MOT Conference Program
Programme de l'atelier MOT

Welcome to the MOT 2022 conference hosted by the University of Ottawa! Provided below is the program for the conference and the abstracts of each talk and poster. Each presenter will talk for twenty minutes, followed by a ten-minute Q&A. Click the title of the talk or poster within the program to go straight to the abstract. Click "Return to Program" at the bottom of the page to return to the program.

Bienvenues à l'édition 2022 de MOT organisée par l'Université d'Ottawa! Vous trouverez ci-dessous le programme et les résumés des présentations et affiches. Les présentations dureront vingt minutes, suivies d'une période de questions-réponses de dix minutes. Cliquez sur les titres dans le programme pour vous rendre automatiquement aux résumés. Appuyez sur « Retour au programme » au bas de chaque page pour y revenir.

Friday, March 25th - Vendredi 25 mars

10:20 to 10:30 — **Opening Remarks - Mot d'ouverture**
Suzy Ahn

10:30 to 12:00 — **Acquisition**
(Session Chair/Modératrice de séance: Yoonjung Kang)
- *Talk 1*: Child and adult unpredictability-associated phonetic enhancement (Gemma Repiso Puigdelliura)
- *Talk 2:* Achieving native-like accuracy of French/English vowels for Canadian bilingual speakers: effects of language proficiency, context of language use and daily exposure time (Jean-François Berthiaume & Daniel Pape)
- *Talk 3:* Child second language phoneme production after accented and native speech exposure **(**Margarethe McDonald et al.)

12:00 to 1:00 — **Lunch - Dîner**
- Gather Town Tutorial – Démo sur Gather Town

1:00 to 2:30 — **Phonology I – Phonologie I**
(Session Chair/Modératrice de séance: Karen Jesney)
- *Talk 1:* ROOT-first Deletion in Harmonic Serialism (Nate Shaftoe)
- *Talk 2:* Vowel Hiatus Resolution in Arabic (Samira Farwaneh)
- *Talk 3:* Turkish (so-called) 'Yumuşak g' (ğ), long vowels, and the phonology-morphology interface (Nicolas Royer-Artuso)

2:30 to 3:00 — **Break - Pause**
- Gather Town Tutorial – Démo sur Gather Town

3:00 to 4:30 — **Phonetics I – Phonétique I**
(Session Chair/Modérateur de séance: Alexei Kochetov)
- *Talk 1:* Bilingual Perception of Voicing Contrasts in English and Spanish Pseudowords (Owen Ward & Vanina Machado)

Return to Program/Retour au programme

- *Talk 2:* Phonetic interference in the production of stops by Western Armenian bilinguals (Talia Lisa Tahtadjian)
- *Talk 3:* Categorization of non-standard Spanish trills: Exploring the variability of Costa Rican Spanish (Mariana Cortes Kandler)

4:30 to 4:45 **Break - Pause**

- Gather Town tutorial – Démo sur Gather Town

4:45 to 6:30 **Posters + Social - Affiches + Activité sociale**
(Opening/Ouverture: Kevin McMullin)

- Posters will close at 5:45 to allow poster presenters to join social

Saturday, March 26th – Samedi 26 mars

10:30 to 12:00 **Phonology II - Phonologie II**
(Session Chair/Modérateur de séance: Lev Blumenfeld)

- *Talk 1:* Loan words and the Perceptual Map: the pattern of cluster repairs in Sranan (Katie Van Luven)
- *Talk 2:* The vowels of Badaga re-examined (Paul Arsenault)
- *Talk 3:* Analogy and Lexical Diffusion in Ring Grassfields Bantu languages: an account for sound change (Liliane Hodieb)

12:00 to 1:00 **Lunch - Dîner**

1:00 to 2:30 **Phonetics II – Phonétique II**
(Session Chair/Modérateur de séance: Daniel Pape)

- *Talk 1:* Pathways to Depalatalization of the French Palatal Nasal in Quebec & Hexagonal French: An EPG Study (Laura Colantoni)
- *Talk 2:* The acoustic realization of the stop voicing contrast in Argentine Spanish (Sophia Sedigh Afshar)
- *Talk 3:* The variability and reliability of listener performance in an AXB assessment of phonetic imitation (Beth MacLeod)

2:30 to 3:00 **Break - Pause**

3:00 to 4:30 **Variation and Change – Variation et changement** (Session Chair/Modérateur de séance: Morgan Sonderegger)

- *Talk 1:* Analysis of Mandarin exclamations: A Comparison of the Pronunciation of "en" in Regions (Jiaying Li)
- *Talk 2:* BAG-raising in Ontario and Colorado (Lisa Sullivan)
- *Talk 3:* Perceptions of sociophonetic variation in the pronunciation of "Toronto" (Caitlin Bergin)

4:30 to 4:45 **Break – Pause**

4:45 to 5:45 **Business meeting – Réunion de travail**
(Session Chair/Modérateur de séance: Marc Brunelle)

Return to Program/Retour au programme

Poster Presentations - Affiches:

- Accent or Expectation? Investigating the Relationship Between Accented English and Psychological Valence Perception (Mariam Bekhet et al.)
- High Vowel Syncope Failure in Urban Jordanian Arabic: A Positional Faithfulness Treatment (Aziz S. Jaber)
- A Preliminary Investigation of Vowel Changes in Eastern New England English (Sydney La Valley)
- On the acquisition of consonant clusters in Williams syndrome: A case study (Eirini Ploumidi)
- The effect of increasing acoustic and linguistic complexity on auditory processing: an EEG study (Fareeha S. Rana et al.)
- Investigating universality of consonant intrinsic F0 effects (Connie Ting)
- Cue primacy effects in Mandarin tone imitation (Wei Zhang)

Return to Program/Retour au programme

The acoustic realization of the stop voicing contrast in Argentine Spanish

Sophia Sedigh Afshar, Carleton University

Context: Consonant lenition is a synchronic and diachronic sound change in which consonants become "weaker" or more vowel-like in certain contexts, especially between vowels. In Spanish, the voiced stops /b d g/ participate in a synchronic lenition pattern in which the full stop allophones alternate with a continuant variant (e.g., [1], [3], [4]). The specific realization of this alternation varies between dialects of Spanish, with some showing extreme lenition and others showing stop allophones even intervocalically. Furthermore, there is some evidence that place of articulation of the voiced stops could affect degree of lenition ([4]). A smaller body of work has shown that, in some varieties, the voiceless stops also sometimes weaken, coming to be realized as continuants and/or voiced (e.g., [3]), that is, more like voiced stops. Given the variability in how Spanish dialects lenite the voiced stops and given that some varieties of Spanish have been shown to weaken the voiceless stops too, this raises the question of how different dialects of Spanish realize the stop voicing contrast. To date, only a handful of studies have explored lenition of both the voiced and voiceless stops together (e.g., [3]). The current study explores this issue for a particular variety of Spanish: Argentine Spanish.

Research questions:

1) How is the stop voicing contrast of Argentine Spanish realized in intervocalic position?
2) Does place of articulation influence which acoustic cues are used to produce the contrast?

Methodology: 467 Spanish words spoken aloud by Argentine Spanish speakers (12 females, 11 males, ranging in age from 19 to 58, median 31) were extracted from the Romance Phonetics Database ([2]). Target words contained voiced or voiceless intervocalic stops in the onset of an unstressed syllable.

Analysis & findings: Two acoustic measures were taken using Praat: the percentage of the stop interval containing glottal pulses (%-voicing) and intensity of the stops relative to the following vowel (relative intensity). Voiced stops are predicted to have higher %-voicing and higher relative intensity (i.e. closer to the following vowel). Furthermore, the more lenited the stop, the higher these two measures will be. A logistic mixed-effects model found that stop voicing was strongly predicted by both relative intensity ($\beta = -31.1$, $p < 0.001$) and %-voicing ($\beta = -12.3$, $p < 0.001$), with no significant interaction with place. These finding suggest 1) that the stop voicing contrast is realized by both relative intensity and %-voicing, with relative intensity being the stronger cue, at least in production, and 2) that the way the stop voicing contrast in Argentine Spanish is realized does not depend on place of articulation.

Contribution: This study contributes to our understanding of lenition processes and contrast maintenance in varieties of Spanish by illustrating how the stop voicing contrast is realized in one particular variety. Future work should explore other varieties and should determine if the acoustic cues found in production align with those used in perception.

References

1. [1] Carrasco, P., Hualde, J. I., & Simonet, M. (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican versus Iberian Spanish. *Phonetica*, 69(3), 149–179.
2. [2] Colantoni, L., & Steele, J. (2004). The University of Toronto Romance Phonetics Database. http://r1.chass.utoronto.ca/rpd/
3. [3] Lewis, A. M. (2002). Contrast Maintenance and Intervocalic Stop Lenition in Spanish and Portuguese: When is it Alright to Lenite? In Current Issues in Romance Languages: Selected Papers from the 29th Linguistic Symposium on Romance Languages (pp. 159–171).
4. [4] Ortega-Llebaria, M. (2004). Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/. Laboratory Approaches to Spanish Phonology. Berlin: Mouton de Gruyter, 237-255.

Return to Program/Retour au programme

The vowels of Badaga re-examined

Paul Arsenault, Tyndale University

Abstract

Emeneau (1939) reported that Badaga (a Dravidian language of India) had a remarkable three- way contrast between ‘normal’ (non-retroflex), ‘half-retroflexed’ and ‘fully retroflexed’ vowels. Thereafter, Badaga became a celebrated example of retroflex vowels, earning recognition in works such as Chomsky and Halle (1968) and Ladefoged and Maddieson (1996), and playing an important role in discussions of phonological typology and feature theory (e.g., Hockett 1958; Bhat 1974; Lindau 1978). Research conducted many decades after Emeneau (1939) found no trace of the retroflex vowels, suggesting that they had been lost in the interim period (Pilot- Raichoor 1988, 1991; Hockings and Pilot-Raichoor 1992). As a result, Emeneau’s account has never been adequately verified despite its typological and theoretical significance.

This paper re-examines evidence from surviving data, which confirms that Badaga had distinctive retroflex (or rhotacized) vowels with a characteristic low third formant. However, the evidence does not provide convincing support for contrasting degrees of retroflexion. The weight of evidence suggests that Badaga likely had a two-way contrast between regular and retroflex vowels, with some allophonic variation in how the later series was realized. This conclusion is supported by evidence from three areas. First, acoustic analysis of field recordings made by Peter Ladefoged and associates in the 1990’s confirms the production of retroflex (or rhotacized) vowels by some speakers but does not provide evidence of a clear or consistent distinction between ‘half-retroflexed’ and ‘fully retroflexed’ vowels (contra Ladefoged and Maddieson 1996). Second, historical-comparative evidence reveals that there is no motivation for two degrees of retroflexion, as both alleged degrees have the same historical antecedents. Finally, cross-linguistic typology indicates that no other language maintains contrast between two degrees of retroflexion in consonants, let alone in vowels. Some implications for phonological typology and feature theory are briefly discussed.

References

Bhat, D. N. S. 1974. Retroflexion and retraction. *Journal of Phonetics* 2. 233–237. Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Emeneau, M. B. 1939. The vowels of the Badaga language. *Language* 15(1). 43–47.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: The Macmillon Company.

Hockings, Paul & Christiane Pilot-Raichoor. 1992. *A Badaga-English dictionary*. Berlin & New York: Mouton de Gruyter.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world’s languages*. Oxford: Blackwell. Lindau, Mona. 1978. Vowel features. *Language* 54. 541–542.

Pilot-Raichoor, Christiane. 1988. An outline of the Badaga language. In Paul Hockings, *Counsel from the ancients: A study of Badaga proverbs, prayers, omens and curses*, 51–84. Berlin & New York: Mouton de Gruyter.

Pilot-Raichoor, Christiane. 1991. *Le Badaga: langue Dravidienne (Inde): Description et analyse*. Paris: Universite de la Sorbonne-Nouvelle. Thèse pour le doctorat.

Return to Program/Retour au programme

Accent or Expectation? Investigating the Relationship Between Accented English and Psychological Valence Perception

Mariam Bekhet, Daniel Pape & Nadia Lana, McMaster Univeristy

Whether we realize it or not, our expectations of accented speakers go beyond the actual phonetic and acoustic differences of their speech. We might associate some English accents with romance and charm, whereas other accents could be perceived as aggressive and unpleasant. Television and films in Western culture have frequently presented Arabic and Russian-accented English as antagonistic or villainous, which implicitly influences our internal categorization of anyone with these accents. Previous research (Hatzidaki et al., 2015) suggests that accents can hinder the perception of positive valence. But what happens when there's no phonetic or acoustic difference in the English sentences we perceive, but rather just the label or expectation of a certain accent is modified?

This research project aims to investigate how psychological valence and labels can influence the perception of accented and unaccented speech. This study explores how different phonetic parameters (i.e. accents) of the same emotional (positive, neutral or negative) sentence influence perception of emotion We also investigate how different 'accent' labels influence the perception of emotionality, and thus whether the accent or simply the expectation or label of an accent would influence valence perception.

