Session 5B Abstracts

Mining corpora for gestural timing differences as a precursor to metathesis

Tyler Lau, University of California at Berkeley

Introduction: This study explores the gestural timing of stop-sibilant (TS) and sibilant-stop (ST) sequences in the Buckeye Corpus (Pitt et al., 2007), corroborating previous studies suggesting that gestural variability and perceptual optimization play a role in the incipience of metathesis.

Background: ST/TS sequences are well-known to metathesize. Stops have weak internal place cues and are best perceived when adjacent to (preferably preceding) a vowel (Fujimura et al., 1978). Furthermore, an adjacent sibilant may mask the acoustics of a stop (Mielke, 2001) and can trigger segmental order confusion through auditory stream decoupling (Bregman and Campbell, 1971). Perceptual explanations argue that metathesis occurs to bring phonemes into more salient positions, as mirrored diachronically (Steriade, 2001; Hume, 2004) and experiments have shown that VTSV sequences are biased to be heard as VSTV (Graff and Scontras, 2012; Jones, 2016).

Yanagawa (2003) explores the role of production in metathesis. She demonstrates that in Hebrew, consonant clusters show greater variability of gestural timing in word-medial vs. word-onset position and across morphemes vs. within the same morpheme. Yanagawa hypothesizes that weaker gestural cohesion can lead to metathesis, particularly for TS/ST clusters, as has happened in Hebrew. We attempt to replicate this finding in a much larger sample size using English.

Methodology: We extracted 17,686 ST/TS sequences from the Buckeye Corpus of 40 English speakers in Columbus, Ohio. The independent variable was the *log ratio of sibilant to stop duration* to normalize for speech rate. The dependent variables were *position in the word*, *morpheme boundary presence*, and *cluster type* (whether the stop precedes (STV) or follows (VTS) a vowel). Greater gestural variability (reflected in variability of the ratio) is expected in non-initial position and heteromorphemically. If gestural variability is a source of metathesis, VTS sequences should also be expected to be more variable as they are perceptually dispreferred to STV sequences.

Results:

Figure 1: Density of Sibilant to Stop Ratio by Word Position

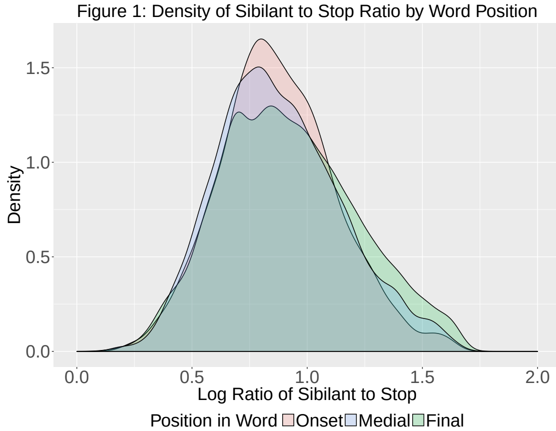
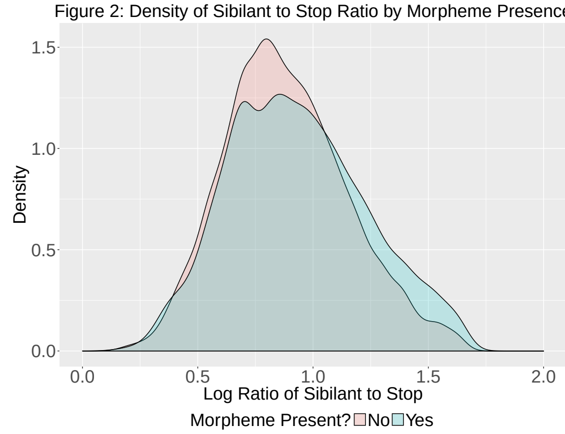


Figure 2: Density of Sibilant to Stop Ratio by Morpheme Presence



A linear mixed effects analysis was carried out with random intercepts for speaker and by-speaker random slopes for the effects of position in the word and morpheme presence. The analysis revealed cluster type and position in the word to be significant (all p's $< .0001$). Postvocalic stop clusters (VTS) have greater sibilant-to-stop ratios than prevocalic ones (STV) (Fig. 4). Medial and final clusters have lower ratios than onset clusters, an unexpected finding. Morpheme presence was not significant ($p = 0.97$). A Levene's Test reveals that the variances of the distributions by word position ($F(2,17673) = 91.94$), morpheme presence ($F(1,17674) = 112.88$), and cluster type

(F(1,13728) = 222.05) are significantly different (all p's < .0001). Clusters show greatest variance in final position, followed by medial, then onset (Fig. 1). Presence of a morpheme boundary also yields greater variance (Fig. 2) and VTS clusters have greater variance than STV ones (Fig. 3).

