# ICPhoV

International Conference on Phonetic Variation
Diversity within and across Languages

19th-21st January 2026, Granada

**BOOK of ABSTRACTS**

UNIVERSIDAD DE GRANADA

Atlas Lingüístico Interactivo de los Acentos de Andalucía

Lengua española

UNIVERSIDAD DE GRANADA
Departamento de Estudios Semíticos

# INDEX

# Contributors

# WELCOME TO ICPhoV

The International Conference on Phonetic Variation: Diversity within and across Languages (ICPhoV), hosted by the Faculty of Arts of the University of Granada (Spain) from 19th until 21st January 2026, is a forum for the presentation and discussion of research on different issues related to phonetic variation within and across languages. ICPhoV aims to explore phonetic variation across a wide variety of languages to understand variation cross-linguistically.

A total of 59 papers are represented in this conference, including plenary talks by three of the world's leading phoneticians: Paul Boersma (University of Amsterdam, The Netherlands), Ghada Khattab (Newcastle University, UK), and Daniel Recasens (Universitat Autònoma de Barcelona, Spain).

The papers presented at ICPhoV cover a range of topics, including phonetic variation within and across languages, speech technology, speech perception, L2 acquisition, diachronic phonetic variation, sociophonetics, and prosodic variation. Likewise, these papers are also varied in terms of methodological approaches (e.g., experimental phonetics, acoustic analyses, articulatory analyses) and perspectives (e.g., synchronic analyses, language change through time).

Researchers from a score of countries across five continents explore phonetic variation in over 30 languages and language varieties from around the world, including Arabic, Basque, Bengali, Czech, Dutch, English, French, German, Hindi, Italian, Khmer, Malayalam, Jin Chinese, Mandarin, Pennsylvania Dutch, Northern Lushootseed, Polish, Q'anjob'al, Russian, Serbian, Spanish, Swedish, Taiwanese Mandarin, Venetan, Vietnamese, and Whistled Gascon Occitan.

I would like to thank all the presenters for their participation in this conference, and for bringing such an impressive collection of papers to the University of Granada. I would also like to thank the extensive group of academics who have served as part of the academic committee and contributed to the high academic standard of this conference. Lastly, my sincerest gratitude goes to the University of Granada for its many contributions, which made this conference possible.

All the best
Alfie


Alfredo Herrero de Haro
Department of Spanish Language, University of Granada

# COMMITEES

- ◆ **Organising committee**

**Alfredo Herrero de Haro (chair)**
Department of Spanish Language, University of Granada

**Daniel M. Sáez Rivera**
Department of Spanish Language, University of Granada

**María Heredia Mantis**
Department of Spanish Language, University of Granada

**Rocío Cruz Ortiz**
Department of Spanish Language, University of Granada

**Antonio Manjón-Cabeza Cruz**
Department of Spanish Language, University of Granada

**Doina Repede**
Department of Spanish Language, University of Granada

**Amina Naciri-Azzouz**
Department of Semitic Studies, University of Granada

**Radka Svetozarovová**
Department of Spanish Language, University of Granada

**María Cardenete**
Department of Spanish Language, Linguistics, and Literary Theory, University of Seville

**Valentina Colonna**
Department of Linguistics and Literary Theory, University of Granada

# ◆ Scientific committee

| | |
|---|---|
| John Hajek | The University of Melbourne, Australia |
| Joshua Penney | Macquarie University, Australia |
| Kathleen Jepson | The University of Queensland, Australia |
| Plinio Barbosa | Universidad de Cambinas, Brazil |
| Alexei Kochetov | University of Toronto, Canada |
| Laura Colantoni | University of Toronto, Canada |
| Domingo Román | Universidad de Santiago de Chile, Chile |
| Radek Skarnitzl | Carles University, Czech Republic |
| Oliver Niebuhr | University of Southern Denmark, Denmark |
| Pärtel Lippus | University of Tartu, Estonia |
| Ander Egurtzegi | French National Centre for Scientific Research, France |
| Jalal Al-Tamimi | Université Paris Cité, France |
| Adrian P Simpson | Friedrich Schiller University, Germany |
| Paul O'Neill | The Ludwig Maximilian University of Munich, Germany |
| Mok Pik Ki Peggy | The Chinese University of Hong Kong, Hong Kong |
| Priyankoo Sarmah | Indian Institute of Technology Guwahati, India |
| Antonio Romano | Università degli Studi di Torino, Italy |
| Maria-Gabriella Di Benedetto | Sapienza University of Rome, Italy |
| Kakeru Yazawa | University of Tsukuba, Japan |
| Nobuaki Minematsu | The University of Tokyo, Japan |
| Takeshi Nozawa | Ritsumeikan University, Japan |
| Karim Bensoukas | Mohammed V University in Rabat, Morocco |
| Karolina Broś | University of Warsaw, Poland |
| Michał Baran | University of Szczecin, Poland |
| Sónia Frota | University of Lisbon, Portugal |
| Ibrahim Al Malwi | Jazan University, Saudi Arabia |

| | |
|---|---|
| Justus Christiaan Roux | Stellenbosch University, South Africa |
| Alfredo Herrero de Haro | Universidad de Granada, Spain |
| Álvaro Arias Cabal | Universidad de Oviedo, Spain |
| Ángeles Vicente | Universidad de Zaragoza, Spain |
| Antonio Pamies | Universidad de Granada, Spain |
| Beatriz Blecua Falgueras | Universitat de Girona, Spain |
| Carmen Aguilera Carnerero | Universidad de Granada, Spain |
| Clara von Essen | Universidad de Málaga, Spain |
| Elena Fernández de Molina Ortés | Universidad de Granada, Spain |
| Eugenia San Segundo | CSIC, Spain |
| Fernando Sánchez Miret | Universidad de Salamanca, Spain |
| Francisco Pedro Pla Colomer | Universidad de Jaén, Spain |
| Isabel Molina Martos | Universidad de Alcalá, Spain |
| Javier Perea Siller | Universidad de Córdoba, Spain |
| Jesús Jiménez | Universitat de València, Spain |
| Joaquim Llisterri Boix | Universitat Autònoma de Barcelona, Spain |
| José María Lahoz Bengoechea | Universidad Complutense de Madrid, Spain |
| Mari Rosa Lloret | Universitat de Barcelona, Spain |
| Maria Assumpció Rost Bagudanch | Universitat Autònoma de Barcelona, Spain |
| Matilde Vida Castro | Universidad de Málaga, Spain |
| Ricard Herrero Arambul | Universitat de València, Spain |
| Paolo Roseano | UNED, Spain |
| Victoria Marrero | UNED, Spain |
| Mikael Roll | Lund University, Sweden |
| Silke Hamann | University of Amsterdam, The Netherlands |
| Ghada Khattab | Newcastle University, UK |
| Niamh Kelly | Newcastle University, UK |
| Patrycja Strycharczuk | The University of Manchester, UK |

Martin Ball      Bangor Univ., Wrexham Univ., University College Cork, UK/Ireland

Janet Watson                University of St Andrews & SQU, Muscat, UK/Oman

Jeff Mielke                North Carolina State University, USA

José Ignacio Hualde        University of Illinois Urbana-Champaign, USA

Matthew Pollock            Louisiana State University Shreveport, USA

Miquel Simonet            University of Arizona, USA

Nicholas Henriksen        University of Michigan, USA

Silvina Bongiovanni        Michigan State University . USA

# INAUGURAL TALK

**How acquiring a bidirectional grammar restricts variation**

Paul Boersma

*University of Amsterdam*

## 1. The setting

When modelling the acquisition of phonetics and phonology by computer simulations, we often model the virtual learners as *probability-matching*, which means that such learners mimic the degree of variation that they observe in their environment. At first sight, this should lead to big problems, because the amount of variation is expected to increase diachronically, i.e. with every generation of learners.

## 2. The problem

To see the problem, imagine first the simplest case of *iterated learning*, where one child learns the language from one adult: if the two speak only in extremely silent conditions, the child will end up having the same amount of noise as her parent. This situation almost never occurs, however.

First, there is typically noise in the *transmission channel* (from adult to child), i.e. there will be wind, animals barking, random muscle activity in the adult's speech organs, and blood flow in the child's ears. This will cause the probability-matching child to end up with more variation than the adult, and this problem will grow over the generations.

Second, there will usually be multiple speakers that the child learns from, and they speak differently. To the extent that the child won't be able to normalize the speaker differences away, she will regard the between-speaker variation as genuine variation in her environment and end up producing more variation than any of her adult models.

## 3. External solutions

To keep the exploding variation at bay, some computational researchers have proposed mechanisms that explicitly reduce noise in the child, such as cooling down the *temperature* in the child's virtual brain, or more teleological (goal-oriented) mechanisms such as built-in functional biases toward communicative success.

## 4. An internal solution: bidirectionality causes emergent restriction of variation

In this talk I assume *bidirectionality of processing*, which is the idea that language users employ the same decision mechanisms (constraint rankings or connection weights) in their speech

production as in their speech comprehension. This bidirectionality idea is explicit in associative networks (Kohonen 1982, Hopfield 1982), in symmetric combinations of instar and outstar learning rules (McMurray et al. 2009), in Deep Boltzmann Machines (Salakhutdinov & Hinton 2009) and in some Optimality Theory work (Smolensky 1996).

Bidirectionality of processing has been proposed for explaining basic phenomena in language, such as auditory dispersion (Boersma & Hamann 2008) or the emergence of the maxims of Grice (1975) in pragmatics (Boersma 2009), and readily transfers to anti-synonymity effects in semantics, morphology and syntax.

As shown by Boersma & Hamann (2008: 250) for a toy example of the evolution of the Polish three-sibilant system, this same bidirectionality *automatically* leads to restrictions on variation. In Boersma & Hamann's Optimality-Theoretic example of one-on-one iterated learning, bidirectionality of constraint weights (using the same connections in production that have optimized your comprehension earlier) turned out to automatically counteract the transmission noise, leading to an equilibrium in the amount of noise transmitted between generations (left picture, showing both dispersion [in black] and noise stability [in grey]):
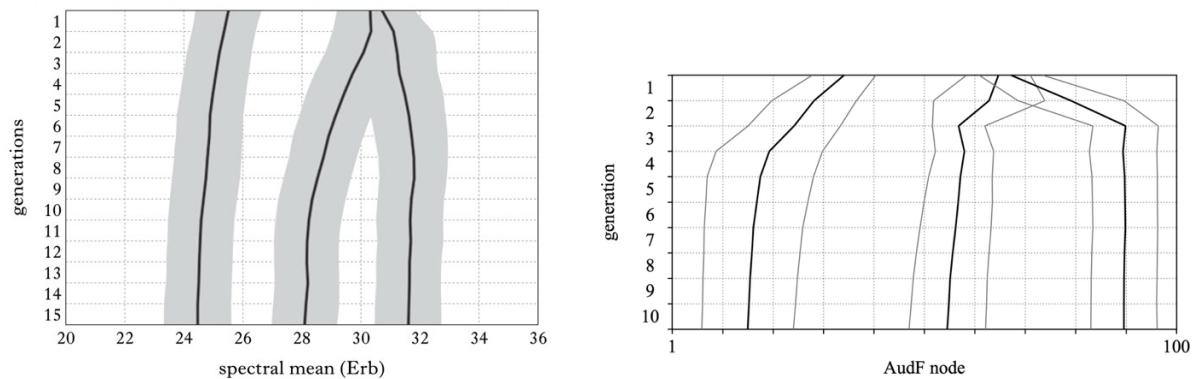


The same type of restriction on variation (for the same example) can also be seen to emerge in the inoutstar-based neural-network simulations by Seinhorst (2012: 58; right picture above) and Boersma, Benders & Seinhorst (2020). The effect is therefore robust: it appears in OT as well as in various types of artificial neural networks, as long as the connections are bidirectional, i.e. as long as the connection from node A to node B has the same strength as the connection from node B to node A.

In this talk I show with live computer simulations that bidirectionality automatically leads to restricted variation, for several more realistic phenomena than the toy examples published about before. The pictures will be much more interesting than the ones above, which work on a single auditory continuum: I show at least F1-versus-F2 spaces, palatalization-versus-place spaces, and duration-versus-F1 spaces, always based on experimental data (where sound change counts as an experiment of nature; Ohala 1993).

My talk also addresses how virtual learners handle additional sources of variation, such as learning from multiple adults who have incompatible sound systems.