We use recordings of both monolingual English speakers and bilingual Russian/English and Arabic/English heritage speakers. For the purpose of this study, the same heritage speaker provides sentences in (1) their natural unaccented English and (2) a truthful imitation of a strongly accented speaker (Arabic or Russian). This procedure was chosen to exclude all speaker-specific effects when comparing the effects of accent on listener ratings. The monolingual English speaker productions are used as filler stimuli and provide a baseline (however including speaker-specific differences). Participants are presented with mismatched recordings and labels (i.e. 'Arabic Speaker', 'Canadian Speaker', 'Russian Speaker', 'English Speaker', and 'Speaker') and then asked to rate the valence of each sentence on a scale of 1 (negative) to 9 (positive). Furthermore, participants are presented with two questionnaires to establish language experience (Marian et al., 2007) and to measure accent prejudice (Ura et al., 2015). Participants are currently being recruited through the departmental undergraduate participation system, and preliminary results will be presented at this conference.

We predict that participants will respond more negatively to accented stimuli, but responses will be exacerbated with prior experience for that specific accent for each participant. We also expect to see a relationship between accent labels and valence ratings, hypothesizing that accents frequently presented in media as villainous will be rated more negatively overall, whether that accent is acoustically present in the audio signal or, in contrast, just provided as a label

References

Hatzidaki, A., Baus, C., & Costa, A. (2015). The way you say it, the way I feel it: emotional word processing in accented speech. *Frontiers in psychology*, *6*, 351.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language, and Hearing Research.

Ura, Masako & Preston, Kathleen & Mearns, Jack. (2015). A Measure of Prejudice Against Accented English (MPAAE): Scale Development and Validation. Journal of Language and Social Psychology. 34. 10.1177/0261927X15571537.

Return to Program/Retour au programme

Perceptions of sociophonetic variation in the pronunciation of “Toronto” Background:

Caitlin Bergin, Carleton University

Listeners are sensitive to linguistic features that are associated with different social groups (e.g., [1]). These linguistic features can influence perception of speaker identity by indexing gender, ethnicity, region of origin, etc. (e.g., [2]). Lenition provides an opportunity for sociophonetic variation, which previous research has found to carry social meaning (e.g., [3]). Anecdotal evidence from popular media and the Internet suggests that the pronunciation of city names could index a variety of social characteristics. This study explores how variations in the pronunciation of *Toronto* index speaker origin.

Research Questions: This study investigates the following: i) which pronunciation(s) do Torontonians perceive to indicate in-group and out-group membership? and ii) do social characteristics of the survey respondents influence their perceptions of variations in *Toronto*?

Methodology: 2740 participants responded to a Qualtrics survey that captured participant perceptions of 7 pronunciations of *Toronto*. Participants listened to an audio clip of a variant being spoken aloud and were asked 3 perception questions: 1. Where would a person who says *Toronto* this way be from?, 2. How old would a person who says *Toronto* this way be?, and 3. Would you personally say *Toronto* this way? The seven variants ([toɹɑnto], [təɹɑnto], [təɹɑnːo], [təɹɑnːə], [tɹɑnːo], [tʃɹɑnːo], [tʃɹɑnːə]) capture lenition of the consonants and vowels and represent plausible attested variants. Variants were presented in a different random order for each participant.

Findings: According to Torontonian respondents, [tʃɹɑnːo] is the variant most likely to be produced by someone from Toronto (62%), followed by [təɹɑnːo] (49%). Torontonians also scored these variants highest when asked if they would personally use them (58% for [təɹɑnːo]; 59% for [tʃɹɑnːo]). These findings suggest that [təɹɑnːo] and [tʃɹɑnːo] signal being local to Toronto, or in-group membership. On the other hand, Torontonians selected [toɹɑnto] and [təɹɑnto] to be most likely used be someone from the United States or elsewhere in Canada (53% for [toɹɑnto] and 49% for [təɹɑnto]). Additionally, when asked if they would say [toɹɑnto] or [təɹɑnto], 89% and 84% (respectively) selected “No,” suggesting that these variants signal out-group membership. Age and gender of the respondents also influenced the findings. The variants [təɹɑnːo], [tɹɑnːo], and [tʃɹɑnːo] became less acceptable the older the participant was. On the other hand, the variants [toɹɑnto] and [təɹɑnto] became more acceptable the older the participant was. The remaining two variants ([təɹɑnːə] and [tʃɹɑnːə]) were the only two variants containing a final schwa and, while these variants were dispreferred by Torontonians generally, female participants appeared to dislike them slightly more than males.

Significance: Results demonstrate that Torontonians ascribe group membership using the pronunciation of their city’s name. While previous research has investigated group membership and sociolinguistic variation, little to no sociophonetic research has connected variation stemming from lenition to group membership. This study provides a first look at how variation stemming from lenition in the pronunciation of city names can be co-opted to signal social information, especially region of origin. This study has the potential to be replicated in other cities whose names experience similar lenition (e.g., Atlanta, Calgary, Edmonton, Sacramento, New Orleans) to investigate lenition and group membership.

References

[1] Labov, W. (1966). *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
[2] Mack, S. (2010). A sociophonetic analysis of perception of sexual orientation in Puerto Rican Spanish. *Laboratory Psychology*, (1)1, 41-63.
[3] Mansfield, J. (2015). Consonant lenition as a sociophonetic variable in Murrinh Patha (Australia). *Language Variation and Change*, 27(2), 203-225.

Return to Program/Retour au programme

Achieving native-like accuracy of French/English vowels for Canadian bilingual speakers: effects of language proficiency, context of language use and daily exposure time

Jean-François Berthiaume & Daniel Pape, McMaster University

A study from MacLeod et al. (2009) on the production of high vowels in English and French bilinguals has shown that the production of vowels of bilinguals can be close to the production of native speakers. However, a study on Spanish and Catalan vowels by Bosch & Ramon-Casas (2011) has shown that bilinguals that learn Spanish and Catalan simultaneously or had acquired Spanish earlier produced vowels with more variability and less accuracy compared to bilinguals that acquired Spanish later. Mayr et al. (2018) found similar results with Spanish-Galician bilinguals.

The current study aims to enlarge the study population to one that is more representative of the general population and expand the focus of the research to all vowels found in both English and French (for Canadian speakers), all to give an accurate picture of second language learning. Our research questions are (1) how close to native-like articulation proficiency bilinguals can achieve with varying levels of L2 proficiency, (2) how strong the effect of second language learning is on their first language and vice versa and (3) which factors influence a native-like vowel production.

To achieve our research goal, we are analyzing all vowels present in both languages in 3 different contexts (sentence-initial focus, sentence-medial focus, and sentence-final unfocused position) with six repetitions for each vowel. We recorded participants with a high-quality microphone and saved the recordings locally as an uncompressed audio format. Due to the pandemic, we could only record a smaller number of participants in a soundproof room, while we recorded others in a quiet room in the participant's house. Participants were selected based on the following criteria: self-identification of their level in both languages, their native language, the contexts of their language use (work/school, friends, family), and their daily language use (exposure). Dominant speakers of one language (English and French) are grouped into either a *weak*, *medium* or *strong* group concerning their L2 proficiency. As additional groups, we also examine a simultaneous bilingual group, i.e., speakers that have acquired both languages before the age of 5, and 2 groups of 'monolingual' speakers whose L2 knowledge is limited to their school education. At the end of this project, we aim to have ten speakers in each group. A subset of this data is currently analyzed and, preliminary results for all speaker groups will be presented at the conference.

Based on the groups, we expect the English group to show less accurate productions and more variability for the rounded front vowels as they are not present in English. We predict higher formants, especially the F2, in the weaker group compared to the stronger ones. Quebec French has lax vowels, unlike metropolitan French. Therefore, we anticipate little to no difference between the production of the French group compared to the English monolinguals. The nasal vowels are another obstacle for the English group; thus, we predict higher F3 for English speakers struggling to target the French nasal vowel qualities.

Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? Journal of Phonetics, 39(4), 514- 526. doi:10.1016/j.wocn.2011.02.001

Macleod, A. A., Stoel-Gammon, C., & Wassink, A. B. (2009). Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers. Journal of Phonetics, 37(4), 374-387. doi:10.1016/j.wocn.2009.07.001

Mayr, R., López-Bueno, L., Fernández, M. V., & Lourido, G. T. (2019). The role of early experience and continued language use in bilingual speech production: A study of Galician and Spanish mid vowels by Galician-Spanish bilinguals. Journal of Phonetics, 72, 1-16. doi:10.1016/j.wocn.2018.10.007

Return to Program/Retour au programme

Pathways to Depalatalization of the French Palatal Nasal in Quebec & Hexagonal French: An EPG Study

Laura Colantoni, Alexei Kochetov & Jeffrey Steele, University of Toronto

Palatal nasals are variably realized in Romance languages including French (Zampaulo, 2019). What makes French particularly interesting is the wide range of geographical and positional variability attested in the patterns of depalatalization. For example, whereas Northern European French (EF) varieties have been described as showing fronting or decoupling of the nasal and palatal gesture in word-initial and final positions (Hansen & Astla Østby, 2016), in Quebec French (QF), velar realizations are observed in word- final position (Picard, 1993). To date, most descriptions have been based on auditory transcriptions, with the exception of the static palatography and X-ray studies reviewed by Recasens (2013), which revealed considerable between-speaker variation in the phonetic realization of /ɲ/. Since existing articulatory data are restricted to EF and involve a small number of tokens and items, it remains to be seen whether the realization of the consonant varies by position or lexical item, both for a given speaker and between speakers, as well as the extent to which depalatalized realizations may merge with coronal/velar nasals.

As a first step towards a more systematic investigation of the phonetics of French /ɲ/, we collected electropalatographic (EPG) data from four female French speakers, two each of EF and QF. The speakers produced six to nine repetitions of six words with the target /ɲ/ in medial (*accompagna, agneau, baignade, grognement*), and final positions (*campagne, montagne*) that were compared to /n/ and /ŋ/ in similar contexts (e.g., *anneau, final, aucune, jogging*). Words were produced in isolation, in a carrier phrase or in a passage. Linguopalatal contact profiles for all nasal tokens were examined qualitatively and quantitatively using measures of the amount of linguopalatal contact (Q) and the contact centre of gravity at consonant midpoint (CoG: higher values = greater anteriority).

The two EF speakers produced a relatively anterior /ɲ/, which differed minimally from /n/ (slightly lower CoG, with little or no difference in Q). The two QF speakers, in contrast, showed variation between fronted alveolopalatal (word-medially) and backed velar realizations (word-medially and finally), with the latter differing minimally from the /ŋ/ in *jogging*.

Our results add to previous descriptions of Northern European Romance varieties (Hansen & Astla Østby 2016; Zampaulo 2019), revealing that depalatalization patterns extend to word-medial position. Moreover, there is evidence that depalatalized nasals may merge with [n+j] sequences, as has been suggested for other Romance varieties (Kochetov & Colantoni, 2011; Bongiovanni, 2021). Our findings build on previous descriptions of QF, showing that velarization resulting from depalatalization is not restricted to word-final position (where depalatalized realizations merge with velar allophones) but is also variably present in word-medial position as evidenced by the fact that one participant velarized all word- medial tokens while the other had variable velar and alveopalatal realizations in this position. We conclude by discussing how these patterns of depalatalization fit within the Romance continuum and interact with other phonological processes.

References

Bongiovanni, S. (2021). An acoustical analysis of the merger of /n/ and /ɲ/ in Buenos Aires Spanish. *Journal of the International Phonetic Association*, *51*(2), 177–201.

Hansen, A., & Asla Østby, K.S. Detey, J. Durand, B. Laks & C. Lyche (2016). Variation in the capital city of France: Paris. In (Eds.), *Varieties of spoken French* (pp. 403–416). Oxford University Press. Kochetov, A., & Colantoni, L. (2011). Coronal place contrasts in Argentine and Cuban Spanish: An electropalatographic study. *Journal of the International Phonetic Association*, *41*, 313–342.

Picard, M. (1993). A velar nasal in Quebec French? yes and no. *Revue québecoise de linguistique théorique et appliquée*, *11*(1-4), 13–19.

Recasens, D. (2013). On the articulatory classification of (alveolo)palatal consonants, *Journal of the International Phonetic Association*, *43*, 1–22.

Zampaulo, A. (2019). *Palatal sound change in the Romance languages: Diachronic and synchronic perspectives.* Oxford University Press.

Return to Program/Retour au programme

Vowel Hiatus Resolution in Arabic

Samira Farwaneh, University of Arizona

The prohibition on consonant clusters in many Arabic varieties has long been the object of examination and reexamination in the literature. In contrast, few attempts has been made to provide a comprehensive and systematic account of vowel clustering or hiatus phenomena in Arabic. This paper aims to contribute to this gap in Arabic phonology research, by reviewing several cases of vowel hiatus and resolution in modern Arabic dialects, particularly Egyptian, Levantine and Gulf dialects. Vowel hiatus resulting from stem+affix or word+word concatenation, trigger a number of resolutions documented for other languages (Bacovic (2007for Chicano Spanish, Tanner (2007) for Ciyao, Casali (1996) for Xhosa), Kabak (2006) for Turkish, and Vago (2017) for Hungarian, ranging from coalescence, glide formation, consonant epenthesis, vowel deletion, or diphthongization.