Figure 3: Density of Sibilant to Stop Ratio by Stop/Vowel Order

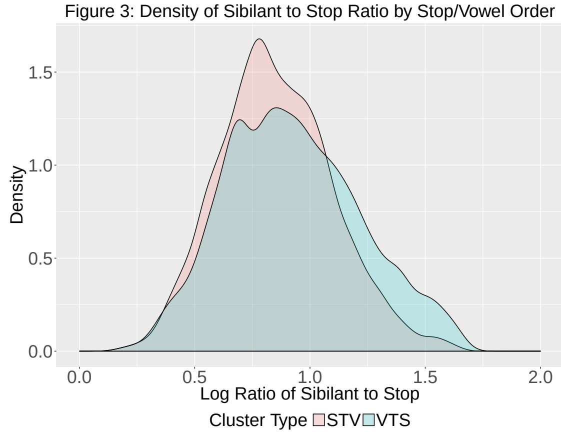
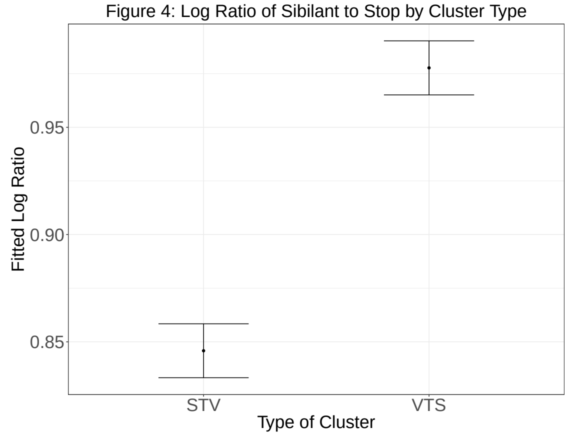


Figure 4: Log Ratio of Sibilant to Stop by Cluster Type



The results corroborate Yanagawa's findings of greater variance in both non-onset and heteromorphemic clusters and also show that word-final clusters have even greater variance than medial ones. Yanagawa's theory of greater variance in unstable clusters is also supported by the greater variance in VTS vs. STV clusters. Furthermore, the model reveals a significantly larger sibilant-to-stop ratio in VTS vs. STV clusters. While the relatively weaker perceptualness of stops in VTS clusters has been attributed to preference for VC over CV transitions, this result also suggests that the longer sibilant may also be more likely to mask the stop and to contribute to segmental order confusion in VTS clusters. These findings offer a view that the common metathesis of VTSV to VSTV may be driven both by segmental order confusion due to longer sibilant noise and gestural variability as well as by perceptual optimization to bring the stop to a position in which it is less masked.

References

Bregman, Albert S and Campbell, Jeffrey. 1971. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of experimental psychology* 89:244.

Fujimura, Osamu and Macchi, M. J. and Street, L. A. 1978. Perception of Stop Consonants with Conflicting Transitional Cues: A Cross-Linguistic Study. *Language and Speech* 21:337–346.

Graff, Peter and Scontras, Gregory. 2012. Metathesis as Asymmetric Perceptual Realignment. In *WCCFL XXVIII*.

Hume, Elizabeth V. 2004. The indeterminacy/attestation model of metathesis. *Language* 80:203–237.

Jones, Kyle. 2016. The perception of stop/sibilant clusters in Modern Hebrew. In *Poster presented at LabPhon 15: Speech Dynamics and Phonological Representations*. Ithaca, NY: Cornell University.

Mielke, Jeff. 2001. Turkish /h/ deletion: evidence for the interplay of speech perception and phonology. In *Proceedings of NELS 32*.

Pitt, Mark and Dilley, Laura and Johnson, Keith and Kiesling, Scott and Raymond, William and Hume, Elizabeth and Fosler-Lussier, Eric. 2007. *Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu]*. Columbus, OH: Department of Psychology, Ohio State University.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: a perceptual account. In *The role of speech perception in phonology*, ed. Elizabeth Hume and Keith Johnson, 219–50. Brill Academic Pub.