## 5. The future
Ideally, one would like the simulations to work on large language models, such as on Whisper (Radford, Kim, Xu, Brockman, McLeavey & Sutskever 2023) or Wav2Vec (Baevski, Zhou, Mohamed & Auli 2020). However, we haven't been able yet to discern bidirectionality effects in such networks, perhaps because they have been trained only on production-like tasks, even when just listening, as these networks are trained to predict what comes next (Boersma & De Heer Kloots submitted). I do expect that such large networks, which have the advantage of working for whole languages, and the small toy networks that we as linguists are studying, will

close in on each other within 7 to 12 years or so. That would be a great day for linguistics and technology alike.

## 6. References

Baevski, Alexei, Henry Zhou, Abdelrahman Mohamed & Michael Auli (2020). wav2vec 2.0: a framework for self-supervised learning of speech representations. In *Proc. 34th Conference on Neural Information Processing Systems*, 12449–12460.

Boersma, Paul (2009). Unidirectional optimization of comprehension can achieve bidirectional optimality. Talk at the 10th Szklarska Poręba Workshop on the Roots of Pragmasemantics, Schreiberhau.

Boersma, Paul, Titia Benders & Klaas Seinhorst (2020). Neural networks for phonology and phonetics. *Journal of Language Modelling* 8(1): 103–177.

Boersma, Paul & Marianne de Heer Kloots (submitted). Computational emergence of the phonology–phonetic interface.

Boersma, Paul & Silke Hamann (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217–270.

Grice, Herbert Paul (1975). Logic and conversation. In P. Cole & J.J. Morgan (eds.), *Syntax and semantics 3: speech acts*. Academic Press, New York NY. 41–58.

Hopfield, John Joseph (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* 79:2554–2558.

Kohonen, Teuvo (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics* 43: 59–69.

McMurray, B., Jessica S. Horst, Joseph C. Toscano & Larissa K. Samuelson (2009). Towards an integration of connectionist learning and dynamical systems processing: case studies in speech and lexical development. In John P. Spencer, Michael S.C. Thomas, and James L. McClelland (eds.), *Toward a unified theory of development: connectionism and dynamic systems theory re-considered*, pages 218–249. Oxford University Press, New York NY.

Ohala, John J. (1993). Sound change as nature's speech perception experiment. *Speech Communication* 13: 155–161.

Radford, Alex, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey & Ilya Sutskever (2023). Robust speech recognition via large-scale weak supervision. *Proc. 40th International Conference on Machine Learning*, 28492–28518.

Salakhutdinov, Ruslan R., and Geoffrey E. Hinton (2009). Deep Boltzmann machines. In *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics*, Clearwater FL.

Seinhorst, Klaas (2012). *The evolution of auditory dispersion in symmetric neural nets*. MA thesis, University of Amsterdam.

Smolensky, Paul (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27: 720–731.

# KEYNOTE

**Accent variation and language processing: a tug-of-war between efficiency and social justice**
Ghada Khattab
*Newcastle University*

In the process of speech perception, production or learning, individuals may use any or all of the sources of variation at their disposal to engage with the task at hand. These range from properties of the speech signal to social and contextual information, and this information is stored, updated and/or created on the fly depending on the demands of the situation. Linguistic and social variability is embedded in most models of language processing, even if there is disagreement with regards to where, how, or indeed if it is represented in the grammar.

If the type of variation described above is useful/required for successful everyday communication, then surely it is central to the repertoire of every speaker/hearer. It seems, however, that not all variation is created equal when it comes to what varieties we expect speakers to aspire to and how we expect speech processing and learning to unravel. The tyranny of 'standard' and so-called 'native' varieties has for generations of research plagued our assumptions regarding the targets that speakers/hearers are aiming for and the patterns exhibited in their speech production/perception. This has influenced everything from notions of linguistic competence to models of L2 attainment and their applications, with their focus on pushing unrealistic homogeneous varieties in the public and private spheres (e.g. education, media, and second language teaching).

While a deficiency-oriented view of non-standard and L2 accents is constantly being challenged, some prevailing assumptions/ideologies about targets are hard to shift and, as a result, stand in the way of multilingual grammars being considered the norm.

In work that explores whether there is any disadvantage in the acquisition of target-like vowels by L2 learners if L2 accents are used alongside their L1 counterparts in training, I present collaborative work using high phonetics variability training (HVPT) in a foreign language classroom with low-intermediate learners under three exposure conditions: 1) a single L1 variety (SSBE), 2) multiple L1 varieties (SSBE, GenAM, AusE) or 3) a mix of L1 and L2 varieties (SSBE, Saudi English). Learners carried out a series of vowel perception tasks (identification, auditory, and category discrimination) before, during and following an extended

period of training. The results show no significant difference between the accuracy of learners across the three conditions. Further, all three groups were able to generalise learning to new speakers and varieties. The study showed that exposure to accent variation in training low intermediate learners on vowels did not hinder their learning; further the inclusion of an L2 variety was also successful and in fact more ecological in terms of representing the type of English that these learners are typically exposed to.

I then move on to psycholinguistic work which explores the cognitive load involved in processing L1 and L2 varieties by L1 listeners in order to test for the processing cost that is expected when interacting with L2 speakers. Using a change detection task with participants who heard stories by L1, L2 advanced, and L2 intermediate speakers, our results showed no difference in accuracy of recall of the speech of L2 speakers compared with that of L1 speakers. However, in a task which involved recalling one's own speech following interacting with L2 versus L1 speakers, accuracy dropped in the L2 condition, but this did not interact with L2 proficiency in the expected way. A closer inspection of the nature of the interactions shows that integrating social-indexical information in the processing of an accent and adaptation to the speaker may lead to more successful communication despite the cognitive cost.

Adaptation also plays a role in processing unfamiliar L1 varieties, as I show in the final set of studies which explores the interplay between familiarity, intelligibility and attitude. Here listeners carry out a range of tasks following listening to familiar or unfamiliar L1 varieties. The main findings demonstrate that familiarity may play an advantage in terms of speed and/or accuracy in some tasks (lexical decision) but not others (recall). Listening conditions, individual voice characteristics and linguistic scaffolding all play a role, and so does attitude towards an accent or a speaker. Adaptation is evident in the way accuracy can increase over the course of the task. These findings suggest that increased exposure to variation combined with changes in attitude towards speakers and varieties can improve speech processing and reduce cognitive cost.

These results point to the potential benefit of exposure to multiple varieties in everyday interaction for increasing familiarity, adaptation and accuracy in speech processing. For a long time, the onus has been put of the L2 (or non-standard) speaker to aim for a model target in order to be understood, and models of L2 attainment have celebrated those who achieve the elusive native-like (standard) target. This is far from realistic in many contexts and does not represent everyday interactions and/or opportunities. A model that celebrates varieties in phonetics training, educational contexts, as well as in public domains promotes social justice. A view of intelligibility and comprehensibility which puts the onus on the listener as well as the speaker is definitely the way forward.

## References

Alghabban, S. (2025). *Improving the perception and production of English vowels by Arabic speakers learning English as a foreign language (EFL) in Saudi Arabia*. PhD dissertation, Newcastle University. https://theses.ncl.ac.uk/jspui/handle/10443/6617

Al-Kendi, A., & Khattab, G. (2021). Psycho-Social Constraints on Naturalistic Adult Second Language Acquisition. *Languages*, 6(3), 129. https://www.mdpi.com/2226-471X/6/3/129

Baxter, F. Khattab, G. Krug, A. and Du, F. (2022). Recall of own speech following interaction with L2 speakers: is there evidence for fuzzier representations? *Frontiers in Communication*. 7:840041. https://doi.org/10.3389/fcomm.2022.840041

Docherty, G.J. and Khattab, G. (2024). Sociophonetics and clinical linguistics. In Ball, M.J., Müller, N. and Spencer, L. (Ed.). *The Handbook of Clinical Linguistics* (2nd Ed.). Wiley. https://onlinelibrary.wiley.com/doi/book/10.1002/9781119875949

Khattab, G. & Foulkes, P. (2024). Sociophonetics and Arabic. In C. Streluf (ed.) The Routledge *Handbook of Sociophonetics*. London: Routledge. https://doi.org/10.4324/9781003034636

Krug A., Khattab G., White L. (2023). The effects of accent familiarity on narrative recall in noise. In: R. Skarnitzl & J. Volín (Eds.) *Proceedings of the 20th International Congress of Phonetic Sciences – ICPhS 2023* (pp. 102-106). https://fu.ff.cuni.cz/icphs2023/www.icphs2023.org/programme/proceedings/index.html

Krug, A. (2024). *The roles of familiarity, intelligibility and attitude in the processing of native and non-native accents*. PhD dissertation, Newcastle University. https://theses.ncl.ac.uk/jspui/handle/10443/6236

Shetewi, O.; Corrigan, K.P.; Khattab, G. (2024). Children's Multilectal Repertoires: Diglossic Style-Shifting by Palestinian Children and Adolescents in Syria. *Languages*, 9, 341. https://doi.org/10.3390/languages9110341
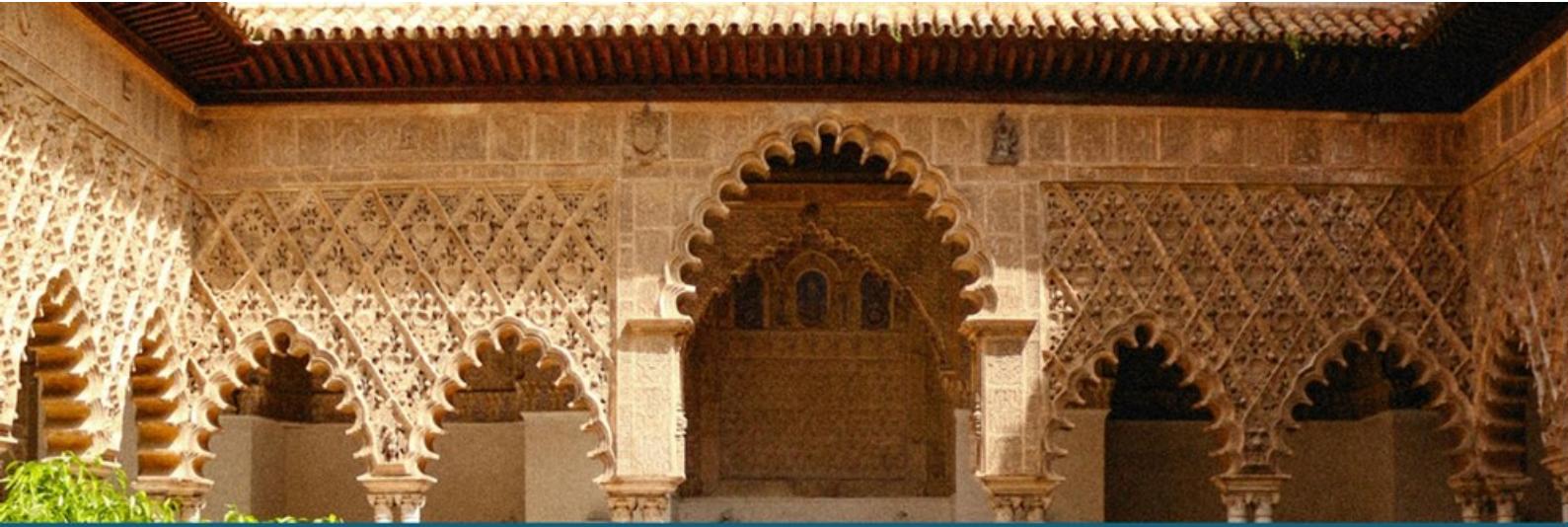
# PLENARY TALK

**From experimental phonetics to sound change: vocalization and deletion of syllable-final dark /l/ in Middle English and Early Romance**

Daniel Recasens
*Universitat Autònoma de Barcelona*

**Abstract**

This research project investigates the phonetic causes of the vocalization and deletion of syllable-final dark /l/ in /VlC/ sequences (where V stands for 'vowel' and C stands for 'consonant') in Middle English and Early Romance. In particular, it deals with an unexpected sound change process which took place in Western Romance in the Middle Ages, i.e., /l/ vocalized into [w] after low and back rounded vowels and before coronals, i.e., dentals, alveolars and palatoalveolars, but not before labials and velars (e.g., Piedmontese [awt] Italian *alto*, [ˈmalva] It. *malva*, [sulk] It. *solco*). In contrast with previous explanations, I hypothesize that the primary reason for this sound change is to be sought neither in the enhancement of the /l/ darkening degree before coronals nor in the develarization of the lateral consonant before labials and velars but in the relative prominence of the F2 (second formant) transitions connecting /l/ to the following consonant. Departing from the intermediate stage [VʊlC] of /VlC/ sequences emerged through glide insertion (which has been assumed to occur in Middle English), the glide was deleted when followed by labials and velars due to the flat and hardly perceptible /lC/ transitions in question. As to the /VlC/ sequences with coronals, the lateral was deleted through gestural merging with the following consonant after which [ʊ] stayed after a low vowel and often merged with the preceding back rounded vowel in accordance with differences in acoustic prominence of the F2 transitions connecting the vowel to the glide (Piedmontese [awt] It *alto*, [ˈvɔta] It. *volta*).