The presentation will first survey the sources of vowel hiatus and the various resolution strategies employed to resolve hiatus conflicts for the goal of establishing a typology of vowel hiatus resolution in Arabic dialects. It will further explore Correlations between hiatus strategy and other linguistic variables such as vowel quality, morphemic status and prominence, while testing the degree to which universals proposed for other languages hold true for Arabic, e.g., retention of category-initial over category-final vowel, retention of root over affix vowel, and low vowels over high vowels. Facts about vowel hiatus and strategies will then be situated within an Optimality-theoretic framework showing how the interaction of faithfulness and markedness constraints accurately and economically account for diverse hiatus situations.

The main source of vowel hiatus is morpheme concatenation. Therefore, the data collected includes vowel-final verbs, nouns, adjectives, active participles, and prepositions inflected with vowel-initial suffixes, to determine the range of strategies employed to remedy *VV sequences. Vowel sequences are blocked by two main forces: OCP, which blocks tautosyllabic vocalic moras, and Onset, an undominated markedness constraint in Arabic, banning vowel-initial syllables; thus *VV cannot be syllabified as *VV. The data confirm that vowel deletion, if applied, targets the first vowel of the sequence, confirming previous analyses. However, Since *VV involves a final stem vowel followed byan initial affixal vowel, persistent deletion of the first vowel supports the faithfulness constraint ranking Max-Affix > Max-Root, contradicting McCarthy and Prince's (1995) universally-fixed metaranking which states that root faithfulness universally dominates affix faithfulness; thus, /rama-u/ > [ramu] not *[rama] '[they threw'.If the first vowel is high, however, then glide formation is favored over vowel deletion; thus, /biki-u/ > [bikyu] *[biku] 'they cried', indicating the dominance of markedness constraint Ident-High preserving the quality of the high vowel. Glide formation is blocked if it runs afoul of *CCC militating against consonant clusters; thus, /yibki-u/> [yibku] with deletion of V1, not *[yibkyu] with glide formation.

The theoretical significance of the project lies in enhancing our understanding of the vocalic system of Arabic, of which little is published, compared to the considerable attention the consonantal system receives. It will contribute more insights into the phonological patterns of vowel+vowel interactions. Beyond the theoretical relevance, the results of the study may have far reaching implications for Psycholinguistics and Applied Linguistics-Verifying the effect of vowel sequencing and the type of ensuing rectification through a global cross-dialectal analysis of different verbal and nominal paradigms will have implications for language processing and acquisition. Moreover, I expect the results to be highly informative for the development of dialect identification as well as voice recognition software.

Return to Program/Retour au programme

Analogy and Lexical Diffusion in Ring Grassfields Bantu languages: an account for sound change

Liliane Hodieb, Institut National des Langues et Civilisations Orientales

Despite their genetic unity and geographic closeness, Ring Grassfields Bantu languages, a group of about twenty languages belonging to the Bantoid group of the Niger-Congo family and spoken in the western region of Cameroon, display considerable differences at every level of the grammar. Phonetically, their noun class system is perhaps the most striking context where such differences appear. For example, in classes 6a, 7, 8 and, 10 Proto-Ring has the vowels *ə, *i, *i, and *i, respectively. Yet, whereas in Wushi and Babanki they are all become ə, in Lamnso we find i in the four classes. Vowel reduction was claimed to be a major cause of sound change in Grassfields languages sometimes yielding to the schwa or to zero (Watters 2003). However, subsequent to vowel reduction is the principle of analogical alignment whose effect in the phonetic and phonology of Ring Grassfields languages is too crucial to be downplayed, yet has failed to be emphasized in previous research. Pozdniakov (2018) is the only author who pointed out the pervasiveness of analogy in the numeral system of Grassfields languages, unattested in other Bantoid languages. But analogy is also very productive in other grammatical contexts like the noun class system as well as in the lexicon where lexical diffusion further comes into play. In Wushi in particular, I argue that the diphthong /ʊə/, unusual in the entire Bantu area, stems from the need to distinguish between morphemes which, except for the back vowel /ʊ/ insertion, would be identical. This vowel insertion began in the grammatical context of noun classification, and through lexical diffusion which is a type of analogy (Kiparsky 1996), extended to the lexicon. This explains the pervasiveness of /ʊə/ in Wushi, unique in the language group. Thus, a series of sound changes needs to be posited as an explanation of the phonetic and phonological pattern of Ring Grassfields languages: vowel reduction > analogy > lexical diffusion. I argue following Kiparsky (1996) that the feature selected for analogy and lexical diffusion depends on the structure of the language. Hence the differences observed in Grassfields languages.

Return to Program/Retour au programme

High Vowel Syncope Failure in Urban Jordanian Arabic: A Positional Faithfulness Treatment

Aziz S. Jaber, Yarmouk University

It has been established in the literature that high vowels are more susceptible to neutralization in world languages due to their phonetic weakness which renders them less audible and less perceptual (McCarthy 2007a). Although several studies have unraveled insightful accounts and analyses of places and environments where high vowel syncope is at play in different Arabic dialects (e.g., Abu Mansour 1995; Abu-Salim 1987; Kiparsky 2003; McCarthy 2007a) they have not identified or accounted for circumstances in which high vowel syncope fails to apply. This paper, thus, aims to fill in this gap in the literature. It investigates the phonetics of high vowels and its ramification in phonology. To achieve this goal, the paper applies positional faithfulness theory to explain the uncharacteristic resisting behavior of certain prominent positions. We will show that the high vowel syncope process causes any weak high vowel to be deleted except when occurring in one of the prominent positions of the 1) stressed syllable (tableaux 1), 2) final syllable (tableaux 2), or 3) proper name subcategory (tableaux 2). However, weak positions like unstressed syllable and nonfinal syllable submit to high vowel neutralizing syncope in Jordanian Arabic. This shows that syncope targets only weak structures in weak positions.

(1) *Ranking required preserving high vowels in stressed syllables
MAX-'σ(HV) >> *Wk-HV >> MAX-I.O

/fihim+u/ ‘they understood’	MAX-'σ(HV)	*Wk-HV	MAX-I.O
a.# 'fi.hi.mu		***!W	L
b.☞'fih.mu		*	*
c. 'fhi.mu	*!W	**W	*

Tableau (1) presents a scenario in which the correct ranking suggested successfully produces the optimal candidate entertained by the grammar of UJA. Candidate (b) wins over the other two for being the only one that satisfies the undominated strong-position constraint MAX-'σ(HV), though explicitly incurring two successive violations of both markedness and general faithfulness constraints. Candidate (c) loses for breaching the top-ranked MAX-'σ(HV), while candidate (a) incurs three violations to the second top-ranked *Wk-HV, and hence is excluded.

(2) *Ranking required preserving high vowels in final syllables
MAX-'σ(HV), MAX-σ-F(HV) >> *Wk-HV >> MAX-I.O

/siliku/ ‘his wire’	MAX-'σ(HV) MAX-σ-F(HV)	*Wk-HV	MAX-I.O
a.#'si.li.ku		***!W	L
b.☞'sil.ku		*	*
c. 'si.lik	*!W	*	*
d.'sli.ku	*!W	**W	*

The competition among candidates to win is decided in favor of (b) for observing the two undominated positional faithfulness constraints MAX-'σ(HV) and MAX-σ-F(HV), though it penalizes one low ranked markedness constraint, *Wk-HV, and another bottom-ranked faithfulness constraint, MAX-I.O. Candidates (c & d), however, are excluded for incurring fatal violations of the undominated MAX-σ-F(HV)and MAX-'σ(HV), respectively. Candidate (a) could have won, hadn’t it incurred three violations of the third top-ranked *Wk-HV, which is active and operative when it comes to a

Return to Program/Retour au programme

weak high vowel in a less-privileged weak syllable like the medial. (3) Ranking for high vowel syncope in proper name subcategory

MAX-PN(HV)>> *Wk-HV >> MAX-I.O

/ riha:b /	MAX-PN(HV)	*Wk-HV	MAX-I.O
a. ☞ # ri.'ha:b	W	*	W
b. ' rha:b	*!	L	*

Incurring one violation of the second top-ranked constraint *Wk-HV notwithstanding, candidate (a) wins out for observing the top-ranked faithfulness MAX-PN(HV). Candidate (b), however, shows compliance to *Wk-HV but is excluded for penalizing the top-ranked MAX-PN(HV).

Return to Program/Retour au programme

Categorization of non-standard Spanish trills: Exploring the variability of Costa Rican Spanish

Mariana Cortes Kandler, University of Ottawa

Trills are articulatorily and aerodynamically complex. This can lead to a myriad of non-standard realizations that are interesting from a phonetic, phonological, and sociophonetic point of view. This study aims i) to categorize the Costa Rican Spanish trill realizations acoustically, ii) to identify the number of variants, and iii) to determine whether they are acoustically distinguishable. Costa Rican trills typically lack vibration of the tongue tip, and are thus categorized as 'non-standard, compared to the 'canonical' or 'normative' Spanish rolled alveolar trill (Chavarría Aguilar, 1950). It has been called 'assibilated'; however, a review of the literature finds no consensus regarding the precise articulation, the acoustic features, or the phonetic symbol that should be used to represent these sounds. Some studies only distinguish between the standard and an assibilated trill (e.g., Vásquez Carranza, 2006), while others make finer-grained distinctions (e.g., the nine allophonic variants identified by Calvo Shadid, 1995). Only impressionistic descriptions have been provided, which poses challenges to cross-study and cross-dialect comparisons.

To examine the variability of trills in Costa Rican Spanish, 18 speakers (9 female, mean age=37.4) participated in a production task. Seventy-two tokens containing the trill were examined in: (i) word-initial/medial position, (ii) stressed/unstressed syllables, and (iii) before and/or after the five Spanish vowels. The study was conducted remotely using Gorilla Experiment Builder (gorilla.sc). Participants recorded their productions through Audacity (audacityteam.org). Praat was used to measure duration, percentage of voicing, center of gravity (fricatives), number of occlusions (trills), and consonant-to-vowel intensity ratio. Tokens were classified into allophones based on visual inspection of the spectrograms and the waveform, using the audio as a guide. Independent samples *t*-tests and ANOVAs were used to examine the distribution of the variants and evaluate the effect of position and stress on the acoustic measurements.

The corpus consisted of 3500 tokens. Ten categories are described and characterized acoustically following Sebregts (2015). The fricative rhotic is the most common allophone, followed by the canonical trill, which was produced significantly more by female speakers than male speakers (t=4.132, $p_{holm}=0.005$). Six of the variants were very rare, appearing in less than 3% of cases. Word stress showed the expected effect of making trills and approximants longer, but fricative variants were shorter in this condition. Word position also played a role: approximants and trills showed a statistically non-significant tendency to be shorter word-initially, but fricatives had the opposite direction of effect and were significantly longer in this position (t=0.002, $p<0.001$); also, fricatives and trills had a significantly higher percentage of devoicing word-initially (t=20.206, $p<0.001$, and t=6.789, $p<0.001$, respectively). In addition, fricatives showed a wide range of variation in center of gravity, (M=2159.2 Hz, SD=705.506, Range=5622.3 Hz). Despite the variation in acoustic features, in general, neighbouring vowels presented considerable coarticulation auditorily and visually (enhancement of F2 through a raising of F1 and a lowering of F3, in line with Heselwood & Plug, 2011), suggesting that, no matter what the trill itself sounds like, rhoticity is ensured.

This acoustic description of the non-standard variants offers insight into the phonological system of Costa Rican Spanish and serves as a parameter for future comparisons with trill variants in other dialects. While no single acoustic feature describes all variants, relationships can still be established in a system of family resemblance (Lindau, 1985). Based on these features, future research could test why such variation is possible in this dialect, but may not be allowed in others.

Calvo Shadid, A. (1995). Variación fonética de /ɾ/ y /r/ en el habla culta de San José. Revista de Filología y Lingüística de la Universidad de Costa Rica, 21(1), 115–134.

Chavarría Aguilar., O. L. (1951). The phonemes of Costa Rican Spanish. Language, 27(3), 248–253. Heselwood, B. & Plug, L. (2011). The role of F2 and F3 in the perception of rhoticity: Evidence from listening experiments. In Proceedings of the 17th International Congress of Phonetic Sciences, ICPhS (pp. 867–870). Hong Kong.

Lindau, M. (1985). The story of r. The Journal of the Acoustical Society of America, 67(S1), S27–S27. Sebregts, K. The sociophonetics and phonology of Dutch r. PhD thesis. Ultrecht University.

Vásquez Carranza, L. M. (2006). On the phonetic realization and distribution of Costa Rican rhotics. Revista de Filología y Lingüística de la Universidad de Costa Rica, 32(2), 291.