Yanagawa, Mariko. 2003. Metathesis in Modern Hebrew: An Analysis in Articulatory Phonology. In *Proceedings of the 15th International Congress of the Phonetic Sciences*, 1671–1674. Rundle Mall: Causal Productions.

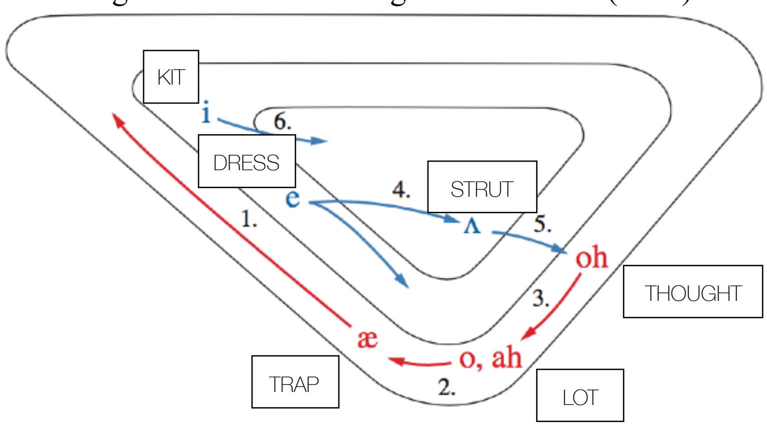
The Northern Cities Shift and Low Back Merger in 3 Cities in Northeastern Pennsylvania

Joseph Butkiewicz, *University of Pennsylvania*

In this study I examine the Northern Cities Shift (NCS) and low back merger in the three largest cities in Northeastern Pennsylvania: Scranton, Wilkes-Barre, and Hazleton. I find that these cities exhibit non-participation in the NCS and merger still in progress.

Background : Labov et al. (2006) categorize Northeastern Pennsylvania with the Inland North. Following this categorization, I expect the region to demonstrate the two defining features of the Inland North: the NCS (Figure 1, using Wells Lexical Sets, 1982) and related resistance to the low back merger (LOT-class and THOUGHT-class merger). However, Dinkin (2009) calls into question the region's participation in the NCS, and Herold (1990) reports low back merger in progress in the region. Further, several recent accounts (e.g., Wagner et al. 2016, Dinkin and Thiel 2017) indicate retraction of the NCS across in Inland North, warranting the collection of new NCS data. To address these competing accounts, I examine, in apparent time, the NCS and low back merger in Northeastern Pennsylvania.

Figure 1: NCS according to Labov et al. (2006)



Hypothesis : Given Labov et al.'s (2006) findings, I predict inconsistent NCS participation, suggesting diffusion of the shift rather than incrementation (Labov 2007). Additionally, given Herzog's Principle (Labov 1994), which states that "mergers expand at the expense of distinctions," I expect further expansion of the low back merger.

Methods : Thirty sociolinguistic interviews (Labov 1984) were conducted, each including a period of spontaneous speech and a period of formal speech (minimal pair judgements and a word list reading). Half of the 30 interviewees were men, and half women. Twelve grew up in Scranton, eleven in Wilkes-Barre, and seven in Hazleton. The interviewees were a wide range of ages: the oldest born in 1930 and the youngest in 1996. The recorded speech samples were force aligned and extracted using the FAVE-suite (Rosenfelder et al. 2014). Formant frequency output from FAVE was normalized in NORM (Kendall and Thomas 2007) using the modified Neary normalization method used by ANAE (Labov et al. 2006). Only primarily and secondarily stressed tokens were included for analysis. Following Herold's (1990) parameters, pre-rhotic tokens were excluded for analysis of low back merger.

Results (Northern Cities Shift) : None of the speakers met any of Labov's (2007) criteria for inclusion in the Inland North. Further, EQ1 index measurements (Dinkin 2009) were consistent with communities outside the Inland North (Figure 2). In apparent-time, I find DRESS class lowering and backing (consistent with NCS), but TRAP class also lowering and backing (inconsistent with NCS). For the other three possible NCS features (LOT class, STRUT class, THOUGHT class), I do not find strong apparent-time trends.

Results (Low Back Merger) : Euclidean distance in apparent time (Figure 3) suggests progression of merger throughout the region. However, speakers of all ages in Scranton continue to make low back production distinctions according to t-tests and Pillai scores. The Scranton judgment data matches the production data. In Wilkes-Barre and Hazleton, I use t-tests to find that speakers continue to make F2 distinctions in spontaneous speech, but are merged in formally elicited speech. Further, I find that their judgment data suggests perceptual merger. The Wilkes-Barre/Hazleton pattern is suggestive of near-merger, with several cases of a Bill Peter's Effect (Labov 1994) (i.e., merger in perception and formally elicited speech, but not spontaneous speech). My t-test findings by city are summarized in Figure 4.