**Keywords**: dark /l/, /l/ vocalization and deletion, Old English, Romance languages

This research project investigates the positional and contextual conditions under which dark /l/ shifted to /w/ or a similar vocoid when occurring syllable-finally in Middle English and Early Romance, as exemplified by the Lengadocian Occitan forms [faws] derived from Latin FALSU and [ˈtawpo] TALPA. The replacement of dark /l/ by /w/ is associated with acoustic similarity, i.e., an increase in /l/ darkening degree causes the second formant (F2) to lower below 1000 Hz, which is the appropriate F2 frequency for back rounded vocoids. This spectral change

follows from an enhancement of the tongue body configuration for dark /l/ involving tongue blade and predorsum lowering in conjunction with postdorsum retraction, which may result in the formation of a dorsopharyngeal or dorsovelar constriction. However, acoustic equivalence does not seem to be a sufficient factor for dark /l/ vocalization to occur. Other articulatory mechanisms in addition to those just mentioned appear to be needed: (a) undershoot of the apical raising gesture causing tongue tip contact loss at the primary alveolar constriction location; (b) tongue tip lag involving the anticipation of the tongue dorsum lowering and retraction movement with respect to the tongue tip raising gesture (Sproat & Fujimura, 1993).

Production and perceived vocalization studies have shown that the likelihood that these articulatory characteristics contribute to the identification of dark /l/ as /w/ becomes most apparent in the following positional and contextual conditions:

(a) Syllable-finally where tongue tip contact loss may be considered an instance of gestural weakening, even in languages which are endowed with a dark allophone of /l/ not only syllable-finally but syllable-initially as well.

(b) After low and back rounded vowels (where /l/ may also be deleted without vocalizing) in so far as these vowels are similar to dark /l/ both articulatorily and acoustically.

(c) Before labials and velars rather than before coronals and thus dentals, alveolars and palatoalveolars (see for example Szalay, 2019), a reason for this contextual effect being that, while labials and velars do not interfere with apical contact loss during the alveolar lateral, coronals prevent it from occurring since they are homorganic with /l/. However, /l/ vocalization is expected to also take place in the latter context condition at least for two reasons. On the one hand, as shown by dorsopalatal contact and F2 data reported in earlier studies, /l/ is also dark when occuring before coronals (it has been found to be as dark as before velars though less dark than before labials). On the other hand, whenever gestural merging between /l/ and a following (homorganic) coronal results into listeners failing to hear the lateral, the /w/-like tongue dorsum configuration associated with the tongue tip lag factor is likely to remain as the sole cue of dark /l/. In that case /l/ vocalization will be accomplished not through a direct replacement of /l/ by /w/ but through the pathway /Vl/ > [Vʊ̯l] > [Vw].

Sound change data for syllable-final /l/ in Middle and Modern English (Jespersen, 1965) are to a large extent consistent with the production and perceived vocalization data summarized so far. Based on graphemic representations like *haulve* and *faulse* for *half* and *false*, scholars agree in that preconsonantal /l/ vocalization in English was achieved through an [Vʊ̯l] stage in /VlC/ sequences with low and mid back rounded vowels and labial, coronal and velar consonants alike. At a later date, the lateral dropped before labials and velars due apparently to articulatory reasons and the little acoustic salience of the F2 transitions connecting /l/ with the following consonant (i.e., the /lC/ transitions), which happen to be flat since labial and back velar consonants are endowed with a low F2 locus; finally, the resulting diphthong became a monophthong yielding [aː] for [aʊ̯l] (*psalm*, *half*) and [ɔː] for [oʊ̯l] (*chalk*, *folk*). As to the coronal context condition, two pathways must be differentiated. According to the more general pathway the lateral was preserved since it was homorganic with the following coronal consonant, the rising F2 transitions connecting the two consonants being quite prominent in this case since most coronals are endowed with a relatively high F2 locus; later, [aʊ̯l] monophtongized into [ɔːl] (*bald*, *false*) and [oʊ̯l] into [oːl], which later shifted to [əʊ̯l] (*gold*, *shoulder*). According to the second pathway, the sequence [aʊ̯l] dropped the lateral through merging with following /t/ yielding forms like [sɔʊ̯t] *salt*, which cooccurred with [saʊ̯lt]; afterwards, those lexical variants with the lateral prevailed upon those without it (and thus [saʊ̯lt] upon [sɔʊ̯t]) or else /l/ was reintroduced after the glide in those forms which had no lateral ([sɔʊ̯t] > [sɔʊlt]); finally, there was diphtong monophthongization into [ɔː] ([saʊ̯lt] > [sɔːlt]). The reason why the lateral could be deleted from [aʊ̯l] but not from [oʊ̯l] may be sought in the fact that the F2 transitions connecting the low vowel to the following glide (i.e, the [Vʊ̯]

transitions) were more acoustically prominent for [aʊ̯] than for [oʊ̯], which may have lead the speaker/listener to consider that dark /l/ was sufficiently cued by the glide.

These Old English facts may throw some light on an unexpected outcome of coda dark /l/ vocalization in several Romance language in the Middle Ages, namely, the substitution of /l/ by [w] before coronals but not before labials and velars throughout Italy (safe for Tuscan), Romansh and most of Ladin, Francoprovençal and Occitan. Thus, for example, Ligurian has [ˈawtu] It. *alto* ALTU and [ˈkawdo] It. *caldo* CALIDU but no /l/ vocalization in [ˈarba] It. *alba* ALBA and [karˈkaɲo] It. *calcagno* CALCANEU where *l* has been replaced by *r*. Other languages underwent coda /l/ vocalization before labials, coronals and velars (French and Occitan dialects such as Gascon) or else before labials and coronals (Portuguese, Old Spanish). In modern times, dark /l/ shifted to clear /l/ in Romance with a few exceptions (Portuguese, Catalan, Romansh dialects).

Two main explanatory hypotheses have been proposed in order to account for this scenario: /l/ was develarized before labials and velars through a dissimilatory process (Rohlfs, 1966: 342, 345-346); /l/ became darker before coronals in order to avoid develarization from taking place in this consonantal context condition (Straka, 1968). I hypothesize that /l/ vocalization before coronals but not before labials and velars may obey another motivation which is in some ways consistent with the Middle English sound change data summarized above. The point of departure is a [Vʊ̯lC] sequence structure, as strongly suggested by lexical forms like *ault* ALTU available in present-day Romansh, *olt* ALTU in Milanese (where *ol* derives most probably from *aul*) and *auldeia* for *aldea* in northern European Portuguese as well as by the graphemic sequence *aul* appearing in Old French, Old Francoprovençal and Old Venetian documents (e.g., Old Venetian *aultro* ALTERU). At a later date, the glide [ʊ̯] dropped before labials and velars owing to the hardly perceptible /lC/ transitions, while before coronals the lateral was deleted for the same reason that [saʊ̯lt] *salt* became [sɔʊ̯t] in Middle English. This would account for why Piedmontese has *aut* ALTU and *faus* FALSU buy *malva/marva* MALVA and *solc/sulc* SULCU. Moreover, differences in acoustic prominence of the [Vʊ̯] transitions may account for why the glide has remained unmodified after a low vowel (Piedmontese *aut*, *faus*) while often merging with a preceding back rounded vowel (Piedmontese *vòta* *VOLVITA, *dus* DULCE).

**References**

Jespersen, Otto (1965). *Modern English grammar on historical principles: sounds and spellings*. London/Copenhaguen: Allen & Unwin/Einar Munksgaard.

Rohlfs, Gerhard (1966). *Grammatica storica della lingua italiana e dei suoi dialetti*, vol. 1: Fonetica, Torino: Einaudi

Sproat, Richard & Fujimura, Osamu (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, 21(2), 291-311.

Straka, Georges (1968). Contribution à la description et à l'histoire des consonnes L. *Travaux de Linguistique et de Littérature, Université de Strasbourg*, 6(1), 267-326.

Szalay, Tünde Orsolya (2019). *Production and perception of lateral-final rimes in Australian English*. PhD Dissertation, Macquarie University.
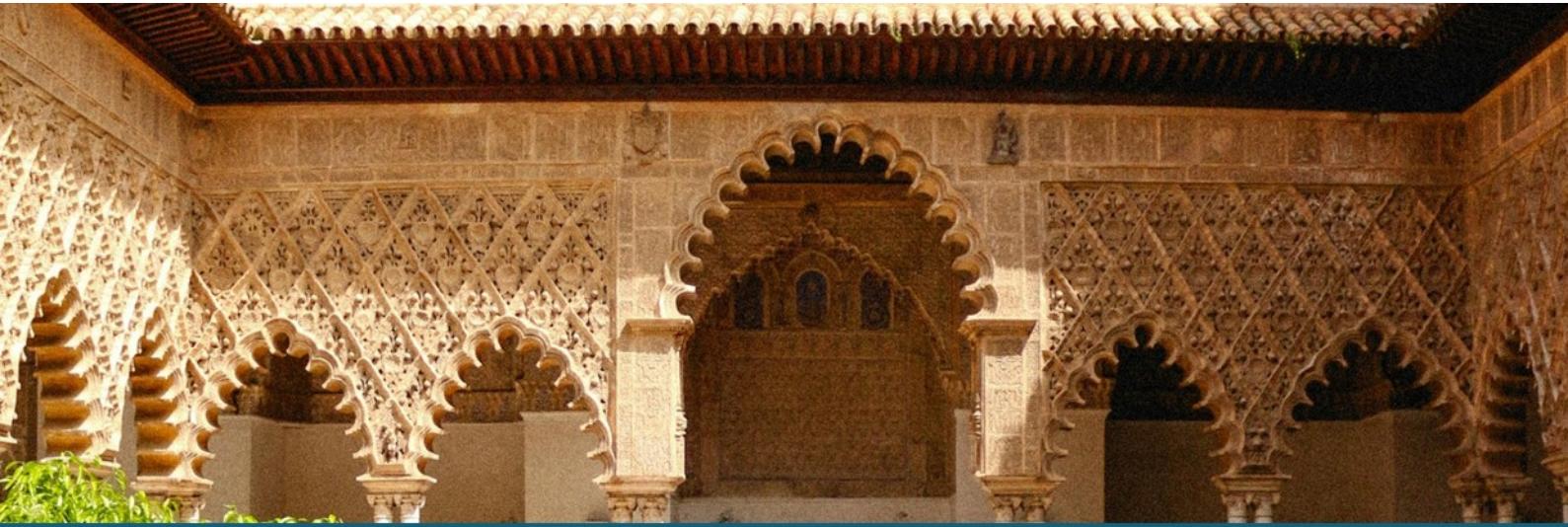
# PHONETIC VARIATION

# Variability in the articulation of Mandarin rhotic vowels

Song Jiang, Alexei Kochetov
*University of Toronto*
soong.jiang@mail.utoronto.ca, al.kochetov@utoronto.ca

**Abstract**
Individual and contextual variability in the choice of tongue shapes – tip-up or tip-down – has been well documented for the North American English /ɹ/ (including [ɚ]). In contrast, the situation with Mandarin rhotic vowels (e.g. [ɚ, ur]) is much less clear. This study explores individual and contextual variability of Mandarin rhotic vowels using ultrasound. Results show individual and, to a lesser extent, contextual variation in the realization of these segments. First, the speakers varied in using either a tip-up or a tip-down tongue configuration, although the former one was more common. Second, [ur] was less commonly produced with a raised and curled back tongue tip, suggesting that vowel quality may impose subtle articulatory constraints. These results demonstrate great variability in the articulation of Mandarin rhotic vowels.