Return to Program/Retour au programme

A Preliminary Investigation of Vowel Changes in Eastern New England English

Sydney La Valley, University of Ottawa

Eastern New England English (ENEE) is a regional dialect of the United States, which appears to have several ongoing sound changes in progress. Some researchers have referred to this as "receding" dialect features or "dialect leveling" (Johnson, 2007; Stanford et al., 2012; Stanford et al., 2014; Wood, 2011). In 2014, Stanford et al. found that birth year was a significant factor in the presence of dialect features such as postvocalic ɹ-lessness, intrusive ɹ, and *Mary/marry/merry* vowel distinctions; interestingly, their results also suggest that distinct vowel productions in pairs such as *hoarse/horse* are relatively stable among generations. The present pilot study conducted preliminary research of vowel properties from two young adult speakers in Lee, New Hampshire to investigate the status of ENEE vowel changes. F1 and F2 measurements were collected from wordlist-elicited tokens to map the vowel space of speakers; current analysis is comparing these to acoustic data extracted from Dartmouth New England English Database (DNEED) (Stanford, 2020).

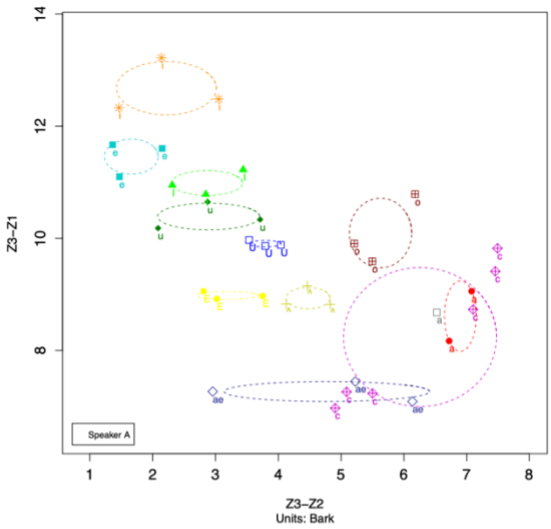


Figure 1: Bark difference metric vowel space for male speaker (A)

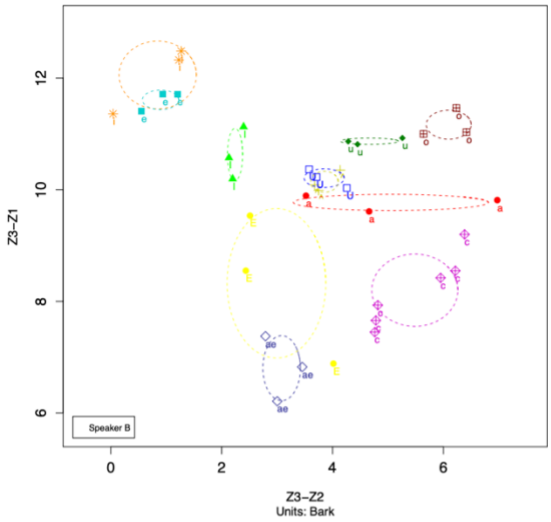


Figure 2 : Bark difference metric vowel space for female speaker (B)

Dependent Variable	Word	F1 p-value	F2 p-value	F3 P-value
speaker	Hoarse	0.005	0.046	0.037
	Horse	0.06	0.276	0.279
	caught	0.008	0.008	0.015
	cot	0.005	<.001	0.095
	Mary	0.416	0.275	0.879
	Merry	0.383	0.023	0.021
	Marry	0.269	0.013	0.97

Table 1 : T-test results comparing formant frequencies by speaker

References

Johnson, D. E. (2007). Stability and change along a dialect boundary: the low vowels of southeastern New England.

Nagy, N. (2001). "Live free or die" as a linguistic principle. *American Speech, 76*(1), 30-41. Nagy, N., & Roberts, J. (2004). New England Phonology. North American mergers in progress. (2005). In W. Labov, S. Ash, & B. Charles, *The atlas of North American English : phonetics, phonology, and sound change : a multimedia reference tool* (pp. 58-73).

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), 126-169.

Stanford, J. N., Severence, N. A., & Baclawski, Jr., K. P. (2014). Multiple vectors of unidirectional dialect change in eastern New England. *Language Variation and Change, 26*, 103-140.

Stanford, J. N. A Modern Update on New England Dialectology: Introducing the Dartmouth New England English Database (DNEED). *American Speech*, 2020. doi: https://doi.org/10.1215/000312838662137

Steindel Burdin, R. (2019). Religion and sound change in eastern New England English. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Ed.), *Proceedings of the 19th International Congress of Phoetic Sciences*, (pp. 775-779). Melbourne, Australia.

Thomas, Erik R. and Kendall, Tyler. 2007. NORM: The vowel normalization and plotting suite. [Online Resource: http://ncslaap.lib.ncsu.edu/tools/norm/]

Wood, J. (2011). Short-a in northern New England. *Journal of English linguistics, 39*(2), 135-165.

Return to Program/Retour au programme

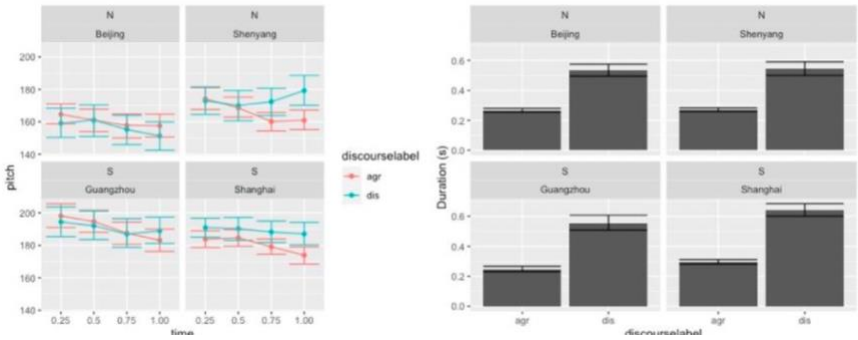
Analysis of Mandarin exclamations: A Comparison of the Pronunciation of “en” in Regions

Jiaying Li, York University

This presentation discusses a quantitative research on a Mandarin phonetic difference resulting in a pragmatic meaning change. The topic focuses on a suprasegmental change (pitch and particle duration) with the meaning of agreement and disagreement associated with the Mandarin discourse particle *en*. From Tang (2016) and Oralova (2016), *en* particle is pronounced with the short duration and a falling tone when it is used with the agreement meaning. In contrast, it is pronounced with the long duration (longer than 350ms) and a raising tone when it is with the disagreement meaning. However, the informal observation suggests that northern people in China tend to use the duration difference only, while people from the south rely on the pitch contrast. Thus, there is a regional difference on the pronunciation of *en* particle. This talk will present a study on the representative cities of each region: Guangzhou and Shanghai are for the south; Beijing and Shenyang are for the north.

Sociolinguistic interviews were conducted to collect data from 40 Mandarin speakers. In terms of the interviewees, there were four groups in regards to the cities. Five of the speakers in each group were females and the other half of them were males. Tag questions and questions containing the specific particle *ba* were used to create interview questions. The duration, maximum pitch, mean pitch and the tone shape for all tokens of *en* were measured and analyzed by using Praat.

The results of this study suggest that the pronunciation difference might not be due to the regional difference (north vs. south), whereas the difference among cities is more obvious. In terms of the particle duration, the results of speakers from four cities are statistically significant; thus, they rely on the duration contrast. In terms of the tone shape, speakers from Shanghai and Shenyang used the level tone to indicate the disagreement; for the meaning of agreement, they used the falling tone. On the other hand, speakers from Guangzhou and Beijing only used the duration contrast. They produced the short particle to indicate the agreement, while they only lengthened the particle to indicate the disagreement. Their tone shapes were unchanged. Figures 1 and 2 show the result of the tone shape in four cities.



Figures 1 and 2: Results of four cities on the tone shape and duration.

References

Oralova, G. (2016). *Minimal Response Token en in Mandarin Conversation*. MA Thesis on University of Alberta. 1-111.
Tang, H. (2016). 谈汉语口语中“嗯“的多义性. *Journal of Language and Literature*, 69-71.

[Talking about the diverse meaning of “en” in spoken Mandarin]

Return to Program/Retour au programme

Loan words and the Perceptual Map: the pattern of cluster repairs in Sranan

Katie Van Luven, Carelton University

Loan words often present speakers with structures that don't exist in their native language. When a loan word contains a structure not permitted in a speaker's L1 there are in principle a variety of strategies available to repair it. For example, a loan containing an illegal consonant cluster (C_1C_2V) could be adapted via deletion of the first consonant (C_2V), deletion of the second consonant (C_1V) or insertion of a vowel (C_1**V**C_2V). Loan adaptations are naturally accounted for in constraints-based models like OT, where the relevant constraints are already part of the grammar and the principle of Richness of the Base requires any input to be handled properly. However, because the constraints are separated from actions that repair their violations, and because all significant generations are on output structures, characterizing the adaptation process in this kind of model will encounter the "too many solutions" problem. Steriade's (2001) P-Map hypothesis is one response to this problem, and it is particularly suitable in the domain of loan words as previous work has already shown that perceptual similarity plays a role in adaptions (Kenstowicz, 2007; others). Just like how some repairs are preferred over others in the derivation of native words, with loan words there exist cross-linguistically preferred adaptations, where the P-Map provides a principled way to distinguish which repair strategy is optimal: the difference in perceptibility between the target and repair are compared, and the repair that leads to the smallest perceptual change is the one chosen (Steriade, 2001).

This paper investigates the role of perceptual similarity in the context of loanword adaptions in Sranan, an English-based creole spoken in Suriname. Interestingly the repair pattern of [s]+stop clusters in Sranan has be suggested to be a problem for the P-Map hypothesis. I contrast the repair pattern in Sranan with Cantonese, where there has been a well-establish P-Map analysis of [s]+stop repairs, suggesting that the generalization regarding the perceptual salience of [s] demonstrated in Cantonese plays a role in the Sranan adaption pattern as well. The strategies used to repair [s]+stop consonant clusters in Sranan are summarized below (examples due to Alber & Plag, 1999):

1. a) **word-finally:** ...VC_1C_2# → ...VC_1# ex. *fir**st*** → [fosi]

b) **word-medially:** ... VC_1C_2V → ...VC_1V... ex. *si**st**er* → [sisa]

... VC_1C_2V → ...VC_2V... ex. *goo**dm**orrow* → [kumara]

c) **word-initially:** #C_1C_2V... → #C_2V... ex. ***st**ory* → [tori]

In Sranan, word-initial [s]+stop clusters are repaired via deletion of [s], whereas other all cluster are imported without adaptation. Word-finally, all consonant clusters are repaired using both deletion and epenthesis, where the second member is deleted and a final vowel is inserted. Two strategies are available for word-medial clusters: sometimes deletion targets the first member of the cluster and other times the second. In cases when deletion targets the second member, the surviving consonant is always [s]. In particular, it has been noted that this pattern of [s]-deletion in (1c) is potentially problematic for accounts of loanword adaptation that appeal to perceptual similarity (Fleischhacker, 2004; others). The P-map theory predicts that a segment with more salient perceptual cues is more likely to be preserved during adaptation than one with less salient cues, and acoustic studies show that [s] is one of the most salient consonants with the strongest cues (Wright, 2004). Moreover, [s] is often preferentially preserved in adaptation in other languages such as Cantonese (Steriade 2001; others). However, in Sranan it is the [s] that is targeted for deletion and the stop that is retained. I suggest that instead of viewing (1c) as the odd one out in that it seems to contradict the perceptual similarity account available for (1a-b), it is in fact (1a) and (1c) that form a group to the exclusion of (1b). Word-initially and word-finally I follow Alber and Plag (1999) in assuming that high ranking contiguity constraints NOINTRUDE and NOSKIP (McCarthy & Prince, 1995) are the ones that determine the adaptation pattern, whereas considerations of perceptual similarity only get to exert their preference word-medially (1b) in the form of context-free faithfulness constraints MAX-S >> MAX-R >> MAX-T (Steriade, 2001), since it is precisely in that environment that contiguity constraints alone underdetermine the final outcome.

Bibliography: • Alber, B & Plag, I. 1999. Epenthesis, deletion and the emergence of the optimal syllable in creole. Rutgers Optimality Archive 335, <http://ruccs.rutgers.edu/roa.html>. • **Fleischhacker, H**. 2000. The location of epenthetic vowels with respect to consonant clusters: an auditory similarity account. UCLA: M.A. thesis. • **Kenstowicz, M**. 2007. Salience and similarity in loanword adaptation: A case study from Fijian. *Language Sciences 29*. 316–340. • **Wright, R**. 2004. A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner & D. Steriade (eds.) Phonetically based phonology, 34–57. Cambridge: Cambridge University Press. • **Steriade, D**. 2001. Directional asymmetries in place assimilation: A perceptual account. In E. Hume & K. Johnson (eds.) The role of speech perception in phonology, 219–250. San Diego: Academic Press

Return to Program/Retour au programme

The variability and reliability of listener performance in an AXB assessment of phonetic imitation

Beth MacLeod, Carleton University

Context & problem: Phonetic imitation occurs when, during an interaction, a speaker's pronunciation shifts to become more like that of the person to whom they are speaking. This phenomenon is commonly assessed using the AXB perception task (e.g. [2]), which relies on the judgements of listeners. Despite its popularity, very few studies using the AXB assessment of imitation have considered variation or reliability in the listeners' performance. [2] suggests that there is variation in the "propensity" of listeners to perceive imitation, but in that study, listeners did not all hear the same experimental trials. The current study applies a test-retest design (e.g. [1]), where listeners all complete the same experiment twice, to explore the variability and reliability of listener performance in an AXB assessment of phonetic imitation.