Discussion : I find that none of the cities surveyed in Northeastern Pennsylvania participates in the NCS according to the criteria used by Labov et al. 2006. From this, I conclude that the NCS was either never fully present or was present but has retracted. I find a state of near-merger in Wilkes-Barre and Hazleton, with evidence of a community-wide Bill Peter's effect; in Scranton, speakers continue to make distinctions. These results support the following conclusions: near-merger is a legitimate path to the low back merger that can occur across a community; and, the low back merger is not necessarily expanding rapidly, even in a region where there's no evident reason for it not to spread.

Figure 2: EQ1 Index

Community	**Mean EQ1**	**Standard Deviation**	***n***
Scranton	-91	15	12
Wilkes-Barre	-93	30	11
Hazleton	-100	13	7
Telsur Inland North	+22	72	61
Telsur non-IN	-111	55	385

Figure 3: Euclidean Distance in Apparent Time

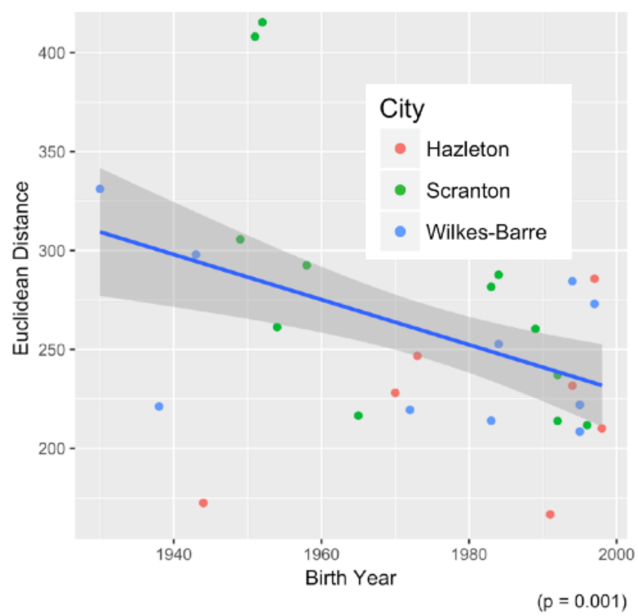
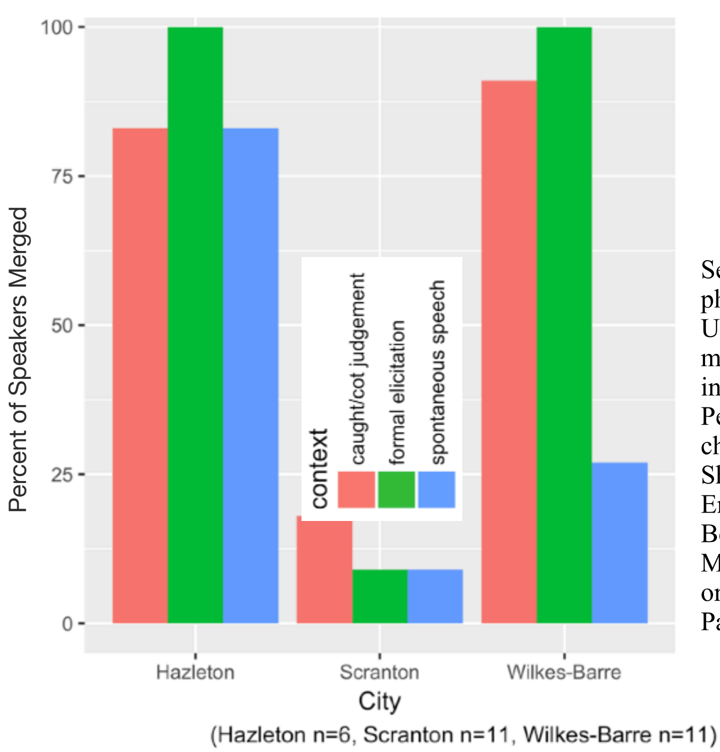