**Keywords:** individual variability, articulation, rhotic vowels, ultrasound, PCA

## 1. Introduction and background

One of the most documented characteristics of the North American English rhotic /ɹ/ is its contextual and/or inter-speaker variability in the choice of tongue shapes – tip-down or tip-up (bunched or retroflex; Delattre & Freeman 1968, Mielke et al. 2016, among others). This variability has been explained by the lack of acoustic differences between the two constriction types, as well as lesser compatibility of tip-up shapes with high/front vowels and non-syllable-initial positions. In contrast, the situation with Mandarin rhotic vowels is much less clear. Some articulatory (EMA or ultrasound) studies of Beijing Mandarin rhotic vowels reported predominantly tip-up (retroflex) articulations (Xing 2022); others found no predominant configuration between tip-up and tip-down (bunched) tongue shapes (Chen et al. 2024). It has also been noted that tip-down tongue shapes are predominantly used in other Mandarin dialects (Northeast: Jiang et al. 2019, Southwest: Huang et al. 2024). To further explore the individual and vowel-quality-specific variability, a systematic ultrasound investigation of various vowel qualities in two Mandarin dialects, Beijing (BJM) and Northeast (NEM), was conducted.

## 2. Methodology

Audio-ultrasound synchronized data were collected from 10 BJM speakers (7F) and 5 NEM speakers (3F) using an EchoB system (Articulate Instruments Ltd.) with an UltraFit head stabilizer. The stimuli were comprised of meaningful words with the vowels [ə, u, ɑ] and their rhotic counterparts [ɚ, ur, ɑr] preceded by bilabial stops (e.g., [pu] 'no' - [pur] 'step-DIM'). Given that [ɑ] in BJM is not rhotacizable, only data from NEM speakers were included for this vowel. The target words were produced in the carrier phrase "__, [mɑ] __ [pɑ]" ("__, curse with the word __") 5 times. Tongue contours were traced using the DeepLabCut method within AAA. For each acoustically defined rhyme, 7 equally timed frames were extracted, converted to polar coordinates, and z-scored within-speaker. Principal Components Analysis (PCA) was employed to capture main characteristics and individual variation of rhotic articulation for each vowel.

## 3. Results

Tip-up, mid-bunched, and front-bunched tongue shapes were found for all three vowels, as shown in Fig 1. Six out of 10 BJM speakers and 3 out of 5 NEM speakers used tip-up

configurations across the board. One mid-bunching speaker and 1 front-bunching speaker were found in each dialectal group. Nonetheless, BJ05 did tip-up for [ʊ˞], but front-bunched for [ɚ], while BJ06 did tip-up for [ɚ], but mid-bunching for [ʊ˞]. Fig. 2 shows the first two PCs of the PCA model for [ɚ], [ʊ˞], and [ɑ˞]. PC1 for [ɚ] indicates that tongue shapes of [ɚ] primarily vary between dorsum-up and tip-up configurations. For [ʊ˞], tongue shapes vary predominantly along the PC1 involving the dorsum and front, with little evidence of a tip curling-up posture. PC1 for [ɑ˞]'s tongue shape shows variation in the tongue dorsum and front, while the PC2 involves tip curling-up posture. Fig. 3 illustrates how individual speakers' tongue shapes are distributed along the PC1/PC2 dimensions, which jointly distinguish the tip-up and tip-down configurations for [ɚ], [ʊ˞], and [ɑ˞]. Compared to mid-bunching speakers, front-bunching speakers showed higher PC1 scores for [ɚ] and [ʊ˞] and lower PC2 scores for [ɑ˞], aligning more closely with the patterns observed in tip-up speakers.

## 4. Discussion and conclusion

Both tip-up and tip-down tongue shapes were identified for [ɚ], [ʊ˞], and [ɑ˞], with some skew toward tip-up articulations. The results reflect between-speaker articulatory variability in Mandarin rhotic vowels, which can be explained by the lack of acoustic differences between the two types (cf. English). Additionally, the tendency for tip-up [ɚ] and [ɑ˞] to involve greater tip raising and curling back than tip-up [ʊ˞] suggests that vowel quality may impose subtle articulatory constraints, regardless of the tip configuration. Compared to English, however, rhotic tongue shapes in Mandarin appear to be much less sensitive to vowel quality, which suggests some language-particular differences in rhotic target specifications.
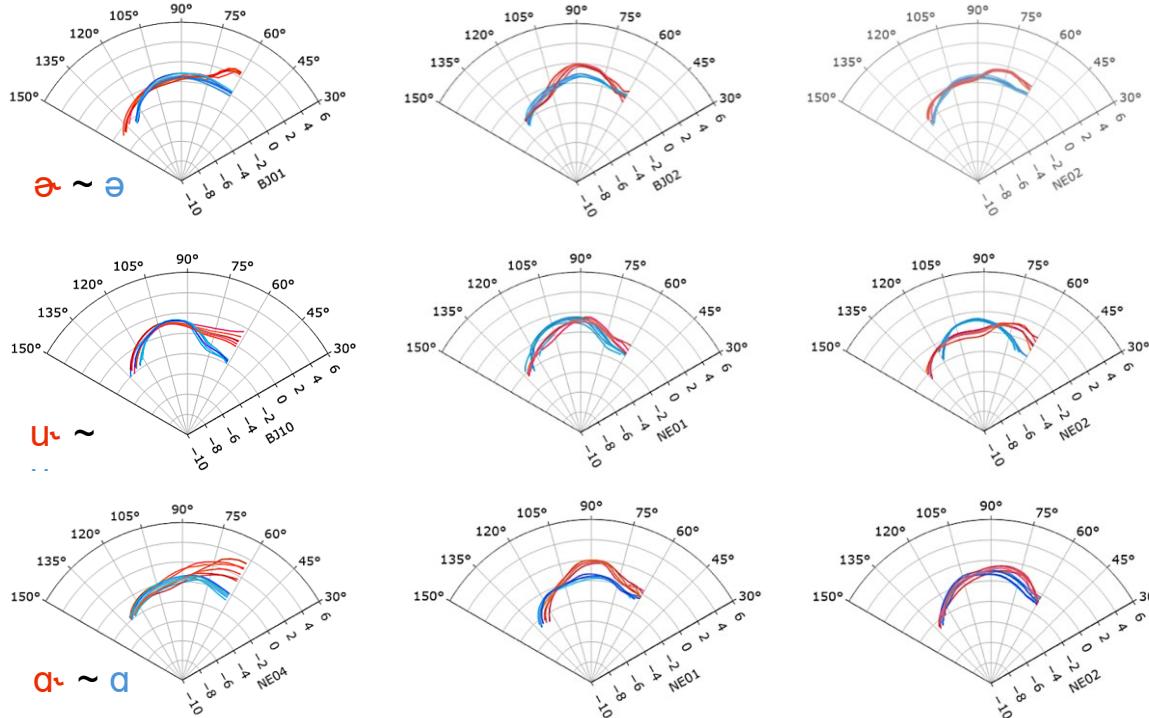


Figure 1 Examples of tip-up (left), mid-bunched (middle), and front-bunched (right) [ɚ, ʊ˞, ɑ˞]
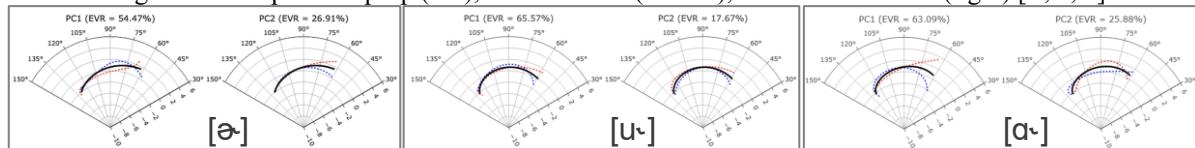


Figure 2 PC1 and PC2 for each vowel (black line: mean, red line: +2sd, blue line: -2sd)

22
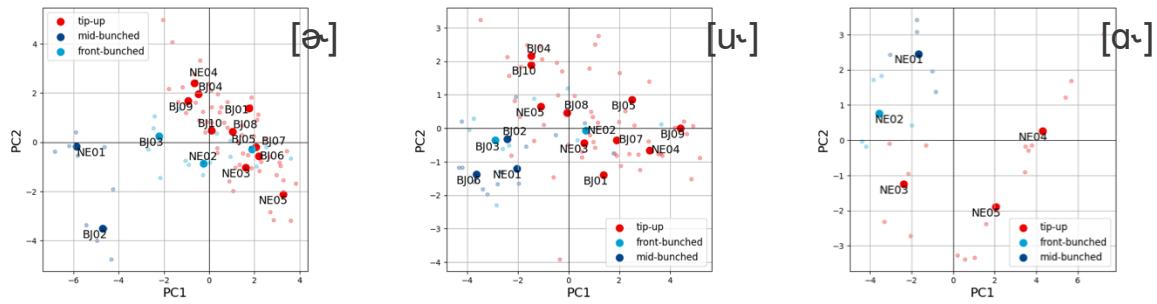
Figure 3　By-speaker (larger dots) and by-token (smaller dots) PC1 and PC2 scores of [ɚ, u˞, ɑ˞]

## 5. References

Chen, S., Whalen, D. H. & Mok, P. P. K. (2024). Production of the English /ɹ/ by Mandarin–English Bilingual Speakers. Language and Speech, 00238309241230895.

Delattre, P., & Freeman, D. C. 1968. A dialect study of American r's by x-ray motion picture. Linguistics, 6(44), 29–68.

Huang, J., Hsieh, F., Chang, Y. & Tiede, M. (2024). On the two rhotic schwas in Southwestern Mandarin: when homophony meets morphology in articulation. Phonetica, 81(1), 43-80.

Jiang, S., Chang, Y. & Hsieh, F. (2019). An EMA study of er-suffixation in Northeastern Mandarin monophthongs. In Proceedings of 19th international congress of phonetic sciences, Melbourne, Australia 2019. Canberra: Australasian Speech Science and Technology Association Inc.

Mielke, J., Baker, A. & Archangeli, D. (2016). Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. Language, 92(1), 101–140.

Xing, K. (2022). Phonetic and phonological perspectives on rhoticity in Mandarin. [Doctoral dissertation: The University of Manchester].

**Realization of /tʃ/ across Andalusia: Geographical distribution and acoustic properties**

Alfredo Herrero de Haro, Matthew Pollock
*Universidad de Granada*, *Louisiana State University Shreveport*
ahh@ugr.es, mpollock@lsus.edu

**Abstract**
Variability in the production of /tʃ/ in Andalusia (southern Spain) has long been documented, with [ʃ] and [tʃ] being common. Extending from an earlier dialectological survey collected in the 1950s, this paper draws from a new corpus, the ongoing Interactive Linguistic Atlas of Andalusian Accents (ALIAA), to reexamine how /tʃ/ production varies 70 years later. Acoustic analyses of 1,162 speakers from 571 locations across Andalusia, including measures such as VOT, segment duration, burst presence, and vowel formant transitions, are underway. A preliminary analysis from 69 locations shows a similar distribution pattern to that described in earlier dialectological work. However, the current analysis has identified six different realizations of /tʃ/, varying from a canonical affricate to a fricative. The fricative variant, previously identified as indexing lower prestige in Andalusian Spanish, has receded in parts of the region and more realization types can be distinguished now thanks to acoustic analyses. This study shows not only how the geographical distribution of the affricate has changed in the last seventy years, but it also contributes to our understanding of deaffrication processes in other languages.

**Keywords:** affricates, deaffrication, Andalusian Spanish, linguistic atlas

## 1. Introduction and background

Deaffrication of /tʃ/ has been widely reported across Andalusia, in southern Spain. In this region, production of /tʃ/ as [tʃ] or [ʃ] varies by geographical origin and, to a lesser extent, based on the sociocultural background of speakers. Alvar et al. (1973: map 1709) offer the most detailed analysis of the geographical distribution of /tʃ/, examining 230 locations. These authors report the highest rates of affricate production in eastern and northern Andalusia, while western Andalusian speakers preferred [ʃ]; it is important to note that Alvar et al. (1973) collected their data in the 1950s. More recent research has shown a change in the distribution of /tʃ/ realizations. Melguizo Moreno (2006, 2007) describes the fricative variant as most common in word-initial position and among older, working-class speakers in Granada, while [tʃ] use was higher among other groups. Alternately, Pollock (2024) found no instances of deaffrication among political speech across Andalusia, although he identified a fronted variant [ts] common among young and female politicians. Regional prestige norms also play a role in deaffrication across Andalusia, as described by Villena-Ponsoda (2013).