Methodology: To create the materials for the AXB task, four participants provided a baseline by reading aloud 20 English words, then shadowed a pre-recorded model talker saying the same words. All talkers were female, first-language (L1) speakers of Canadian English. The AXB task was run in OpenSesame online via the Remote Control feature in Zoom. On each trial, participants heard a word repeated three times: AXB. X was always the model talker recording, while A and B were the baseline and shadowed productions (counterbalanced) by one of the shadowers. The listeners' task was to choose which of A or B was pronounced more like X. The proportion of trials where they chose the shadowed token reflects the strength of the evidence of the shadower having imitated the model. Thirty listeners (all L1 Canadian English) participated in the AXB task, completing four blocks of 80 trials, for a total of 320 trials each. All listeners participated in the same experiment, hearing the same trials in the same order. Next, the same participants completed the same experiment a second time, 2-3 weeks after the first session.

Findings: Preliminary results from 16 listeners find that the percentage of shadowed tokens chosen in the AXB task ranged from 41% to 89% in the first session and 44% to 90% in the second session. These percentages were very strongly correlated ($r = 0.96$, $p < 0.001$), showing that the listeners performed similarly across sessions. Furthermore, a linear mixed-effects model found that the trial-by-trial choices in the first session were a strong predictor of the choices in the second ($\beta = 0.997$, $p < 0.001$). The by-listener slopes for the effect of the first session on the second ranged from 0.82 to 1.11, showing that each listener had a similar relation between sessions. Lastly, imitation studies often test if the group-level percentage of shadowed tokens chosen is significantly higher than 0. In this study, using the listeners' first and second session behaviour both yielded a finding of imitation at the group level (Session 1: intercept = 0.55, $p < 0.05$; Session 2: intercept = 0.67, $p < 0.01$).

Significance & contribution: This study is the first to explicitly focus on listener performance in an AXB test of phonetic imitation. The results suggest that individuals differ in their ability to perceive imitation in this task, but that this variation reflects stable characteristics of the individual listener, rather than random fluctuations. The findings indicate that AXB is a reliable way to assess imitation, as long as enough listeners are tested to include a range of perceptual abilities, supporting the suggestion of [3].

References

[1] Amitay, S., Irwin, A., Hawkey, D. J. C., Cowan, J. A., & Moore, D. R. (2006). A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naive listeners. *The Journal of the Acoustical Society of America*, *119*(3), 1616–1625.

[2] Babel, M., & Bulatov, D. (2012). The role of fundamental frequency in phonetic accommodation. *Language and Speech*, 55(2), 231–248.

[3] Pardo, J. S. (2013). Measuring phonetic convergence in speech production. *Frontiers in Psychology*, 4(AUG), 1–5.

Return to Program/Retour au programme

Child second language phoneme production after accented and native speech exposure

Margarethe McDonald (1), Eon-Suk Ko (2), and Margarita Kaushanskaya (3)
(1) University of Ottawa, (2) Chosun University, (3) University of Wisconsin-Madison

Children acquiring a second language around the world do so in a variety of different ways. Some receive more exposure to native speakers of the language they are acquiring and some receive more exposure to accented speakers. At a perceptual level, there are some indications that adults in the initial stages of language learning comprehend accented speech better than native speech (Hayes-Harb, Smith, Bent & Bradlow, 2008; Pinet, Iverson, & Huckvale, 2011). In addition, there are indications that children benefit from accented speech exposure when differentiating difficult contrasts, especially at the initial stages of language acquisition (XXX, in prep). This study aimed to examine how lab-based exposure to native and accented speech affected the *production* of easy and difficult L2 contrasts in children.

Fifty-eight Korean-native speaking children between ages 6-9 living in Kwangju, South Korea who were acquiring English as a second language were included in the final sample of the study. The 'difficult' phonemes /f/ and /v/ were targeted because they do not exist in Korean and are therefore often produced more similarly to the 'easy' phonemes /p/ and /b/ by native Korean speakers in the initial stages of English acquisition. In one block, children were exposed to alternating /f/- and /p/-initial words presented by either a native North American English speaker or by a Korean-accented English speaker. After a set of 4 alternating words, children were prompted to produce the minimal pair fan-pan through images representing the two objects. The sequence of exposure to 4 words and production of the minimal pair was repeated 6 times for each exposure condition. In another block, children followed the same procedure with exposure to /v/- and /b/-initial words and the production of the minimal pair vee-bee. Each child's productions were recorded and later presented to native English speakers. Twenty-two native English speakers judged the intelligibility of each production by denoting if they heard the word fan, pan, vee, or bee. In addition to the experimental task, children also completed a battery of tasks including an English articulation screener and the Peabody Picture Vocabulary Test (PPVT-4, Dunn & Dunn, 2007) to index English proficiency.

Results indicated a significant interaction between phoneme type (easy, difficult) and exposure type (native, accented) ($\chi 2 = 48.77$, $p < .001$). The interaction was such that for the easy phonemes /p/ and /b/, children were more significantly intelligible after exposure to Korean-accented English than after exposure to native English ($OR = 1.433$, $p < .001$). On the other hand, for the difficult phonemes /f/ and /v/, children were significantly more intelligible after exposure to native English than exposure to Korean-accented English ($OR = 0.79$, $p = .02$). English proficiency was also examined. Children with higher English proficiency, as measured by the PPVT-4 were more intelligible for all contrats, and this effect did not significantly interact with exposure type for either phoneme type. Results indicate that accented speech may be used as a tool for improving speech production in the short-term; however, the usefulness of the tool depends on the interaction between the phonological inventory of the two languages.

References

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of phonetics, 36*(4), 664-679.

XXX (in prep). Perceptual shift in children after short-term exposure to accented and native speech.

Pinet, M.,

Iverson, P., & Huckvale, M. (2011). Second-language experience and speech-in-noise recognition: Effects of talker–listener accent similarity. *The Journal of the Acoustical Society of America, 130*(3), 1653-1662.

Return to Program/Retour au programme

On the acquisition of consonant clusters in Williams syndrome: A case study

Eirini Ploumidi, University of Crete

The present study investigates the acquisition of consonant clusters of rising, plateau and of reversed sonority in the speech of a monolingual atypically developing Greek-acquiring girl (age: 9) diagnosed with Williams syndrome. The data, drawn from Andriola (2015), are cross-sectional, obtained from a picture-naming task. Previous studies show that children with Williams syndrome avoid the realization of clusters using several simplifications strategies, e.g., cluster reduction (e.g., Velleman and Mervis 2011; Garber 2018). In this study, we show that distinct patterns of (un)faithful realizations are attested depending on the target cluster and its position within the word. The realization patterns are as follows:

i) The *rising sonority* clusters are not faithfully realized. Rather, they are reduced. Two reduction patterns are attested.
PATTERN 1: The [OBSTRUENT + LIQUID] clusters are reduced to the *less* sonorous cluster member, i.e., the sonority pattern is attested (1a-b).
PATTERN 2: The [OBSTRUENT + NASAL] clusters are reduced to the *more* sonorous cluster member, i.e., the consonant which is adjacent to the syllable nucleus is realized. Thus, contiguity is obtained (1c-d).
ii) The *sonority plateau* clusters are not faithfully produced. Thus, they undergo reduction.
PATTERN 1: The sonority plateau clusters are always reduced to the rightmost member of the cluster, i.e., the cluster member which is adjacent to the vowel of the syllable is realized. Hence, contiguity-driven reductions occur (1e-f).
iii) The clusters of *reversed sonority* are realized in salient positions reflecting a POSITIONAL FAITHFULNESS effect whereas they undergo reduction in non-salient ones (for salient positions see e.g., Beckman 1998; Smith 2000). The attested patterns are as follows:
PATTERN1: In prominent positions, i.e.,instressedandinitialsyllables,bothclustermembersarepreserved in the output (1g-h).
PATTERN 2: In non-prominent positions, i.e., in unstressed and non-initial positions, cluster reduction occurs. In this case, the consonant which surfaces is the one which is adjacent to the vowel (1i-j).
Taking the findings together, we claim that the consonant clusters have not been acquired since they undergo reduction systematically. The only exception is the clusters of reversed sonority which in prominent positions are faithfully realized. All in all, it is argued that sonority, contiguity as well as positional faithfulness effects govern the child's (un)faithful realizations.

DATA

(1)	Cluster	Adult form	Child form	Gloss
a.	Rising sonority	**fli**.ˈdzani	**fi**.ˈdzani	cup
b.		xo.ˈ**dri**	ko.ˈ**di**	fat
c.		pe.ˈ**xni**.ðja	pe.ˈ**ni**.ðja	toys
d.		ˈci.**kn**os	ˈci.**n**os	swan
e.	Plateau sonority	ˈli.**mn**i	ˈli.**n**i	lake
f.		a.ˈ**vɣ**o	a.ˈ**ɣ**o	egg
g.	Reversed sonority	er.ɣo.ˈ**st**a.si.o	el.ɣo.ˈ**st**a.sio	factory
h.		a.ˈ**ft**o	a.ˈ**ft**o	this
i.		ˈfa.da.**zm**a	ˈfa.da.**m**a	ghost
j.		ˈna.**ft**is	ˈna.**t**is	able seaman

REFERENCES

Andriola, Vasiliki. 2015. *"The linguistic characteristics of Williams syndrome. A case study"*. Master thesis. Technological Educational Institute of Western Greece • Beckman, Jill. 1998. *"Positional Faithfulness"*. PhD diss. Amherst, GLSA. • Garber, Claudia. 2018. *"Early Phonological Systematization in Children with Williams Syndrome: A Longitudinal Study"*. Master thesis. University of Vermont. • Velleman, Shelley L, and Carolyn Mervis. 2011. "Children with 7q11.23 Duplication Syndrome: Speech, Language, Cognitive, and Behavioral Characteristics and their Implications for Intervention." *Perspectives on language learning and education* 18 (3): 108-116. doi:10.1044/lle18.3.108 • Smith, Jennifer. (2000). "Prominence, augmentation, and neutralization in phonology". In *Annual Meeting of the Berkeley Linguistics Society*, 26 (1): 247-257.

Return to Program/Retour au programme

Child and adult unpredictability-associated phonetic enhancement

Gemma Repiso Puigdelliura, McMaster University

Predictability is known to be in an inverse relationship with phonetic enhancement, less predictable words are more likely to be produced with more phonetic prominence. Predictability is conceptualized here as having a listener-oriented dimension by which speakers guide phonetic effort in considering the listener's needs (Turnbull & Clopper, 2015). Under this account, the perception of the listener's needs is paramount to employ strategies of phonetic enhancement, suggesting a relationship between theory of mind and predictability-based phonetic enhancement. In a study on phonetic reduction, however, Turnbull (2019) did not observe the expected relationship between semantic predictability-induced phonetic reduction and individual variation in theory of mind. In this project, we seek to address this question by comparing groups of children (i.e., developing theory of mind) and adults (i.e., mature theory of mind) in Spanish cross-word interactions. In consonant-to-vowel word junctures (i.e., C#V/), glottal marking enhances the prosodic boundaries of the vowel-initial word. Although vowel-initial glottalization is more present in prominence-driven languages, Spanish speakers have also been found to present vowel-initial glottalization (Garellek, 2014). We predict that high predictability words will be glottalized less often than low predictability words, and that child Spanish speakers will glottalize less often than adult Spanish speakers, as they will be less likely to apply unpredictability-based phonetic enhancement.

Twenty adult Mexican Spanish speakers (14F, 6M, mean age = 20.87) and 44 child Mexican Spanish speakers (19F, age range = 5;1 to 11;8 years, mean age = 8;6 years) participated in two production tasks eliciting sequences of function words + real words (i.e., 'el árbol' *the tree* high predictability) and function words + novel words (i.e., 'el anbo' low predictability). Tokens were coded as presenting glottal phonation (i.e., creaky phonation or complete glottal stop) or modal phonation at the word juncture. We submitted the 975 resulting tokens to a generalized logistic regression was with the variables age (i.e., younger child, older child, adult), stress (i.e., stressed, unstressed), and type of word (i.e., real words, novel word) as fixed effects. Our findings showed that real words were less likely to be glottalized (M = 1.14%, SE = 0.46) than novel words (M = 21.60%, SE = 1.94) (β = -3.80, z = -7.97, $p < 0.001$). Adult speakers were more likely to glottalize (M = 13.88%, SE = 1.92) than younger child Spanish speakers (M = 5.80%, SE = 1.37) (β = 1.62, z = 2.27, p = 0.02). Younger child Spanish speakers did not significantly differ from older child Spanish speakers (M = 11.85%, SE = 1.68) (β = 0.88, z = 1.6, p = 0.11).

Our results demonstrate that word unpredictability is associated with phonetic enhancement in cross-word junctures. Our findings also suggest that children are less likely to be attuned to unpredictability-associated phonetic enhancement than adults, supporting a relationship between developing theory of mind and predictability-based phonetic prominence.