Figure 4: Percent of Speakers Merged by City



Selected References: **Dinkin**, Aaron. 2009. Dialect boundaries and phonological change in Upstate New York. PhD dissertation, University of Pennsylvania.; **Herold**, Ruth. 1990. Mechanisms of merger: The implementation and distribution of the low back merger in eastern Pennsylvania. PhD dissertation, University of Pennsylvania.; **Labov**, William. 1994. Principles of linguistic change, vol.1: Internal factors. Oxford: Blackwell.; **Labov**, William, Sharon **Ash** & Charles **Boberg**. 2006. The atlas of North American English: Phonetics, phonology, and sound change. New York and Berlin: Mouton de Gruyter.; **Wagner**, Suzanne, Alexander **Mason**, Monica **Nesbitt**, Erin **Pevan**, & Matt **Savage**. 2016. Reversal and re-organization of the Northern Cities Shift in Michigan. Penn Working Papers in Linguistics 22.2:171–179.

Production and perception of word-initial Korean stops undergoing sound change

Andrew Cheng (UC Berkeley)

Korean exhibits a three-way "laryngeal" contrast in manner of stop and affricate articulation [1] that has been undergoing a tonogenetic change in production and perception in the Seoul metropolitan area. In this sound change, lenis stops' VOT increases in word-initial position, nearly merging with the VOT of aspirated stops, and at the same time, f0 of the following vowel of aspirated stops rises, making f0 the primary cue for contrast [2,3]. Thus, average *aspirated-lenis* VOT is decreasing, while mean *aspirated-lenis* f0 is increasing. This change is manifested in both production and perception. As a change in progress, not only do younger speakers advance the change, female speakers also have been shown to lead over males [5], and speakers with L2 English experience lead over those without [6]. While the bulk of this research focuses on the sound change in peninsular Korean, only one study has examined its progress in the Korean diaspora and found evidence in Korean Canadian speakers of "VOT drift" toward the phonological patterns of English [7]. Determining the extent to which members of a diasporic language community participate in a sound change may shed light on how factors such as migration and unstable L1 input inhibit generational transmission of changes in progress (the "transition problem" [8]).

This study focuses on Korean L1, English L2 (heritage Korean) speakers' changing use of VOT and f0 to distinguish lenis and aspirated stops in both production and perception, as well as the social valuations of these acoustic variables. Thirty-six Korean sentences were designed to elicit the lenis and aspirated stops and affricates in prosodically unmarked position. They were read aloud by thirty-two speakers of Korean who had come to the United States at varying ages: adolescence (first generation immigrants), from birth to 2 years old (second generation), or at any point in between (1.5 generation). Their speech was then played back to Korean American listeners recruited through Mechanical Turk, who judged the utterances on speaker attributes such as native-like proficiency and perceived generational group.

Production results show that second generation Korean Americans do not participate in the sound change of their native South Korean same-age counterparts. A significant effect of age of immigration was found for use of f0 contrast in females (p=0.02); second generation speakers did not demonstrate the increased mean *aspirated-lenis* f0 (Fig. 1a). A significant effect of generation was found for VOT of aspirated stops for both genders (p=0.007), but not for lenis stops; however, most speakers maintained a similar, near-zero amount of VOT contrast between both stops, as no significant effects of generation were found for mean *aspirated-lenis* VOT (Fig. 1b). Thus, while second generation speakers may demonstrate the VOT merger, it is not occurring in tandem with the rise in f0 contrast.