## 2. Methodology

The Interactive Linguistic Atlas of Andalusian Accents (ALIAA) is a project that aims to update and expand upon the analysis of Alvar et al. (1973). Over 493,000 audio files from more than 4,400 Andalusian speakers were collected, from which over 183,000 audios from 1,162 speakers in 571 locations across Andalusia were selected for the atlas. Data were collected using a mixture of self-guided online surveys and in-person interviews. The current study focuses on examining five words with /tʃ/ produced by all speakers: champú 'shampoo,' champús 'shampoos,' hecha 'made,' chapuz 'botched job,' and coche 'car.'
Acoustic measures in Praat (Boersma & Weenink, 2025), including VOT, segment duration, burst presence, and vowel formant transitions, were used to categorize realizations. Additionally, following an analysis of the Venezuelan affricate (Díaz-Campos et al. 2023), /tʃ/

was described not just categorically but also through a continuous measure, classifying productions based on a ratio of frication to stop portions of the segment.
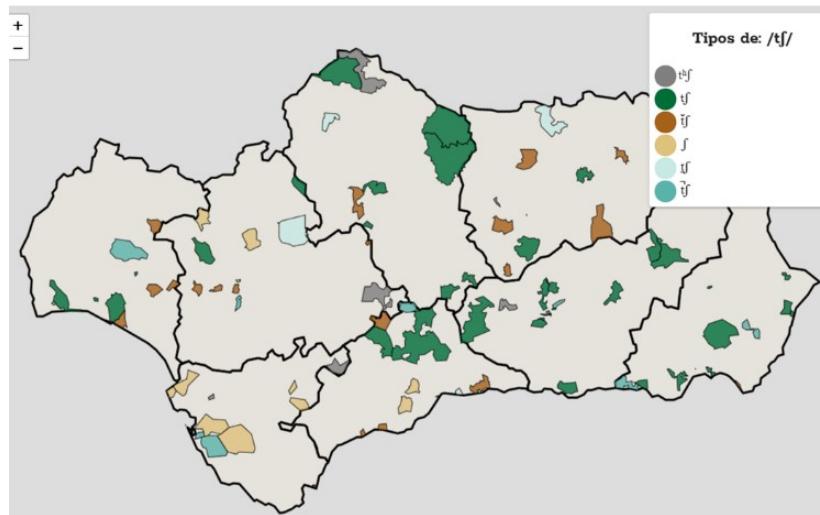


Figure 1. Distribution of /tʃ/ in 69 locations.
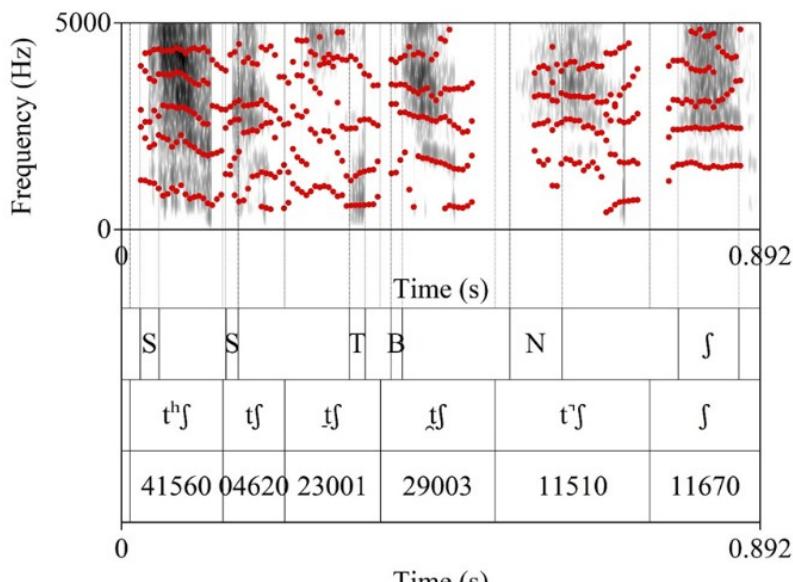


Figure 2. Realizations of /tʃ/.

## 3. Results

Preliminary analyses in 69 locations show a similar pattern of distribution to that of Alvar et al. (1973: map 1709) (Fig. 1). However, an area around central Andalusia that Alvar et al. described as having interchangeable [tʃ] and [ʃ] appears to produce intermediate pronunciations. Additionally, to drill down into variable production, the current analysis has identified six different realization types, varying from an aspirated [tʃ] to [ʃ] (Fig. 2)

## 4. Discussion and conclusion

This study shows how the geographical distribution of each realization of the affricate has changed in the last seventy years and identifies the range of realizations that exist for /tʃ/. Despite descriptions of a reduction in deaffrication across southern Spain, deaffrication still enjoys regular use across some areas of Andalusia. In addition to documenting variation, this study also contributes to our understanding of the broader process of deaffrication, which languages such as French have historically undergone. Our description of this change in Andalusia, similar to others elsewhere in the Spanish speaking world, conceptualize it as part of an ongoing, gradient process with numerous phonetic steps between complete affricate and fricative production. The gradual spread of change described in this article can, in turn, also be applied to historical analyses in other languages.

In Andalusia, the cross-linguistic phonetic pattern of evolution from [tʃ] to [ʃ] is challenged, stopped, and even reverted by Europe-wide tendencies of regional varieties to merge with national standards due to prestige. This has caused [ʃ] to surface as [tʃ] in some areas where [ʃ] was traditionally present and it calls for a reanalysis of the complex relationship between phonetic universals and sociolinguistic pressures.

## 5. References

Alvar, M., Llorente, A. & Salvador, G. (1973). Atlas lingüístico y etnográfico de Andalucía, vol. 6. Granada: Universidad de Granada/Consejo superior de investigaciones científicas.

Boersma, P. & Weenink, D. (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4.31, http://www.praat.org/. Accessed on 05/01/2025.

Díaz-Campos, M., Cole, M. & Pollock, M. (2023). Re-conceptualizing affricate variation in Caracas Spanish. Hispania, 106(1), 9–26.

Melguizo-Moreno, E. (2006). La fricatización de /č/ en una comunidad de hablantes granadina. Interlingüística 17, 748–757.

Melguizo-Moreno, E. (2007). Convergencia y divergencia dialectal: Estudio de la variación lingüística granadina. Lingüística Española Actual 29(2): 261–285.

Moya Corral, J. A. & García Wiedemann, E. (1995). La 'ch' fricativa en Granada: Un sonido del habla masculina. In A. Ward (ed.), Actas del XII Congreso de la Asociación Internacional de Hispanistas (270–283). Madrid: Centro Virtual Cervantes.

Pollock, M. (2024). Buenas no[tʃ]es y mu[ts]isimas gracias: Variable Affricate Production in Peninsular Spanish Political Discourse. Phonetics and Phonology of Ibero-Romance Languages: Languages, 9(6). DOI: https://doi.org/10.3390/languages9060218

Regan, B. (2020). El [ʃ]oquero: /tʃ/ Variation in Huelva Capital and Surrounding Towns. Estudios de Fonética Experimental, 29, pp. 55–90.

Villena-Ponsoda, J. A. (2008). Sociolinguistic patterns of Andalusian Spanish. International Journal of the Sociology of Language 193–194, 139–160.

Villena-Ponsoda, J. A. (2013). Actos de identidad: ¿Por qué persiste el uso de los rasgos lingüísticos de bajo prestigio social? Divergencia geográfica y social en el español urbano de Andalucía. In R. Guillén Stuil (ed.), Estudios Descriptivos y Aplicados Sobre el Andaluz (173–207). Seville: The University of Seville Press.

# Weakening of intervocalic /b/ and /v/ in conversational French: reduction or lenition?

Suyuan Dong, Nicolas Audibert
*Laboratoire de Phonétique et Phonologie (CNRS & Sorbonne Nouvelle)*
suyuan.dong@sorbonne-nouvelle.fr, nicolas.audibert@sorbonne-nouvelle.fr

## Abstract

This study investigates weakening of /b/ and /v/ in conversational French. Based on 2761 manually annotated intervocalic /b/ tokens from 10 female and 10 male speakers in the NCCFr corpus, the results reveal substantial variation, including 20% of approximant-like realizations with distinct acoustic profiles and shorter durations compared to canonical forms. A speaker-level correlation with /v/ weakening ($r = .68$) suggests shared dynamics beyond speech rate. Positional effects are limited across 3848 /b/ tokens annotated in all position for female speakers. These patterns, consistent with /v/ weakening data, suggest an incipient sound change in French.

**Keywords:** voiced labial obstruent, approximant, conversational French, lenition

## 1. Introduction and background

This study investigates the weakening of voiced labial obstruents /b/ and /v/ in conversational French. Phonetic variation is commonly viewed as a potential source of sound change, shaped by articulatory and perceptual constraints (e.g., Ohala, 1990; Blevins, 2004; Lindblom, 1990; Beddor, 2023). Lenition is reported as frequently observed in intervocalic contexts (De Carvalho, 2008, pp.208–212). One of the lenition trajectories for voiced obstruents — stop > fricative > approximant > zero (Trask, 2006, p.191; Crystal, 2008, p.274) — reflects a progressive relaxation of constriction related to manner of articulation (Kirchner, 2004, p.313; Broś et al., 2021). While this process is well documented in Romance languages (e.g., Simonet et al., 2012; Wayland et al., 2023), it remains underexplored in French despite its relative frequency (e.g., Wu & Adda-Decker, 2021).

## 2. Methodology

Building on a prior study of 5504 tokens of /v/ produced by 10 female (n=2970) and 10 male (n=2534) speakers, we apply the same methodology to 2761 /b/ tokens produced by the same speakers from the NCCFr corpus (Torreira et al., 2010), including 1572 occurrences from female and 1189 from male speakers. Manual annotation was based on spectrogram inspection to classify realization types and extract acoustic cues, including closure phase, burst, presence of the first four formants and energy in frequency bands between 5kHz and 15kHz.

In addition to duration, 20 acoustic parameters were extracted at 6 equidistant measurement points throughout the closure phase to distinguish canonical voiced stops from approximant-like realizations of /b/. These parameters include spectral moments, harmonic-to-noise ratio, zero crossing rate, relative intensity in different frequency bands, and measures of spectral and amplitude change in higher frequencies. To assess syllabic and lexical positional effects on the weakening process, we also manually annotated 2277 /b/ tokens produced by female speakers in non-intervocalic contexts. Analyses of male speakers are ongoing.

## 3. Results and discussion

The results reveal substantial variation in /b/ realizations: among 2044 effective tokens, 48% were realized as voiced stops, 20% as approximants (with no spectrographic evidence of burst and with stable formant structure), 7% as intermediate types (featuring both approximant-like characteristics and a detectable burst — see Figure 1), and 22% were nasalized realizations. By comparison, 35% of /v/ tokens were realized as approximants and 49% as voiced fricatives.