References

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, *45*(1), 106–113. https://doi.org/10.1016/j.wocn.2014.04.001

Turnbull, R. J. (2019). Listener-oriented phonetic reduction and theory of mind. *https://doi.org/10.1080/23273798.2019.1579349*, *34*(6), 747–768. https : / / doi . org / 10.1080/23273798.2019.1579349

Turnbull, R. J., & Clopper, C. G. (2015). *Assessing the listener-oriented account of predictability-based phonetic reduction* (Doctoral dissertation).

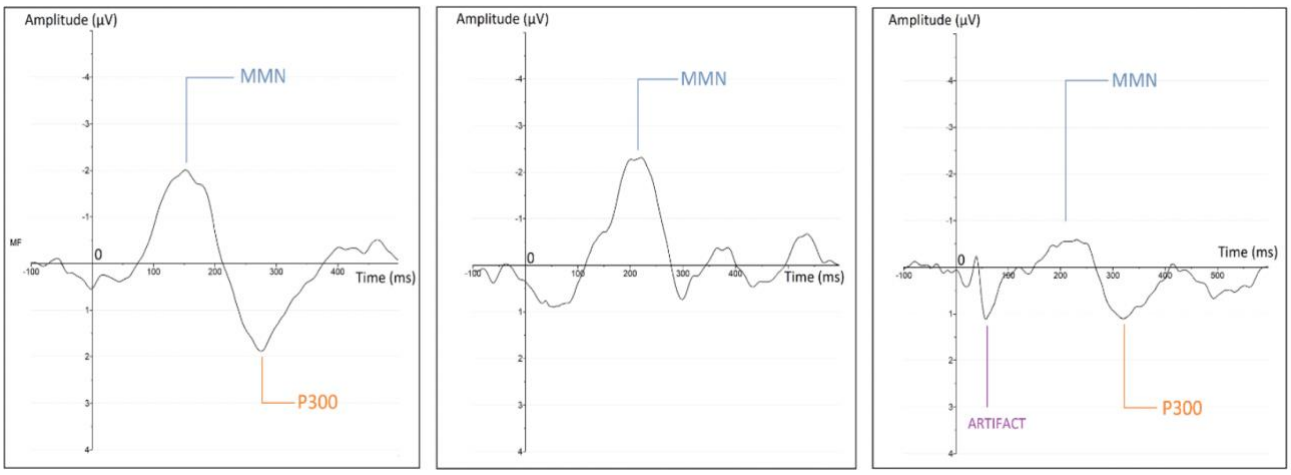
Return to Program/Retour au programme

The effect of increasing acoustic and linguistic complexity on auditory processing: an EEG study

Fareeha S. Rana, Dr. Daniel Pape, and Dr. Elisabet Service, McMaster University

The mismatch negativity (MMN) is a neurophysiological marker of pre-attentive processing that is elicited when a rare, deviant stimulus is presented in a stream of more frequent, standard stimuli (oddball paradigm). Recorded using electroencephalography (EEG), the MMN is sensitive to acoustic differences between oncoming stimuli. In linguistic research, MMN has often been used to investigate individuals' perceptual boundaries of phonemes, or to study phoneme learning. MMNs have been elicited by syllables as well, showing a larger response to contexts that form words as opposed to those that form pseudowords. Previous work has compared speech and non-speech sounds and investigated how the MMN response differs. However, the results are not conclusive. Some studies have shown a larger response to speech compared to complex waves, but smaller in comparison to simple tones (e.g. Sorokin et al., 2010). Other works comparing non-speech stimuli have shown complex waves to elicit an MMN that is larger compared to those elicited by simple tones (e.g. Tervaniemi et al., 2000). The current study investigated pre-attentive processing of stimuli with a systematic (stepwise) increase in acoustic (and linguistic) complexity. Three different types of acoustic stimuli were presented in an oddball paradigm where deviants differed in pitch only: simple (pure sine wave) tones, complex harmonic waves, and artificially produced syllables generated with an articulatory synthesizer. The syllables had consonant- vowel (CV) structure, beginning with voiceless stops. The simple tones had a frequency of 1000 Hz (deviants: 1200 Hz). The complex waves were generated with a fundamental frequency of 100 Hz (deviants: 120 Hz), and their intensity envelopes were matched to those of the syllables. Finally, syllables with no semantic content (/ti tu tə ka/; fundamental frequency 100 Hz) were generated using VocalTractLab (Birkholz et al., 2019), a software that allows highly controllable manipulation of vocal tract and laryngeal parameters and configurations to generate natural-sounding artificial speech. Thus, complex waves and artificial syllables were acoustically very comparable, the main difference being the presence versus absence of both vocal tract influences and laryngeal control (e.g. voiceless VOT phase with transient and noise elements driven by articulator interactions). We hypothesised that the pitch deviant for each stimulus type would evoke an MMN, and that the size of this response would differ depending on the nature of the stimuli (simple, complex wave, or syllable). Results supported our hypothesis. An MMN was elicited for all three types of stimuli (**Figure 1**). The largest response was evoked to simple tones, and the smallest to the syllables. The responses to the complex waves were more similar to those evoked to sine waves than to syllables. This suggests differences between how linguistic stimuli (syllables) are processed compared to non-linguistic stimuli (simple, complex waves). The results of this study demonstrate how incremental increases in acoustic complexity and resemblance to speech of auditory stimuli are reflected in the MMN response.

Figure 1. *The figures show the MMN response evoked to the frequency (pitch) deviant for the three stimulus types: simple tones (left-most), complex waves (middle), and syllables (right-most). Time ms is plotted on the x-axis, and amplitude in µV is plotted on the y-axis.*



Return to Program/Retour au programme

Turkish (*so-called*) '*Yumuşak g*' (ğ), long vowels, and the phonology-morphology interface

Nicolas Royer-Artuso, UQAC – Université du Québec à Chicoutimi

While Turkish ***yumuşak g*** 'soft g' (**ğ**) has received some attention in the literature, I will argue that some of its properties have not been connected properly to other phenomena in the language. This constitutes one of the missed opportunities to understand Turkish's vowel system and phonology: when we take into account the *synchronic* properties of **ğ** (i.e. when we do not map diachrony onto synchrony), some other problematic facts of the language receive straightforward solutions.

In Turkish reference grammars, we find statements like the following:

(1) "The so-called 'soft g' lacks a corresponding 'consonantal' sound in standard Turkish [...] When it is in word-final or syllable-final position, it lengthens a preceding back vowel [...] Between identical back vowels it is inaudible [...]", and so on. (Göksel & Kerslake [1]:7).

(2) "Long vowels occur in words borrowed from Persian and Arabic" (Ibid.:11).

(1) and (2) contradict each other. What is described in (1) corresponds phonetically to a long vowel and *do occur* in 'native' words. No reasons other than 1) the Phonemic Principle; what we could call 2) the 'Contrastive Principle' and the 'Complementary Principle'; and 3) Etymology, are responsible for the current synchronic analyses of Turkish **ğ**. Once we abandon or at least question some of these principles –as many frameworks tend more and more to do- we get a different picture of Turkish phonology and its processes: there is no **ğ** phonemically. Arguing thus, as the work of some phonologists and/or phoneticians allows us, permits a new perspective on other phenomena of the language.

For example, Pöchtrager [3] shows that the traditional analysis of the 'k-∅ alternation' (∅=**ğ**) as the result of a productive phonological rule/process does not hold. If we further modify our statements about the Turkish system of vowels as well as our statements about its phonology, there is no 'k-∅ alternation' problem remaining (something Inkelas showed indirectly–see [2] pp. 84; 86 and 97).

Some problems regarding Turkish Vowel Harmony(VH) also receive a different light once we modify our statements about **ğ**. VH is sometimes blocked in the context $CV:_1CV_2$, but not always. We also find free variation. Once we accept the fact that long vowels do exist in Turkish, another picture of VH emerges.

I will conclude by arguing that the data I present give weight to exemplar-based and/or word- based approaches to phonology and/or morphology.

References cited

[1]Göksel,A. & C.Kerslake.2005.*Turkish:A comprehensive Grammar*.Routledge.
[2]Inkelas,S.2011.The Interaction Between Morphology and Phonology.In Goldsmith, Riggle & Yu(eds.)*The Handbook of Phonological Theory.*Blackwell.
[3]Pöchtrager,M.2013.k-∅ and What Phonology Can Do.*Dilbilim Araştırmaları Dergisi*2013/1.

Return to Program/Retour au programme

ROOT-first Deletion in Harmonic Serialism

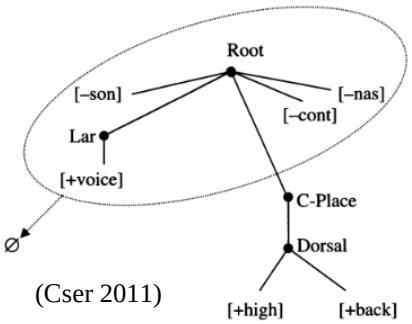
Nate Shaftoe, York University

It is well-established in Harmonic Serialism (HS) research that deletion is a two-step process, with the PLACE node being deleted in one step, and the ROOT node in a following step (McCarthy 2008, 2019). This can be seen in the alternation /pat.ka/→[pa.ka], which in McCarthy's analysis has an intermediate form of [paʔ.ka] (*ibid*). This process has many interesting implications for HS, such as the prediction that where full deletion exists glottal deletion will necessarily occur as well, not all of which have yet been properly explored. One of these is an outstanding question which seems to have not been discussed at all: the possibility of deleting the ROOT node first.

As discussed in McCarthy (2008, 2019), in HS the PLACE node is deleted first, followed by the ROOT node. This makes intuitive sense, and is borne out rather clearly in the data from which the assumption was developed, which focused on cluster simplification and the universal tendency for codas to reduce rather than onsets. The ROOT node should be deleted last, since it is the defining foundation of the segment, and can be pronounced without the presence of a PLACE node, as demonstrated by the existing of the glottal segments. McCarthy (2019) provides the evidence that PLACE nodes can be observed to be deleted first, with ROOT deletion always occurring second. This pattern holds synchronically and, quite importantly, diachronically. It also leaves open a logically possible alternative: ROOT-first deletion. In such a situation, the ROOT node would be deleted in the first step, leaving behind a bare PLACE node. In a following step the PLACE node could be deleted, leading to the same output as PLACE-first deletion, albeit with different implications for the intermediate form. The other option is for the PLACE node to be left behind 'stranded' by the deletion of its ROOT node.

Such a rootless segment is functionally impossible to pronounce: a segment which consists only of a PLACE node lacks any kind of manner of articulation. Rather, the vocal tract would only have the information for where the sound should be produced, but have no means by which to produce the sound, or even to articulate (since even the slight raising of the tongue for an alveolar segment would be disallowed, due to a lack of information as to the stricture of the segment). It would, however, exist for phonological purposes: it could in theory trigger spreading or assimilation, which would only manifest in other segments which still have intact ROOT nodes.

This proposed process has, in fact, some empirical grounding. It has been argued that in Latin, the word-initial sequence [gn] underwent ROOT deletion in the [g], leaving behind a rootless dorsal PLACE node (Cser 2011). This rootless PLACE node then can attach to preceding segments after affixation: *in+gnoscere→ignoscere*, while *re+gnoscere→ renoscere*, lacking the velar place. Note that in *ignoscere*, the <g> would have been pronounced as [ŋ], which more clearly demonstrates the place-assimilation process. Unless there is some preceding segment susceptible to PLACE-assimilation, the stranded PLACE node is left unpronounced. This project provides an HS analysis of Cser's Latin data, demonstrating how it fills the niche of ROOT-first deletion. It will investigate why the rootless PLACE node remains unpronounced, as it could either be forced to delete or could simply remain unparsed.



(Cser 2011)

This analysis could have useful implications for floating features in HS, allowing them to be analyzed in terms of PLACE nodes being stranded after ROOT deletion. It also provides insight into the function of feature nodes in phonology, implying that the ROOT node is responsible for pronunciation in the output. As well, it furthers the reintegration of feature geometry and HS by adapting a feature-geometric analysis into an HS context, demonstrating how the two models can interact.

Cser, A. (2011). The floating C-place node in Latin. *Journal of Linguistics 47*:65-85.
McCarthy, J. J. (2008). The gradual path to cluster simplification. *Phonology 25*:271-319. McCarthy, J. J. (2019). How to delete. *Perspectives on Arabic Linguistics XXX:7-32*.