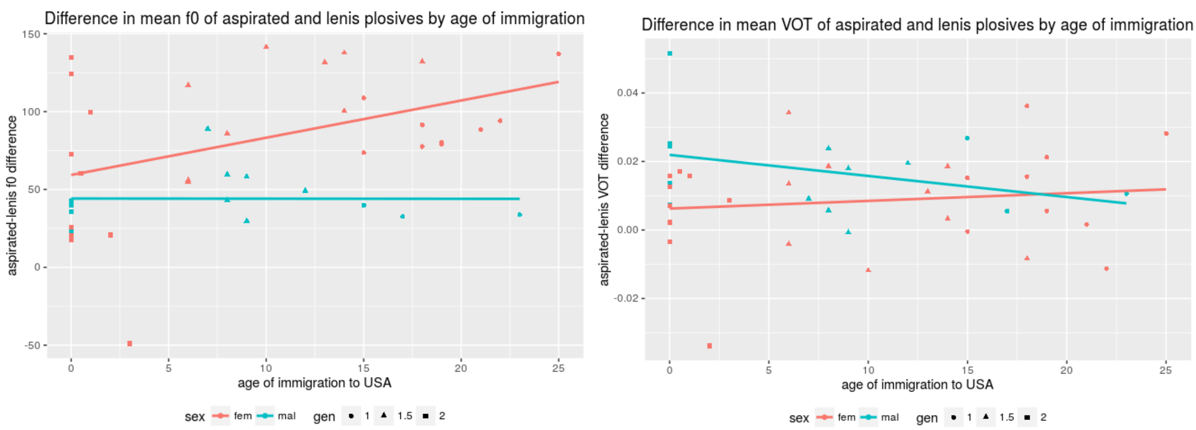


Figure 1a. *F0 difference of word-initial aspirated and lenis stops and affricates by age of immigration.* Figure 1b. *VOT difference by age of immigration.*

In the perception task, a speaker's generational group was very easily identifiable by listeners (Pearson's R=0.681, p<0.001 (Fig. 2a)). However, the correlation between a speaker's perceived age of immigration and their use of VOT and f0 to contrast lenis and aspirated stops was mild (for VOT: R=-0.439, p=0.025 (Fig. 2b); for f0: R=0.422, p=0.025).

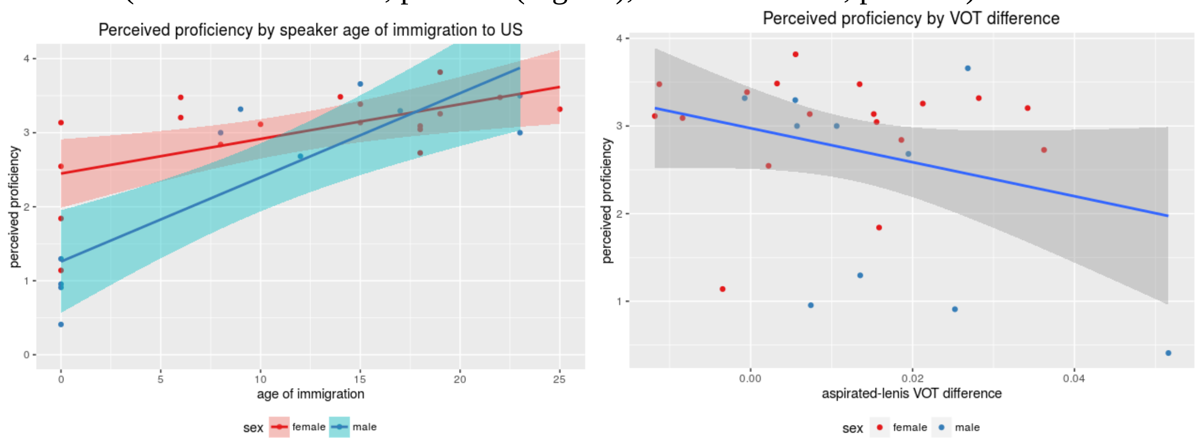


Figure 2a. *Speakers' perceived proficiency in Korean is positively correlated with their age of immigration*.
Figure 2b. *Correlation between perceived proficiency and aspirated-lenis VOT difference (greater difference indicates less merged) is mildly negative.*

A closer look at the individual differences among speakers reveals that those with the highest proficiency rankings demonstrate the tonogenetic sound change and closely resemble first generation speakers, but the lowest-ranked speakers diverge from the norm in varying ways: some speakers ranked low in proficiency still demonstrated progressive use of VOT and f0, indicating that their proficiency evaluation may depend on other speech factors.

It is concluded that while the tonogenetic sound change is robust in Seoul Korean, its progress has been impeded in the diasporic second generation community, and that these acoustic cues are not used as sociophonetic markers for native Korean identity. These findings add to the recent literature that focuses on heritage speakers [9, 10] and promotes further study of immigrant language contact situations such as this in order to elucidate how the heritage speaker acquires and maintains a bilingual phonological system.

[1] Taehong Cho, Sun-Ah Jun, and Peter Ladefoged. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30:193-228, 2002.

[2] David J. Silva. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23(2):287-308, 2006a.

[3] Hye-Young Bang, Morgan Sonderegger, Yoonjung Kang, Meghan Clayards, and Tae-Jin Yoon. The effect of word frequency on the timecourse of tonogenesis in Seoul Korean. In Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS), 2015.

[4] Midam Kim. Correlation between VOT and F0 in the Perception of Korean Stops and Affricates. In INTERSPEECH 2004. 8th International Conference on Spoken Language Processing, 2004.

