate voices of the deceased. Especially with respect to these latter use-cases, it is important to know whether these technologies are accurately modeling the diverse ways in which we speak, including the use of socially meaningful language variables.

Voice cloning systems rely on deep neural networks pre-trained on hundreds of hours of speech recordings from many different speakers (Neekhara et al. 2021; Casanova et al. 2022; Betker 2022). The reference audio sample, which usually only spans a few seconds to a few minutes, is usually encoded as an internal voice representation that the model then uses to generate new speech audio. Because the reference audio is so short, this encoding process requires the model to make inferences about features of the reference speaker's speech. The present question is, do these models accurately represent the systematic phonological variation present in language?

This project aims to understand the extent to which voice cloning systems preserve phonological variation for American English speakers. As a pilot study, I focus on YourTTS, an open-source voice cloning model (Casanova et al. 2022). YourTTS was trained on the voice cloning toolkit (VCTK) dataset, which includes speech data from 110 English speakers with various accents (Yamagishi, Veaux, and MacDonald 2019). I use short clips from the Switchboard corpus as reference voices to synthesize new speech, and measure how the presence of the PIN-PEN merger changes between the reference speech and synthesized speech, as measured by the Bhattacharyya affinity of the midpoint F1 and F2 formants between /ɪ/ and /ɛ/ (Godfrey, Holliman, and McDaniel 1992; Austen 2020). Figure 1 shows clear differences between the reference and synthesized speech. I find that, on average, cloned voices using YourTTS exhibit a decrease in the merger, and this disproportionately affects speakers who have a stronger merger in their reference speech. These preliminary results suggest that YourTTS output might be biased towards the standard variety of American English.

The next phase of research includes expanding the analysis to other phonological variables and probing the outputs of more voice cloning systems, including those currently being offered as products. Additionally, the Switchboard corpus used for the preliminary analysis contains low-fidelity phone recordings, which may unfairly diminish the performance of models. I plan to use a higher-fidelity speech corpus for future experiments.

Austen, Martha. 2020. “Production and Perception of the Pin-Pen Merger.” *Journal of Linguistic Geography* 8 (2): 115–26.  
<https://doi.org/10.1017/jlg.2020.9>

Betker, James. 2022. “TorToiSe Text-to-Speech.” <https://github.com/neonbjb/tortoise-tts>.

Casanova, Edresson, Julian Weber, Christopher Shulby, Arnaldo Candido Junior, Eren Gölge, and Moacir Antonelli Ponti. 2022. “YourTTS: Towards Zero-Shot Multi-Speaker TTS and Zero-Shot Voice Conversion for Everyone,” February.  
<https://doi.org/10.48550/arXiv.2112.02418>.

Godfrey, John J., Edward C. Holliman, and Jane McDaniel. 1992. “SWITCHBOARD: Telephone Speech Corpus for Research and Development.” In *Proceedings of the 1992 IEEE International Conference on Acoustics, Speech and Signal Processing - Volume 1*, 517–20. ICASSP'92. USA: IEEE Computer Society.

Neekhara, Paarth, Shehzeen Hussain, Shlomo Dubnov, Farinaz Koushanfar, and Julian McAuley. 2021. “Expressive Neural Voice Cloning.” In *Asian Conference on Machine Learning*, 252–67. PMLR.

Yamagishi, Junichi, Christophe Veaux, and Kirsten MacDonald. 2019. "CSTR VCTK Corpus: English Multi-Speaker Corpus for CSTR Voice Cloning Toolkit (Version 0.92)." The Rainbow Passage Which the Speakers Read out Can Be Found in the International Dialects of English Archive: ([Http://Web.Ku.Edu/~idea/Readings/Rainbow.Htm](http://Web.Ku.Edu/~idea/Readings/Rainbow.Htm))., November. <https://doi.org/10.7488/ds/2645>.