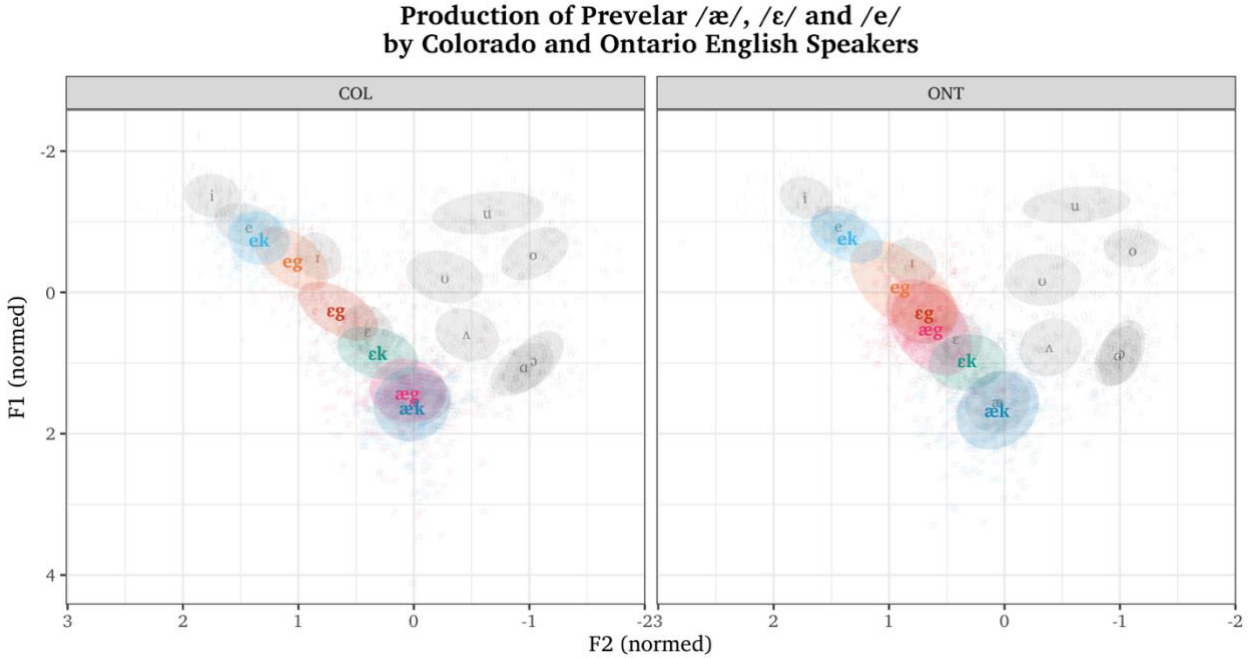
Return to Program/Retour au programme

BAG-raising in Ontario and Colorado

Lisa Sullivan, University of Toronto

BAG-raising is a process in which some speakers of North American English raise /æ/ before /g/, but not /k/. This process occurs in parts of Canada (particularly BC) (e.g. Boberg 2008; Esling & Warkentyne 1993), and in the Pacific Northwest (e.g. Riebold 2018; Freeman 2021), Montana (Bar-El et al. 2017) and Upper Midwest (e.g. Koffi, 2013, Bauer & Parker, 2008) in the US. Beyond these regions, evidence of the distribution of BAG-raising is limited, but does suggest that it may be localized to Canada and US states close to its border (Stanley 2018; 2019). The current study extends our knowledge of BAG-raising by documenting its production in Ontario and Colorado. BAG-raising and Canadian raising aside, speakers from Ontario and Colorado are expected to have similar vowel systems (Holland & Brandenburg, 2017; Boberg 2010). Data from previous studies (Boberg, 2008; Sullivan 2020), suggests Ontarians should BAG-raise. There is no data for Colorado; however, as it is not part of a region where BAG-raising has been documented, we expect that Coloradans will not BAG-raise.

23 Ontarians and 22 Coloradans were recorded saying monosyllabic words with /æ, ɛ, e/ before /g/ and /k/, as well as 11 English vowels before /d/ and /t/ using Gorilla (Anwyl- Irvine et al. 2019). The recordings were annotated using the Montreal Forced Aligner (McAuliffe et al. 2017) and manually corrected in Praat (Boersma & Weenink, 2021). F1 and F2 at the midpoint of each vowel was extracted using Praat, converted to Bark (Zwicker & Terhardt, 1980) and z-score normalized after excluding outliers. The results are displayed in the figure below which shows that is /æg/ is raised almost to /ɛg/ for Ontarians, but is unraised for Coloradans. This suggests that Ontarians BAG-raise, but Coloradans do not, thereby confirming previous findings that BAG-raising is present in Ontario English and demonstrating that BAG-raising is not a feature of Colorado English.



Return to Program/Retour au programme

Phonetic interference in the production of stops by Western Armenian bilinguals

Talia Lisa Tahtadjian, University of Ottawa

Armenian is a relatively understudied language that constitutes an independent branch of the Indo-European language family. It has two standard dialects today: Western Armenian is spoken mainly in the diaspora (WA); while its counterpart, Eastern Armenian, is mainly spoken in Armenia, Russia, and Iran (EA). The most salient differences between the production of WA and EA primarily include laryngeal contrasts in stops and affricates. EA traditionally has a 3-way laryngeal contrast in stops and affricates (voiced vs. voiceless vs. voiceless aspirated), while WA has a 2-way contrast (voiced vs. voiceless aspirated) [1,2].

Voice onset time (VOT) is an established acoustic property that distinguishes laryngeal contrasts in different languages [3,4]. Bilingual speakers have shown that they produce language-specific VOT differences in voicing, even if they sound like monolinguals [5]. While VOT is a property distinguishing how L1 and L2 phonetically interfere with one another, the degree and direction of the interference is not always the same for each bilingual population [6].

The present study examined the VOT of voiceless and voiced stops in two populations of WA speakers who were simultaneous bilinguals in either Canadian English or Lebanese Arabic, residing in Canada and Lebanon, respectively. Twenty-six participants (with an additional 5 Canadian participants' data still being analyzed) produced word-initial stops /th,d,kh,g/ phrase-initially and phrase-medially, in unilingual and code-switching conditions (e.g. an Armenian target word in an English carrier phrase). Armenian-English bilinguals (n = 9) showed some VOT differences between the unilingual conditions – a higher long lag VOT for voiceless stops in English than Armenian and a similar long negative VOT for voiced stops in English and Armenian, while Armenian-Arabic bilinguals (n = 17) did not show significant voicing differences between their L1 and L2 – producing a contrast between lead and short lag in both languages (Figure 1). In code-switching conditions, Armenian- English bilinguals produced no significant shifts in VOT, while shoring a tendency towards higher voiceless VOT in English contexts. Some limited differences between code-switching conditions were observed for Armenian-Arabic bilinguals.

Figure 1. Mean VOT (ms) in Western Armenian bilinguals

	Arabic	***WA (Lebanon)***	***WA (Canada)***	***English***
Voiced stops	-70	-99	-94	-76
Voiceless aspirated stops	+31	+40	+73	+96

Thus, results addressed that the degree of phonetic contrast between L1 and L2 affected the realization of phonetic interference, leading to different patterns in the two Western Armenian bilingual groups, and that phonetic interference does not always occur in the speech of simultaneous bilinguals. In WA bilinguals who code-switch, the influence of English might be more evident than the influence of Arabic, since WA spoken in Canada and English are phonetically different, while WA spoken in Lebanon and Arabic are not.

References

[1] Hacopian, N. (2003). A three-way VOT contrast in final position: Data from Armenian. *Journal of the International Phonetic Association*, 51-80.

[2] Vaux, B. (1998). *The Phonology of Armenian*. Oxford University Press.

[3] Lisker, L. & Abramson, A. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements, *WORD*, 20:3, 384-422.

[4] Cho, T., Whalen, D. H., & Docherty, G. (2018). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics*, *72*, 52-65.

[5] Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism*, *9*(1), 97.

[6] Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*, *39*(4), 558-570.

Return to Program/Retour au programme

Investigating universality of consonant intrinsic F0 effects

Connie Ting, McGill University

It is well known that vowels vary systematically in F0 based on intrinsic properties of a preceding obstruent; F0 following voiceless obstruents tends to be higher than F0 following voiced obstruents [1,2,3]. Within the vast literature investigating this consonant intrinsic F0 effect (henceforth CF0 effect, following Kingston 2007), this pattern has been individually attested in most of the languages in which it has been studied, raising the possibility that it is a universal phenomenon stemming from a physiological source. We might thus expect that the effect is robust such that its effect size and duration are similar across languages. However, the details of previous studies' results on CF0 effects vary in several ways, calling into question this idea of 'universality'. For example, while many report the effect to exist primarily in the onset of the vowel, some have found that the effect can persist to the mid-point or even through the whole vowel [4]. There have also been cases where the effect was found to be in the opposite direction [5,6], further conflicting with the notion of 'universality'. Moreover, the CF0 effect may have the potential to be phonologized into a tone contrast on the following vowel, leading languages to undergo sound change (e.g. tonogenesis and tone splitting [7,8]). It is thus necessary to consider these differences and establish the cross-linguistic robustness of the effect, which can tell us more about the CF0 effect as a possible phonetic precursor to sound change and about the path of sound change processes more generally. To do this, we need a single study comparing many languages. The present study offers a comparison of the CF0 effect across 16 languages, using large corpora of read speech.

Data & Methods: This study examines read speech of 16 languages (Croatian, English, French, German, Hausa, Korean, Mandarin, Polish, Portuguese, Russian, Spanish, Swahili, Swedish, Thai, Turkish, and Vietnamese) from the GlobalPhone [9] and LibriSpeech [10] language databases. All data was force-aligned using the Montreal Forced Aligner [11]. F0 contours were extracted using Praat for [a, i, u] vowel tokens in utterance-initial obstruent-vowel syllables. F0 measurements were then transformed to semitones (within-speaker) and further data-processing was done to exclude unreliable F0 measurements (e.g. excluding vowels shorter than 50ms, excluding vowels where pitch extraction could not be obtained for at least 50% of the vowel). The resulting datasets include 69-132 speakers and 0.7-8.1k vowels per language. For each language, the trajectory of F0 over time was modeled as a function of voicing. More specifically, a Generalized Additive Mixed Model was fit, including F0 (semitones) as the dependent variable, with VOICING (of the preceding consonant), VOWEL, and TONE (where applicable) as predictor variables. Two smooths were included to (i) model the non-linear pattern of F0 over time (normalized) and (ii) model F0 over time separately for the different levels of VOICING. Variation across speakers and words was accounted for by random smooths and random intercepts, respectively.

Results Figure 1. shows the estimated difference in F0 (semitones) between voicing categories over the time-course of the vowel for each of the 16 languages. This shows us not only the size of the difference at the onset of the vowel, but also the time at which the difference is estimated to be no longer significant. The results suggest that the CF0 effect is fairly robust; it is present (and positive) in all 16 languages examined here, supporting the universality of the CF0 effect. However, the effect shows variability across languages both in terms of effect size (0.64~4.52 ST) and duration (0.12~1 normalized time). The comparison of effect size and duration across languages is further summarized in Figure 2. Crucially, there does not seem to be a clear correlation between effect size and duration. This suggests that the CF0 effect, while present across languages, can vary in at least two independent ways. Studies of this kind offer a more comprehensive examination of the robustness and variability of the CF0 effect across languages and provide insight on the status of the CF0 effect as a phonetic precursor to sound change.

Return to Program/Retour au programme

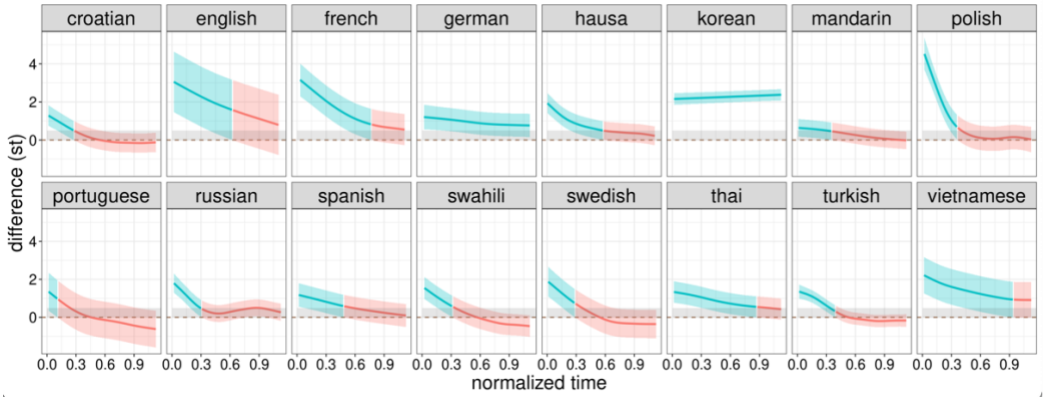


Figure 1. Difference smooths of F0 between voicing contrasts, with shaded pointwise 95%-CI. Colour of shading reflects significance of difference between voicing contrasts; **blue**: significant difference, **red**: non-significant difference (the effect is considered not significant once the CI overlaps with 0). For languages with a three-way laryngeal contrast (e.g. Korean, Thai), contrast of interest is between the least and most voiced categories. Gray shading between 0-0.5 difference represents approximately the area of just noticeable difference (following Jongman et al. 2017).