As with /v/, approximant-like realizations of /b/ show distinct acoustic profiles compared to canonical forms (Figure 2), with shorter durations and important overlap across categories regardless of speaker. Moreover, individual-level correlations between the proportion of weakened variants in /v/ and /b/ reveal a positive relationship ($r = .68$), especially among female speakers ($r = .93$), suggesting shared dynamics of variation beyond speech rate effects. For female productions across all positions (n=3848), /b/ is most realized in simple word-initial onset; positional effects on realization types remain limited.
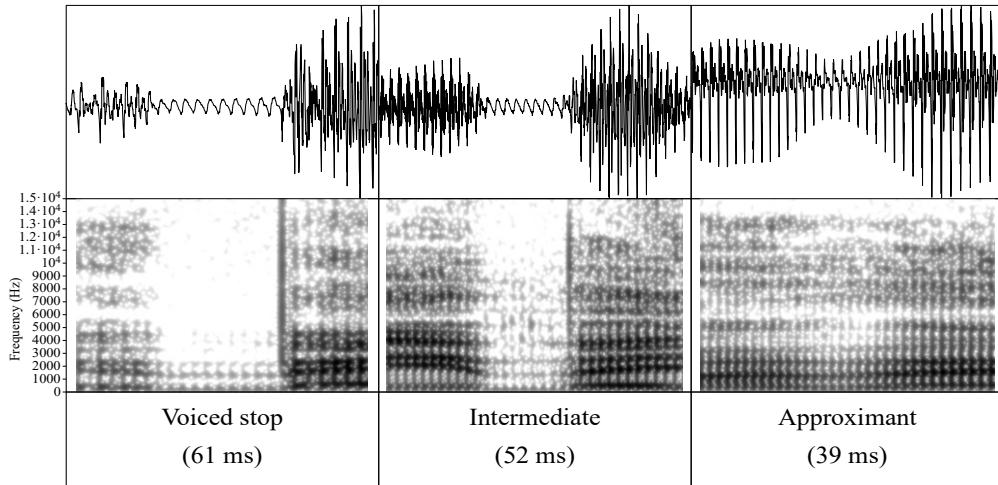


Figure 1: Examples of different /b/ realization types in the /ɛ_a/ context, produced by a female speaker. Consonant duration in parentheses.
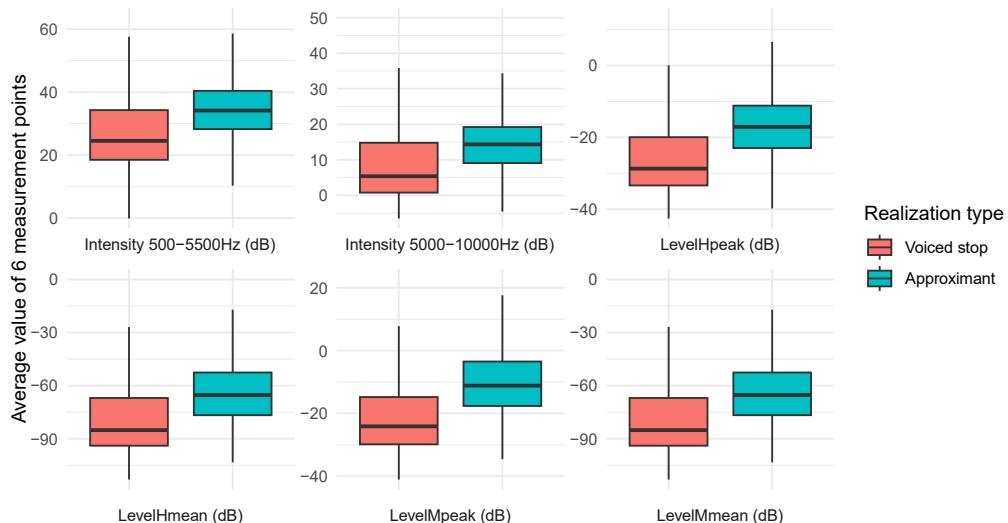


Figure 2: Acoustic measurements of intervocalic /b/ during the closure phase. LevelMpeak and LevelHpeak: spectral peak values in the 3–5.5kHz and 5–10kHz; LevelMmean and LevelHmean: average energy in these bands.

Spectrographic inspection of intermediate and approximant-like /b/ tokens, which maintain the same place of articulation as the canonical form, aligns with Duez (1995) and suggests that the burst may become a secondary perceptual cue compared to other articulatory gestures. These preliminary findings are consistent with the hypothesis of an incipient sound change.

## 4. References

Beddor, P. S. (2023). Advancements of phonetics in the 21st century: Theoretical and empirical issues in the phonetics of sound change. Journal of Phonetics, 98, 101-228. DOI: org/10.1016/j.wocn.2023.101228.

Blevins, J. (2004). Evolutionary Phonology: The Emergence of Sound Patterns. Cambridge University Press.

Broś, K., Żygis, M., Sikorski, A. & Wołłejko, J. (2021). Phonological contrasts and gradient effects in ongoing lenition in the Spanish of Gran Canaria. Phonology, 38(1), 1-40. DOI: 10.1017/S0952675721000038.

Crystal, D. (2008). A dictionary of linguistics and phonetics (6th edition). Blackwell Pub.

De Carvalho, J. B. (2008). Western romance. In J. B. De Carvalho, T. Sheer & P. Ségéral (Eds.). Lenition and fortition (pp. 207-233). De Gruyter Mouton.

Duez, D. (1995). On spontaneous French speech: aspects of the reduction and contextual assimilation of voiced stops. Journal of Phonetics, 23(4), 407-427. DOI: org/10.1006/jpho.1995.0031.

Kirchner, R. (2004). Consonant lenition. In: Hayes, B., Kirchnern, R. & Steriade, D. (eds.), Phonetically Based Phonology (1st edition), pp. 313-345. Cambridge University Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, W. J. & Marchal, A. (eds.), Speech Production and Speech Modelling, pp. 403-439. Springer Netherlands. DOI: 10.1007/978-94-009-2037-8_16.

Ohala, J. J. (1990b). The phonetics and phonology of aspects of assimilation. In Kingston, J. & Beckman, M. (eds.), Papers in laboratory phonology, vol. 1: Between the grammar and the physics of speech, pp. 258–75. Cambridge: Cambridge University Press.

Simonet, M., Hualde, J. I. & Nadeu, M. (2012). Lenition of /d/ in spontaneous Spanish and Catalan. 13th Annual Conference of the International Speech Communication Association (Interspeech 2012), 1414-1417. DOI: 10.21437/Interspeech.2012-50.

Trask, R. L. (2006). A dictionary of phonetics and phonology (Reprinted). Routledge.

Wayland, R., Tang, K., Wang, F., Vellozzi, S. & Sengupta, R. (2023). Quantitative acoustic versus deep learning metrics of Lenition. Languages 2023, 8(2), 98. DOI: 10.3390/languages8020098.

Wu, Y. & Adda-Decker, M. (2021). Réduction des segments en français spontané : apports des grands corpus et du traitement automatique de la parole. Corpus [En ligne], 22. DOI: 10.4000/corpus.5812.

# From Zurich to Valais: Phonetic Variation and its Impact on Forensic Automatic Speaker Recognition

Sophie Möller[1], Radek Skarnitzl[2], Andrea Fröhlich[3], Gea de Jong-Lendle[1] & Adrian Leemann[4]

[1] *Institut für Germanistische Sprachwissenschaft, Philipps-Universität Marburg, Germany*
[2] *Institute of Phonetics, Charles University, Czech Republic*
[3] *Speech & Audio unit, Zurich Forensic Science Institute, Switzerland*
[4] *Institute of Germanic Languages and Literatures, University of Bern, Switzerland*

Moeller9@students.uni-marburg.de, radek.skarnitzl@ff.cuni.cz, andrea.froehlich@for-zh.ch, dejong@staff.uni-marburg.de, adrian.leemann@unibe.ch

**Abstract**

Within the German-speaking world, Switzerland stands out for its sustained and socially prestigious use of dialects whose persistence, along with their considerable variation in phonetic, phonological, and lexical features, poses both opportunities and challenges for phonetic research and speech-processing systems. This study examines how dialectal variation within Swiss German affects the performance of automatic speaker recognition (ASR) systems. Although automatic speaker recognition systems are often considered language-independent, practical evaluations involving less widely spoken languages, such as Swiss German, or distinct dialect groups remain limited. In particular, the impact of dialect mismatches between reference datasets and test samples during likelihood estimation is not yet well understood, leaving many open questions. Understanding these effects is essential, particularly in applied contexts such as forensic phonetic casework. This practical relevance provided the motivation for the present study. Developed in collaboration with the Forensic Institute Zurich (FOR), this study aims to provide insight into reference data selection in contexts of low resource varieties such as Swiss German. Using speech data from the SDATS corpus (Swiss German Dialects Across Time and Space), ASR performance across four dialect regions (Zurich, Bern, Valais, Graubünden), based on recordings from 102 male speakers, was analysed. The dataset includes samples that vary in speaking style, degree of dialectality, and recording quality, closely resembling realistic forensic conditions. Preliminary results suggest good overall system performance, but subtle differences depending on the inclusion of standard Swiss German speech in the comparisons. Findings align with previous work and highlight the practical value of insights into phonetic variation when working with ASR systems.

**Keywords:** dialectal variation, Swiss German, phonetic diversity, ASR, reference data

## 1. Introduction and background

Language use in Switzerland represents a notable exception within the German-speaking world. While the standard variety is primarily reserved for formal and written contexts, dialects constitute the unmarked, default mode of communication in everyday life. They are highly prestigious, central to identity construction, and used with a matter-of-factness rarely observed in other German-speaking countries (Herrgen 2019: 246). This special form of diglossia results in few fluent transitions between dialect and standard varieties (Ammon et al. 2018: LXV), although exceptions in certain sociolinguistic contexts (Ruch et al. 2023) and due to language change (Leemann et al. 2024; Graf 2015) do exist. Unlike in Germany, there is no comparable shift toward regionally coloured standard varieties, and the Swiss dialect landscape remains relatively robust (Ruch et al. 2023).

Ultimately, this results in a high degree of phonetic diversity within Swiss German. Swiss German dialects can be grouped into three major Alemannic varietal groups, yet significant variation also exists along west–east and north–south axes. These varieties differ along several

phonetic dimensions, e.g. in vowel quality and quantity, consonant realizations, intonation and speech rate (Leemann et al. 2024; Ruch et al. 2023).

This diversity might pose a challenge for automatic speaker recognition (ASR) systems. In order to effectively apply ASR in forensic phonetic casework in German-speaking Switzerland—or in any low-resource dialect region—the interchangeability of reference corpora becomes a crucial issue. It must be assessed whether a general corpus of Swiss Standard German provides sufficient reference material, or whether dialect-matched corpora yield improved performance.

## 2. Methodology

This study examines the extent to which regional phonetic variation influences the performance of ASR systems. Focusing on four Swiss German dialect regions, it evaluates the effects of dialectal differences (dialect vs. standard), speaking style (read vs. spontaneous), and recording conditions (different types of background noise) that resemble forensic scenarios.

The speaker data stem from the SDATS corpus (Leemann et al., 2020) and include recordings from 102 male speakers aged 17 to 43, distributed across the four dialect regions (Zurich: n = 24, Bern: n = 32, Valais: n = 22, Graubünden: n = 24). While we refer to four regional groups, it should be noted that these categories are a simplification: dialect boundaries are dynamic and shaped by factors such as age, locality, education, and sociolinguistic context (Leemann et al., 2024; Steiner et al., 2023). For each speaker, three recordings were selected, representing distinct speech modalities: read speech in Swiss Standard German (~25 sec.), read speech in dialect (~2 min.) and spontaneous speech (~1 min.). Due to COVID-19 constraints, SDATS data were collected remotely via a supervised smartphone app, leading to variable audio quality resembling forensic conditions. To account for these factors, a noise classification scheme based on Jessen (2012) was developed and applied during annotation. The categorized noise types and their distribution will be presented at the conference as part of a broader diagnostic assessment of potential confounding variables.

Testing was conducted using the VOCALISE 2021, which is a state-of-the-art software program for performing speaker comparisons using Deep Neural Network (DNN) based x-vectors. Performance evaluation was done with the complementary software Bio-Metrics 2024. The values reported correspond to standard performance metrics commonly used in ASR evaluation, namely equal error rate (EER, the value when false rejection and false acceptance rate are identical) and the log-likelihood-ratio cost (Cllr, which reflects the accuracy of an ASR recognition system). Two testing approaches were used. First, the performance of the automatic speaker recognition system Vocalise was evaluated quantitatively using EER and Cllr metrics. Second, the impact of different reference databases on the resulting likelihood ratios (LRs, not log-transformed) was examined qualitatively.

## 3. Preliminary results

Initial results after cross-validation indicate strong baseline ASR performance, with EERs between 0.4% and 1.3% (convex hull) and Cllr values from 0.02 to 0.09. Slightly lower values (indicating even better performance) were observed when Swiss Standard language samples were excluded from the reference data, possibly reflecting effects of the diglossia in German speaking Switzerland—though this requires further investigation. Initial 1-vs.-1 comparisons using varying reference sets also produced very high likelihood ratios. Taken together, these preliminary findings suggest that the ASR system performs well on dialectal speech data, in line with previous findings (Watt et al. 2020). The limited number of speakers meant that calibration aspects could only be addressed to a restricted extent. Nevertheless, observations of outlier speakers suggest that calibration should be a key consideration in future studies. Further analyses on how dialectal variation affects ASR outcomes across different dialect groups and

speech modalities will be presented, focusing on additional influencing factors such as background noise and speaking style.

## 4. References

Ammon, Ulrich, Bickel, Hans, Lenz, Alexandra N. & Fink, Juliane (eds.). (2018). Variantenwörterbuch des Deutschen. Die Standardsprache in Österreich, der Schweiz, Deutschland, Liechtenstein, Luxemburg, Ostbelgien und Südtirol sowie Rumänien, Namibia und Mennonitensiedlungen. 2nd completely revised and expanded edition. Berlin/Boston: De Gruyter Mouton.