[5] Eunjin Oh. Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics*, 39:59-67, 2011.

[6] Mi-Ryoung Kim. Interspeaker variation on VOT merger and shortening in Seoul Korean. In Proceedings of Meetings on Acoustics, volume 19. Acoustical Society of America, 2013.

[7] Yoonjung Kang and Naomi Nagy. VOT Merger in Heritage Korean in Toronto. *Language Variation and Change*, 28(2):249-272, 2016.

[8] Uriel Weinreich, William Labov, and Marvin Herzog. Empirical foundations for a theory of language change. In W. Lehmann and Y. Malkiel, editors, Directions for Historical Linguistics, pages 97-195. University of Texas Press, Austin, 1968.

[9] Terrence G. Wiley and Guadalupe Valdès. Editors' Introduction: Heritage Language Instruction in the United States: A Time for Renewal, 2000.

[10] Maria Polinsky and Olga Kagan. Heritage Languages: In the 'Wild' and in the Classroom. *Language and Linguistics Compass*, 1(5):368-395, 2007.

Korean vowel mergers: contrastive hierarchies and distinctive features

Joy Kwon (joy.kwon@wisc.edu)
University of Wisconsin-Madison

The present paper provides phonological representations accounting for Korean vowel changes. Especially, I suggest phonological representations employing contrastive hierarchy of distinctive features (Ko 2009; Oxford 2015). Data from Middle Korean (MK), Early Modern Korean (EModK), and Modern Korean (ModK) substantiate how phonological models can provide systematic explanations for diachronic/synchronic changes.

The representations are built upon two theoretical frameworks: the model of distinctive features (MDF) (Avery & Idsardi 2001) and the contrastive hierarchy derived from the Successive Division Algorithm (SDA) (Dresher 2009). The former model provides building blocks while the latter sketches a skeleton of segmental representations. A building block is called "dimension" in accordance with Avery & Idsardi's term. Each dimension node is marked/unmarked with a specific dimension having its terminal dependent gestures. These dependent gestures are what we know as features (e.g., [high], [round], [spread]). The realization of a dimension into a specific gesture is referred to as a completion rule. Moving on to the contrastive hierarchy, a Dresherian analysis generates tree-like structures for phonemes. Just like syntactic trees, branches are binary so a mother node bears two daughter nodes (hence, sisters). The order of the feature ranking is determined by SDA.

Based on the idea of features and hierarchy, Oxford (2015) proposes four hypotheses for a sound change model. Among those four, I introduce two relevant hypotheses in this paper: 1) sisterhood merger hypothesis (SMH), meaning structural mergers apply to contrastive sisters, and 2) contrast shift hypothesis (CSH), meaning contrastive hierarchies can change over time (p. 317).

To test Oxford's model, four cases of Korean vowel mergers are examined. Firstly, I reanalyze an example of /ʌ/ sound change in MK (to /ɨ/) and EModK (to /a/) presented in Ko (2009). He argues that /ʌ/ in MK take two distinct paths of merger. The first merger, /ʌ/→/ɨ/, happens in the 16th century (MK). Ko suggests a feature hierarchy of [coronal] > [low] > [labial] > [RTR] which I reanalyze it as Tongue Thrust (TT, completed with [front]) > Tongue Height (TH, completed with [low]) > Labial (Lab, completed with [round]) > Tongue Root (TR, completed with [RTR]). According to this representation, /ʌ/→/ɨ/ in MK are sisters sharing the same mother node. Thus, the first step of /ʌ/→/ɨ/ merger satisfies the sisterhood merger hypothesis (SMH). The second merger, /ʌ/→/a/, takes place in the mid-18th century (EModK). Evolving from MK to EModK, the order of features alters (i.e., CSH) and /ʌ/ and /a/ forms a new sisterhood under the following hierarchy: [coronal] > [low] > [high] > [lab], which I restate it as TT > TH > TR > Lab with the same completion rules (Fig. 1 & Fig. 2). The examples below illustrates the /ʌ/→/ɨ/ and /ʌ/→/a/ merger.