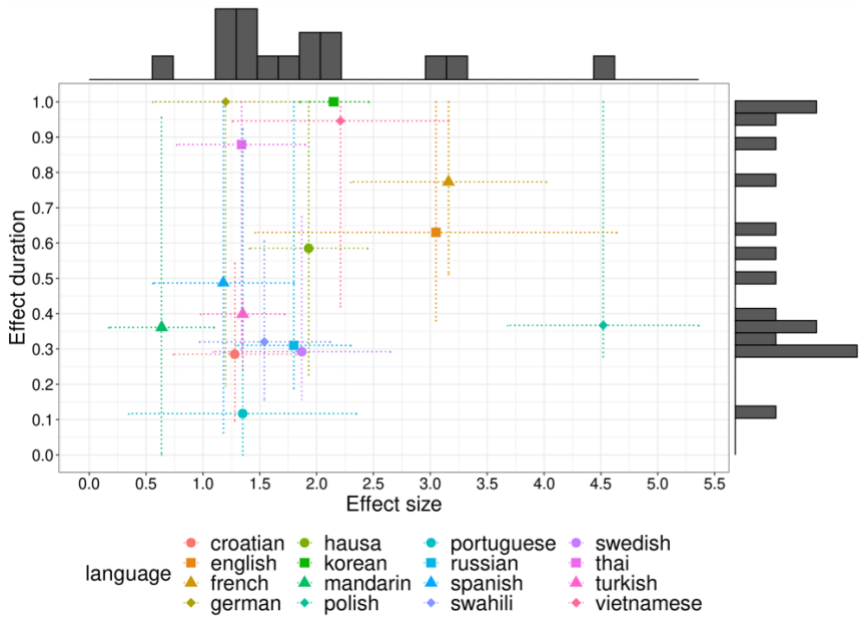


Figure 2. Variation of CF0 effect size (F0 difference between voicing contrast at vowel onset) and effect duration (time when difference is no longer significant) across languages. Error bars represent 95% CI.

References

[1] House, A. S., & Fairbanks, G. (1953). The influence of consonant environment on secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America 25*, 105-113.

[2] Hombert, J.-M., Ohala, J. J., & Ewan, W. G. 1979. Phonetic explanations for the development of tones. *Language 55*, 37-58.

[3] Mohr, B. 1971. Intrinsic variations in the speech signal. *Phonetica 23*, 65–93.

[4] Jessen, M., and Roux, J. C. 2002. Voice quality differences associated with stops and clicks in Xhosa, *Journal of Phonetics 30*, 1–52.

[5] Xu, C. X., and Xu, Y. 2003. Effects of consonant aspiration on Mandarin tones. *The Journal of the Acoustical Society of America 33*, 165-181.

[6] Francis, A. L., Ciocca, V., Wong, V. K. M., & Chan, J. K. L. 2006. Is fundamental frequency a cue to aspiration in initial stops? *The Journal of the Acoustical Society of America 120*, 2884-2895.

[7] Maddieson, Ian (1984). The effects on F0 of a voicing distinction in sonorants and their implications for a theory of tonogenesis. *Journal of Phonetics 12*(1), 9–15.

[8] Kingston, John (2011). Tonogenesis. van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), The Blackwell Companion to Phonology, Wiley-Blackwell, Oxford, 2304–2333.

[9] Schultz. T (2002) "Globalphone: A Multilingual Speech and Text Database Developed at Karlsruhe University," in Proceedings of the ICSLP, pp. 345–348.

[10] Panayotov, V., Chen, G., Povey, D., & Khudanpur, S. (2015). "Librispeech: an ASR corpus based on public domain audio books," in Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2015.

[11] McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, and Morgan Sonderegger (2017). Montreal Forced Aligner [Computer program]. Version 0.9.0, retrieved 17 January 2017 from http://montrealcorpustools.github.io/Montreal-Forced-Aligner/.

Return to Program/Retour au programme

Bilingual Perception of Voicing Contrasts in English and Spanish Pseudowords

Owen Ward and Vanina Machado
University of Toronto

In this study, experimental evidence is presented showing how language experience variables like age of acquisition (AOA), language dominance (LD) and language confidence (LC) affect bilingual perceptual accommodation. Following Gonzales and Lotto's (2013) procedures, 10 Spanish-English bilinguals' perception of voice-onset-time (VOT) contrasts was examined to see if the results of that study -that early bilingual experience promotes language-specific phonetic systems-, can be corroborated in late bilinguals varying in AOA, LD and LC.

VOT in Spanish is shorter than in English, with Spanish voiceless stop VOT corresponding to that of English voiced stops (Gonzales & Lotto, 2013), bilinguals need to overcome this overlap in VOT to achieve native-like perceptual accuracy. While early bilinguals are expected to pattern closely with monolinguals, late bilinguals are likely to vary, as representing a second language is more challenging once native-language categories are firmly established (Flege, 1995; Kuhl, 2008. Casillas and Simonet (2018) show that late bilinguals effectively manage to perceptually differentiate VOT like early bilinguals but with more variability. In addition, Duncan (2019) confirms that the level of LC is directly correlated with bilinguals' perceptual categorization.

Our aim then is to see if AOA, LD and LC can predict trends in bilingual VOT perception. We also intend to compare them with 10 purely monolingual speakers. To achieve this, participants completed an identification task in LabVanced (Finger et al., 2017), where manipulated pseudowords /bafri/-/pafri/ along a VOT continuum, in two language contexts, were randomly presented one at a time. Unlike other studies, bilinguals were not made aware of both language contexts. Instead, instructions were written in Spanish and English so that linguistic judgment was not preconditioned, controlling for language context effects.

Our pilot results confirm that Gonzales and Lotto's (2013) observations do not always apply to late bilinguals. Like in Casillas and Simonet (2018) late bilinguals can perceptually discriminate but present difficulties, e.g., perceiving more /bafri/ at 20ms in the English context. AOA did not seem to predict accuracy as much as the other variables, as higher accuracy rates were seen for responses in their dominant and more confident language context, confirming our predictions. Percentage of crossover VOT difference and level of English–Spanish confidence strongly correlated, showing larger phonetic boundary shifts between language contexts. Monolinguals showed phonemic boundary shifts that were unexpected, possibly an effect of the type of task.

References

Casillas, J. V., & Simonet, M. (2018). Perceptual categorization and bilingual language modes: Assessing the double phonemic boundary in early and late bilinguals. *Journal of Phonetics*, *71*, 51-64.
Duncan, G. (2019). Bilinguals' double phonemic boundary: Not one of normal nature. *Master's Theses*, 1389.
Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). *LabVanced - Online Experiments Made Easy* [Computer Software]. Retrieved from: https://www.labvanced.com/.
Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.
Gonzales, K., & Lotto, A. J. (2013). A bafri, un pafri: Bilinguals' pseudoword identifications support language- specific phonetic systems. *Psychological science*, *24*(11), 2135-2142.
Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979-1000.

Return to Program/Retour au programme

Cue primacy effects in Mandarin tone imitation

Wei Zhang, McGill University

Phonetic imitation involves perception, memory and production processes, so it can provide a window into both phonetic and phonological effects. Studies have identified several factors affecting the accuracy of phonetic imitation. In this paper, we are interested in two of them: feature type effects and phonological contrast effects. In terms of feature type effect, supra-segmental features such as fundamental frequency (F0) and duration were found to be easier to imitate than segmental features such as vowel formants (e.g. Sato et al., 2013). Phonological contrast effects are hypothesized to constrain imitation so that phonologically ambiguous sounds (e.g. between two phonemes) are not well imitated. For example, Nielsen (2011) found that lengthening the voice onset time (VOT) of /p/ was imitated by native English speakers but shortening VOT was not, possibly because the shortened VOT threatened the /p/ category. The phonological contrast effect was also observed in imitation of intonational contours (Pierrehumbert and Steele, 1989). Recent studies have investigated the role of cue primacy in imitation. Kim and Clayards (2019) explored the imitation of an /ɛ/-/æ/ continuum in English which they implemented by varying formants, the primary cue, and duration, the secondary cue. They found that for the ambiguous steps in the continua, formants were poorly imitated, whereas duration was well imitated. The accurate imitation of duration was consistent with the feature type effect. However, another possibility for why duration is not constrained by the phonological contrast effect could be that duration is a secondary cue to the vowel contrast. In this study, we aimed to tease the two possibilities apart.

Mandarin has four lexical tones primarily distinguished by their F0 contours (Chao, 1965). Syllable duration was found to serve as a secondary cue for the flat and falling tonal contrast (Yang 1989). In this study, we carried out an imitation experiment on flat-falling tonal continua created by manipulating the F0 fall range and duration. Since both cues are supra-segmental, we made the following predictions: (1) If both the F0 range and the duration are well (linearly) imitated, the feature type effects dominate the phonetic imitation; (2) if rather, only the F0 range but duration is poorly (categorically) imitated, cue primacy plays a role in the phonetic imitation.

To avoid creaky voice in utterance-final position, which could complicate the interpretation of extracted F0, we inserted the target syllable 'ba' (which carried the tonal continuum) at the beginning of a tri-syllabic phrase: *ba de sheng* [pa.de.ʂɤŋ] 'the sound of ba'). There are 7 steps for F0 falling range and 5 steps for duration (see Figure 1). Eighteen native Mandarin speakers (age: 21-29, gender-balanced) participated in the experiment through Prolific platform.

Figure 2 shows the imitation results of the target syllable. There is a clear difference in the imitation of the two features: the relationship between the target and imitated duration is roughly linear, while that between the target and imitated F0 range is non-linear, and seems to follow a bi-modal distribution. To examine the modes in the distributions of duration and F0 range, we fit one unimodal and one bi-modal model for both distributions, using the brms package (Bürkner, 2017) in R (R Core Team, 2013). The model fits were compared by means of leave-one-out (LOO) cross validation. Results showed that for the duration distribution the unimodal model significantly surpassed the bi-modal one (Expected Log Predictive Density difference = -673.4, SD = 44), and for F0 falling range, on the contrary, the bi-modal model surpassed the unimodal one (ELPD difference = -787.4, SD = 42.3). These results show that the imitation of the primary cue F0 falling range was affected by the tonal contrast, whereas the secondary cue duration was not. They also suggest that feature type does not dominate the phonetic imitation, and that cue primacy plays a significant role in phonetic imitation.

Return to Program/Retour au programme

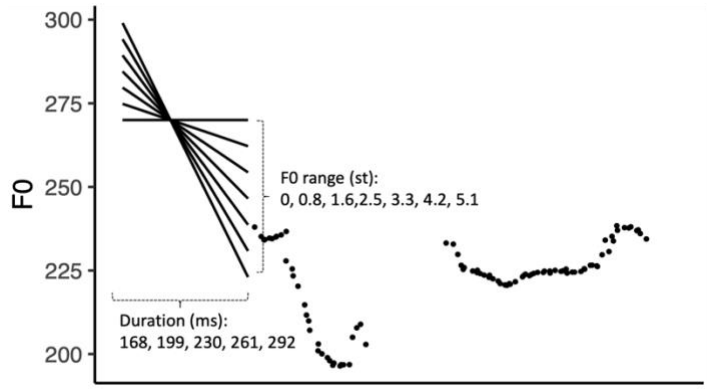


Figure 1: *The scheme of the stimuli. There are 7 steps for F0 contour and 5 steps for duration. The first syllable is the target syllable 'ba'. The last two syllables are the carrier 'de shēng'.*

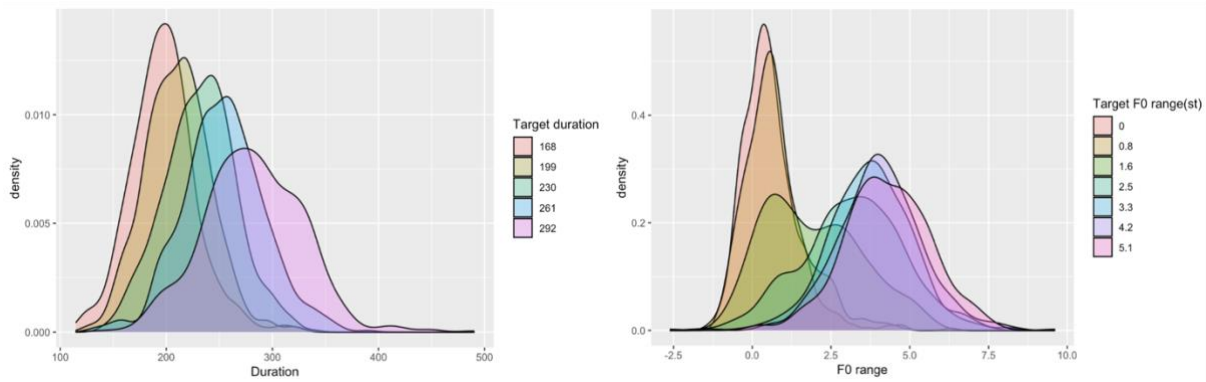


Figure 2: *Density distribution of the imitated duration (on the left) and F0 range (on the right) of the continua. Colors indicate the levels of the target stimuli in the imitation*.

References

Bürkner, P.C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Chao, Y. R. (1965). *A grammar of spoken Chinese*. Univ of California Press.

Kim, D., & Clayards, M. (2019). Individual differences in the link between perception and production and the mechanisms of phonetic imitation. *Language, Cognition and Neuroscience*, *34*(6), 769–786.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132–142.

Pierrehumbert, J. B., & Steele, S. A. (1989). Categories of tonal alignment in English. *Phonetica*, *46*(4), 181–196.

Sato, M., Grabski, K., Garnier, M., Granjon, L., Schwartz, J.-L., & Nguyen, N. (2013). Converging toward a common speech code: Imitative and perceptuo-motor recalibration processes in speech production.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria.

Yang, Yufang. 1989. "The vowels and the perception of Chinese tones." ACTA Psychologica Sinica (1): 29- 33. (In Chinese)

Return to Program/Retour au programme