Graf, Martin Hannes (2015). Aktuelle Tendenzen des Sprachwandels im Schweizerdeutschen. Schriebe und Schwetze im Dialekt. Die Sprachsituation in der heutigen Deutschschweiz (11–19).

Herrgen, Joachim (2019). Deutsch. Sprache und Raum – ein internationales Handbuch der Sprachvariation. Berlin/Boston: De Gruyter Inc. Accessed on 29/05/2025. https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5157606.

Jessen, Michael (2012). Phonetische und linguistische Prinzipien des forensischen Stimmenvergleichs. München: LINCOM Europa

Leemann, Adrian, Steiner, Carina, Studerus, Melanie, Oberholzer, Linus, Jeszenszky, Péter, Tomaschek, Fabian & Kistler, Simon (2024). Dialäktatlas. 1950 bis heute. Zollikon: vdf Hochschulverlag AG.

Leemann, Adrian, Studerus, Melanie, Messerli, Jan, Jeszenszky, Péter & Steiner, Carina (2020). SDATS Corpus – Swiss German Dialects Across Time and Space. With contributions from Open Science Framework. https://doi.org/10.17605/OSF.IO/S9Z4Q Accessed on 31/05/2025.

Ruch, Hanna, Fröhlich, Andrea & Lim, Sarah (2023). Grosse sprachliche Vielfalt auf kleinem Raum. Chancen und Herausforderungen für die forensische Phonetik in der Schweiz. Kriminalistik, 4, 236–244. Accessed on 29/05/2025. https://www-wiso-net-de.ezproxy.ub.unimarburg.de/document/KRIM__69de2f72bb7f4241a1d17455f373dcee 6eeafcf1.

Steiner, Carina, Studer-Joho, Dieter, Lanthemann, Corinne, Büchler, Andrin & Leemann, Adrian (2023). Sociodemographic and attitudinal effects on dialect speakers' articulation of the standard language: Evidence from German-speaking Switzerland. In Proceedings of Interspeech 2023 (3542–3546).

Watt, Dominic, Harrison, Philip, Hughes, Vincent, French, Peter, Llamas, Carmen, Braun, Almut & Robertson, Duncan (2020). Assessing the effects of accent-mismatched reference population databases on the performance of an automatic speaker recognition system. The International Journal of Speech, Language and the Law, 27(1), 1–34.

# Phonetic variation in Andalusian Spanish: The first maps of the Interactive Linguistic Atlas of Andalusian Accents (ALIAA)

Alfredo Herrero de Haro
*Universidad de Granada*
ahh@ugr.es

**Abstract**

Andalusian Spanish pronunciation varies from Central and Northern Spanish (e.g. coda lenition or deletion, /s/ and /θ/ merging). The most detailed account of geographical distribution of phonetic features in Andalusian was published as a linguistic atlas in 1973 with data collected in the 1950s in 230 points; however, recent studies have shown the need to update this atlas. This presentation will discuss the first maps of the Interactive Linguistic Atlas of Andalusian Spanish (ALIAA), a project whose aim is to update and expand the 1973 atlas, describing phonetic variation in more than 500 points. A total of 493,433 audios from 4469 speakers of Andalusian Spanish have been collected through online and in-person surveys; 138,171 audios from 1162 speakers from 571 points have been selected for the atlas. The Webmaus force alignment tool was used to segment audios and the textgrids were corrected manually. Praat scripts have been used to extract acoustic information. ALIAA is expected to be completed in December 2026 and to have 292 maps. The 27 maps created so far show similar distribution patterns to those from the 1973 atlas, but some features of Andalusian Spanish have receded in some areas. Some features from Eastern Andalusian Spanish (e.g. vowel lowering) seem to have extended to the west. Apart from describing phonetic variation in Andalusia, ALIAA will show how the geographical distribution of different phonetic features has changed since the 1950s. Many of the features analysed in ALIAA will help us understand phonetic processes in other languages (e.g. deaffrication in French). The interactive tool created to display ALIAA can be used to display linguistic and non-linguistic information anywhere in the world and the methodology developed for ALIAA has shown that it is possible to reduce the time and cost of producing linguistic atlases when compared with traditional research methods.

**Keywords:** linguistic atlas, phonetic variation, Andalusian Spanish

## 1. Introduction and background

Andalusian Spanish has been widely described in the literature and its most salient features are coda lenition or deletion, gemination following consonant lenition, vowel lowering before underlying consonants, right-to-left vowel harmony, deaffrication of /tʃ/, pronouncing /x/ as [h], and merging /s/ and /θ/ into [θ] or into [s]. Although several studies analyse specific features of Andalusian Spanish, the most complete description of the geographical distribution of phonetic variation across Andalusia is Alvar et al. (1973), which describes phonetic variation across 230 points of Andalusia; it should be noted that the materials for Alvar et al. (1973) were collected between 1953 and 1958. More recent research has shown that the distribution of the phonetic features described in Alvar et al. (1973) has changed and that the descriptions in Alvar et al. (1973) need to be updated. The Interactive Linguistic Atlas of Andalusian Accents (ALIAA) is a project whose main aim is to update and expand the descriptions of Alvar et al. (1973) by describing phonetic variation across 571 points in Andalusia.

## 2. Methodology

A list was created with the phonetic phenomena described in Alvar et al. (1973) and with other phonetic phenomena discovered since Alvar et al. (1973), such as morpheme dislocation of plural morphemes (Herrero de Haro 2020). After this, a survey containing 242 words was designed to cover all those phenomena. The site www.phonic.ai was used for participants to complete a self-guided online survey. After 10 months, the areas for which insufficient data

had been collected online were identified and data were gathered for those areas through in-person interviews. A total of 493,433 audio files from 4469 speakers were collected, and 138,171 files from 1162 speakers across 571 points from Andalusia were selected for the atlas; some audios were discarded because speakers had lived outside their postcode of origin. The Webmaus force alignment tool was used to generate automatic segmentation of the audios and the textgrids were manually corrected. Scripts were used to analyse the audio on Praat (Boersma & Weenink 2024) and extract various measures from the audios. The results for each phenomena were mapped using a tool created for this atlas; the tool shows the space covered by a specific postcode (Figures 1 and 2), it displays linguistic information for each postcode (e.g. transcription of a word), it shows additional comments (e.g. males say [tʃ]) and, by clicking on it, users can hear an example of a phenomenon pronounced by a local speaker (e.g. /x/ being pronounced [h]).

## 3. Results

ALIAA is expected to have 292 maps and, as of May 2025, 27 maps have been created. Two maps are shown below as examples, but more maps will be discussed in the presentation.
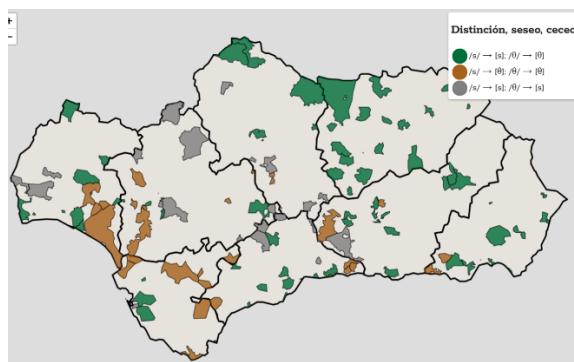


Figure 1. Realisation of /s/ and /θ/ as [θ] in brown, /s/ and /θ/ as [s] in grey, and /s/ as [s] and /θ/ as [θ] in green. 132 points covered.
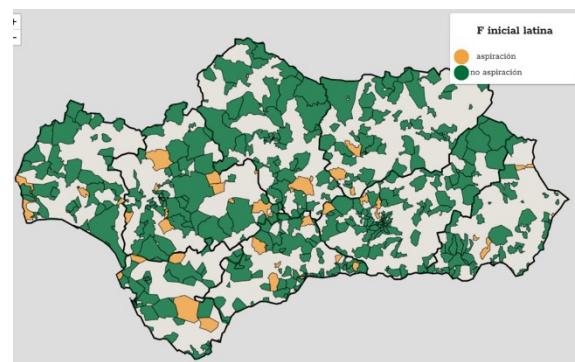


Figure 2. Pronunciation of humo "smoke" as [ˈumo] in green and as [ˈhumo] in orange. 545 points covered.

## 4. Discussion and conclusion

Figure 1 explains the geographical distribution of the /s/-/θ/ merger; ALIAA will analyse this in 571 points, but this map only includes 132 points so far. Figure 1 shows a very similar pattern to Alvar et al. (1973: map 1705), which analysed the same phenomenon in 230 points with data collected in the 1950s: the /s/-/θ/ distinction is maintained in the east and north, it merges into /θ/ in the south and west, and it mergers into /s/ in the centre, northwest, and west. However, there are also some differences. Some studies (Regan 2017) explain how a demerger is occurring in some parts of western Andalusia and the distinction /s/ - /θ/ is being reintroduced; this can be seen in Figure 1: the green sections in western Andalusia.

There is evidence that word-initial /f/ derived from Latin words started being pronounced [h] in the 9th century in Castilian Spanish (Jodl 2020). Towards the middle of the 16th century, this aspiration was receding (Penny 2003: 45), and it was eventually lost in Standard Central and Northern Spanish, but it has survived in some parts of Andalusia. Alvar et al. (1973: map 1715) show in which areas of Andalusia it was maintained in the 1950s and Figure 2 shows where it is maintained in 2025. Although the area of distribution coincides between both atlases, Figure 2 shows that [h] has disappeared in many parts of Andalusia since the 1950s.

Both maps show a move from more geographically marked phonetic features to a national standard; this pattern is found all across Europe (Regan 2017).

The Interactive Linguistic Atlas of Andalusian Accents can show how phonetic features are distributed across Andalusia, how these features have changed since the 1950s, and how these changes can explain phonetic variation in other languages. Furthermore, the methodology developed for ALIAA has proven effective in reducing time and costs of creating linguistic atlases when compared with traditional research methods.

## 5. References

Alvar, Manuel, Llorente, Antonio & Salvador, Gregorio (1973). Atlas lingüístico y etnográfico de Andalucía (Vol. 6). Granada: Universidad de Granada/Consejo superior de investigaciones científicas.

Boersma, Peter & Weenink, David (2024). Praat: Doing phonetics by computer. Version 6.4.19. https://www.fon.hum.uva.nl/praat/ Accessed on 10/09/2024

Herrero de Haro, Alfredo (2020). Morpheme dislocation in Eastern Andalusian Spanish plurals. Lingua, 238.

Jodl, Frank (2020). Fisiología y cognición –el cambio fonético y las teorías acerca del cambio lingüístico más recientes: el cambio [f]–[h]– en castellano antiguo. Revista de Filología Románica, 37, 81-98.

Penny, Ralph (2003). Variation and Change in Spanish. Cambridge: Cambridge University Press.

Regan, Brendan (2017). A study of ceceo variation in Western Andalusia (Huelva). Studies in Hispanic and Lusophone linguistics, 10, 119-60.

# Redrawing the Map of Regionally Accented Standard Dutch: Fine-tuning Whisper for Inclusive ASR

Golshid Shekoufandeh, Paul Boersma, Antal van den Bosch
*University of Amsterdam, University of Amsterdam, Utrecht University*
g.shekoufandeh@uva.nl, paul.boersma@uva.nl, a.p.j.vandenbosch@uu.nl

**Abstract**

In this study we investigate how fine-tuning the Whisper large-v2 automatic speech recognition (ASR) model on specific regional accents affects its accuracy in recognizing other accents within Standard Dutch (SD). Using the Spoken Dutch Corpus (CGN), which includes approximately 900 hours of speech from all regions in the Netherlands and Belgium, we examine accent-specific and comprehensive fine-tuning. Results demonstrate substantial improvements in Word Error Rates (WER), especially when trained on phonetically similar accents. Training on regions with linguistic centrality, such as East Utrecht & Veluwe, significantly enhances recognition across various accents. Conversely, more linguistically distinct areas like Friesland and Limburg benefit from specialized fine-tuning. A multidimensional scaling (MDS) visualization reveals intricate relationships between geographic proximity and phonetic similarity, highlighting central accents as optimal training data for generalization. The study emphasizes the importance of targeted fine-tuning strategies to improve ASR inclusivity and accuracy, advocating for broader and more representative datasets to mitigate biases.