- 1st merger: /han**ʌ**l/ > /han**ɨ**l/ 'sky' /tar**ʌ**-/ > /tar**ɨ**-/ 'different'
- 2nd merger: /p**ʌ**ram/ > /p**a**ram/ 'wind' /t**ʌ**l/ > /t**a**l/ 'moon' (from Ko, 2009, p. 9)

In terms of synchronic change, two instances of ModK vowel mergers corroborates Oxford's model. First example is the loss of Labial dimension (completed with [round]) of front vowels. ModK consists of 10 vowels including two high front rounded vowels /y/ and /ø/. Overtime, they have been diphthongized to [wi] and [we]. For example, /ky/ 'ear' becomes [kwi] and /nø/ 'brain' realizes as [new] in ModK. Originally /y/ and /i/ are sisters and /ø/ and /e/ are sisters (Fig.

3); therefore, SHM explains why those two sound merged to /i/ and /e/. The other example is the most recent merger, the /e/~/æ/ merger (Fig. 4). /e/ and /æ/ are sisters governed by the unmarked Labial dimension (Oral Place). Although the orthography of /e/ and /æ/ remains distinctive (note that Korean graphemes are phonemic), the non-high front vowels are allophonic, i.e., [e]~[æ] and some linguists now transcribe the sound as /ɛ/. For example, /ke/ 'crab' and /kæ/ 'dog' are minimal pairs but phonetically they are pronounced as [ke]~[kæ]~[kɛ]. Again, SHM supports the merger of these two vowels.

In short, Oxford's model of sound change (2015) is applied to data from MK to ModK based on contrastive hierarchy of distinctive features. The results strongly support that Oxford's model can provide systematic accounts for the diachronic and synchronic vowel mergers of Korean vowels as well. Therefore, the present study strengthens the argument that the contrastive illustration posited on distinctive features can provide a coherent and systematic tool for analyzing sound changes.

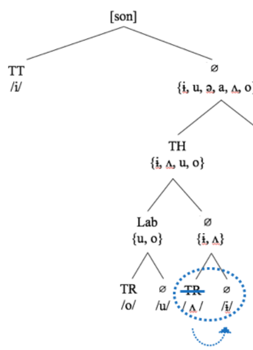


Fig. 1 MK /ʌ/→/ɨ/ merger

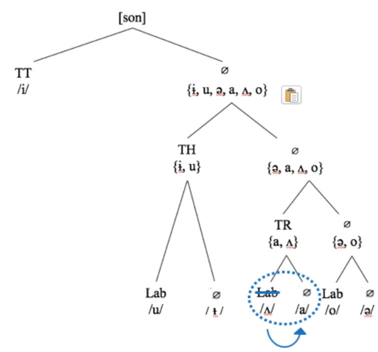


Fig. 2 EModK /ʌ/→/a/ merger

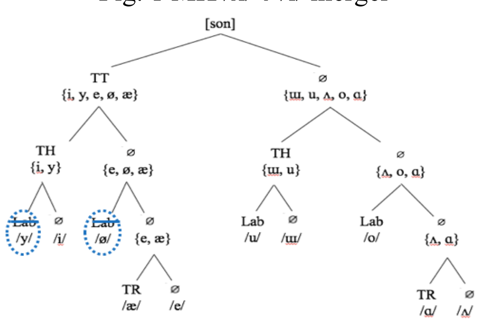


Fig. 3 ModK /y/→[wi] and /ø/→[we] mergers

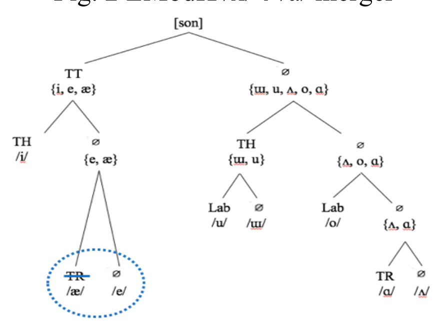


Fig. 4 ModK /e/~/æ/merger

References

Avery, P. & Idsardi, B. (2001). Laryngeal dimensions, completion and enhancement. In T. A. Hall (Ed.). *Distinctive feature theory* (pp. 41-70). Berlin, New York: Mouton de Gruyter.

Dresher, B. (2009). *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press

Ko, S. (2009). A contrastivist view on the evolution of the Korean vowel system. In H. Maezawa & A. Yokogoshi (Eds.), *MIT Working Papers in Linguistics 61: Proceedings of the 6th Workshop on Altaic Formal Linguistics,* 181-196.

Oxford, W. (2015). Patterns of contrast in phonological change: Evidence from Algonquian vowel systems. *Language, 91*(2), 308-357.

Purnell, T. & Raimy, E. (2015). Distinctive features, levels of representation, and historical phonology. In P. Honeybone and J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 522-544). Oxford: Oxford University Press.