**Keywords:** speech recognition, regional accents, linguistic centrality, fine-tuning, Standard Dutch

## 1. Introduction and background

Regional accents significantly impact the accuracy and inclusivity of automatic speech recognition (ASR) systems. Effective recognition of diverse accents is crucial, particularly as ASR becomes more integrated into daily life and diverse linguistic communities. As Feng et al. (2021) explain, bias in speech recognition systems often arises due to inadequate representation of diverse linguistic variations.

## 2. Methodology

We base this study on the Spoken Dutch Corpus (CGN), which was developed by Oostdijk (2000) and includes speech data from Dutch and Flemish speakers. The dataset was divided into regionally accent groups based on geographic and phonetic intuition. Fifteen speech components, including read speech, interviews, and dialogues, were incorporated to maximize data diversity. Transcriptions in .ort format were converted to plain text. Audio files longer than 30 seconds were segmented using sentence boundaries, adhering to Whisper's input constraints. To control for speaker overlap, training, evaluation, and test splits were made based on unique speaker IDs. In the first experimental phase, the Whisper large-v2 model was fine-tuned separately on each regional accent group for five epochs to balance performance and training efficiency. Each fine-tuned model was then tested on all regional subsets. In the second phase, a model was fine-tuned on the complete CGN dataset and evaluated similarly across all regional groups. Word Error Rates (WERs) were measured and compared with zero-shot performance.

To analyze accent proximity, odds ratios were calculated by comparing the WERs of fine-tuned and non-fine-tuned models. These values were inverted to construct a dissimilarity matrix, which was then visualized using monotone multidimensional scaling (MDS) in Praat.

## 3. Results

Accent-specific fine-tuning substantially reduced WER, notably enhancing performance across phonetically similar regions. The greatest improvements occurred in cases such as fine-tuning on the Holland & West Friesland & West Utrecht accent, which drastically improved recognition of the Friesland accent, achieving a 62.9% WER reduction. Figure 1 below illustrates these findings with a heatmap of odds ratios, showing accent similarity across regions (more blue is more enhancement).
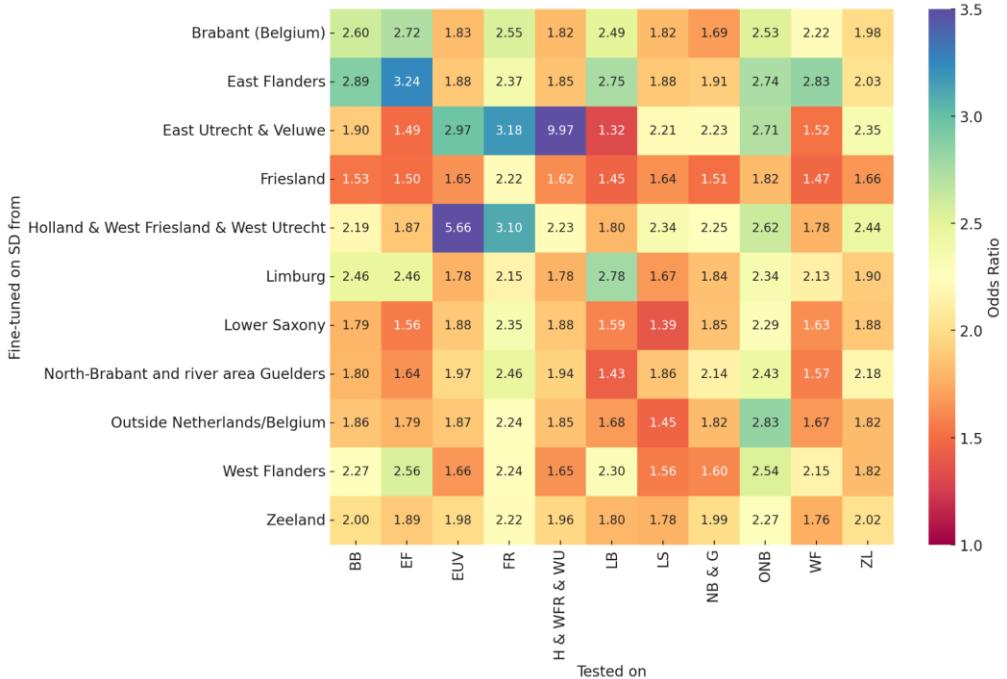


Comprehensive fine-tuning further improved results, significantly lowering WER across all accents. Zeeland showed the most improvement, with an 87.4% relative WER reduction. A multidimensional scaling (MDS) visualization (Figure 2) clarifies these complex relationships between phonetic similarity and geographic proximity, confirming certain central accents as optimal for generalized model training (ONB=Outside the Netherlands and Belgium).
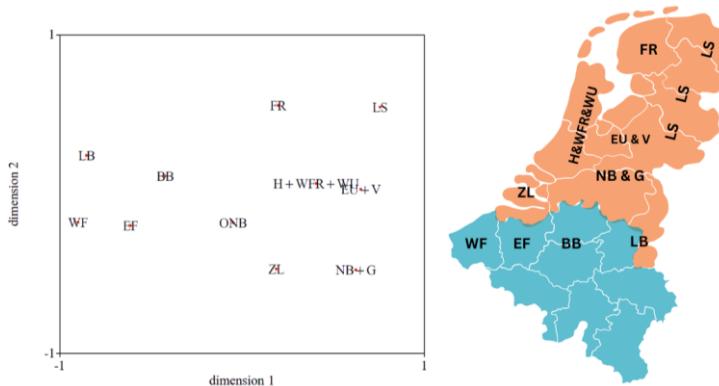


## 4. Discussion and conclusion

The results suggest that the region East Utrecht & Veluwe is phonetically central, facilitating generalized ASR performance to most of the other regions. In contrast, areas such as Friesland, with distinct phonetic features influenced by Frisian, require targeted fine-tuning and give the best results for their own regional accent, as Ten Bosch (2000) notes in his exploration of phonological distance. Future research should address dataset limitations by considering

speaker mobility and ensuring balanced regional representation, ultimately advancing inclusive ASR technology, as emphasized by Herygers et al. (2023).

## 5. References

Feng, S., Kudina, O., Halpern, B.M. & Scharenborg, O. (2021). Quantifying bias in automatic speech recognition. arXiv preprint arXiv:2103.15122.

Herygers, A., Verkhodanova, V., Coler, M., Scharenborg, O. and Georges, M. (2023). Bias in Flemish automatic speech recognition. Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung, pp.158-165.

Oostdijk, N. (2000). May. The Spoken Dutch Corpus. Overview and First Evaluation. In LREC (pp. 887-894).

ten Bosch, L. (2000). ASR, dialects, and acoustic/phonological distances. In INTERSPEECH (pp. 1009-1012).

# Regional phonetic variation: analysis of Murcian /-s/

Ana Bravo[1], Gemma Meseguer[2]

[1]*University of Murcia,* [2]*HiTZ Basque Center for Language Technology - Aholab Signal Processing Laboratory, UPV/EHU*

ana.bravo@um.es, gmeseguer001@ikasle.ehu.eus

**Abstract**

This talk aims to provide a qualitative description of the different realisations of implosive /-s/ in Murcian, an understudied phenomenon: retention [s], aspiration [h], elision [Ø], and marked elision [Ø |]. The analysis is based on spoken samples of 20 speakers from different areas of Murcia extracted from radio episodes that will be later analyzed and classified with the Audacity and Praat programs. Phonological processes affecting Murcian /-s/ are traditionally made dependent on phonological processes affecting Oriental Andalusian /-s/. Nevertheless, our study shows the singularity of the Murcian language in the phonetic and phonological plane, confirming that it deserves being studied in its own right. Namely, the analysis concludes that marked elision [Ø |] is almost totally dominant in most of the contexts. Secondly, we show how marked elision [Ø |] interacts with phonological rules of syllabification, which had not been taken into account until now in the study of Murcian Spanish. This way we explain the the resistance of Murcian Spanish to the CV rule. Finally, we address the consequences that the marked elision [Ø |] has for the morphological system.

**Keywords:** Murcian, /-s/ implosive, /-s/ loss, opacity, marked deletion

## 1. Introduction and background

Contemporary Hispanic dialectology has studied Murcian as a transitional dialect (Muñoz Garrigós, 1994: 317), whose phonological phenomena have always been justified by the influence of Andalusian. According to dialectologists, the phonetic link between Murcian and Andalusian dialects has resulted in an exchange of features, notably the weakening of the implosive /s/ (ibid.: 320). Based on this assertion, it is understandable that, in a peninsular context, phenomena related to /-s/ have been analyzed exhaustively with reference only to the Andalusian variety: Manuel Alvar (1994). In RAE-ASALE (2011: § 5.6i) it is mentioned in passing that the loss of /-s/ is frequently in Murcian, while Lloret & Martínez-Paricio (2025: 651-656) and Martínez-Gil (2025: 210-215) concentrate exclusively in Oriental Andalusian with no references at all to Spanish Murcian nor to either the realisations of /-s/ or to the compensatory lengthening in the vowels.

As for works dedicated specifically to the study of Murcian, only the articles by Hernández-Campoy (2002, 2015) are available. His work represents an advance over the other dialectology studies reviewed, as it analyzes several situations that occur in Murcian related to the elision of /s/ in implosive position. One of these is the gemination of the consonant following the lost /-s/, or the specific behavior of voiced plosive phonemes, since the loss of /-s/ favors the maintenance of the plosive realization (Rafael Monroy and Hernández-Campoy 2015: 232). He also comments on the resulting eight-vowel system due to the loss of the fricative, and its possible diachronic explanation (ibid.: 234); however, even in these studies, the phenomenon of /-s/ loss is too limited, so a broader description of it is necessary.

## 2. Methodology

Materials were collected online using episodes of La radio al sol, a Cadena Ser radio program that interviews locals from the towns and villages they visit, all of which are in the Region of Murcia. For this purpose, episodes recorded in Jumilla, Lorca, Mazarrón, Cieza, and Mula were selected. In accordance with Gómez Ortín's (2004) division of the Murcian dialect domain, this

selection covers four of the five areas established by the author: the central area (Cieza and Mula); the coastal area (Mazarrón); the Guadalentín area (Lorca) and the Altiplano area (Jumilla). Speakers were selected using random sampling. Four speakers were randomly selected from among all those interviewed in each of the five episodes mentioned above.

Approximately two minutes of each audio fragment were extracted for data collection. The audio files were edited and manipulated using Audacity software. To organize the information, the frequency of the different realizations of /-s/ recorded for each speaker in these four possible contexts was taken into account: when followed by a consonant, [_C]; when followed by a consonant across the word boundary, [_#C]; when followed by a vowel across the word boundary, [_#V]; and in absolute final position [_##].

## 3. Results

Contexts

|  | _C | | _#C | | _#V | | _## | |
|---|---|---|---|---|---|---|---|---|
|  | Núm. | % | Núm. | % | Núm. | % | Núm. | % |
| [s] | 3 | 5% | 0 | 0 | 18 | 50% | 10 | 42% |
| [h] | 0 | 0 | 0 | 0 | 1 | 3% | 0 | 0 |
| [Ø] | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 58% |
| [Ø |] | 62 | 95% | 153 | 100% | 17 | 47% | 0 | 0 |

Table 1. Provisional results

## 4. Discussion and conclusion

Acoustic analysis of the occurrences with Praat confirm that the variant that has occurred most frequently is marked elision [Ø |]. Only the context [_#V] behaves differently, and up to three variants have been recorded. The total number of times that /-s/ has been preserved as [s] in this context, as reflected in Table 1, is noteworthy, as it is the only context in which a higher percentage has been obtained than the marked elision variant [Ø |]: 50% versus 47%, respectively. These results confirm findings in previous studies (Erker 2010 and references cited therein), namely, that in those varieties in which the loss of /-s/ occurs, it is more likely to be retained in a context in which it precedes a vowel across the word boundary [_#V] than when the segment that follows it is a consonant [_#C]. Finally, the resistance of Murcian to the CV rule or Attack Rule (Colina 2016) when the /-s/ is lost (Rafael Monroy y Hernández-Campoy 2015) in the context VC#V reveals that Murcian and Central Castillian applied the syllabification rule and the Deletion Rule in coda in reversal orders, a proposal that had not been taken into account until now in the study of Murcian. Regarding vowel opening, although it was demonstrated in the theories presented that this phenomenon always occurs as a result of the elision of /-s/, in this study it was identified more clearly when the context was that of absolute final position.

## 5. References

Alvar, M. (1996). Manual de Dialectología hispánica: El español de España. Barcelona: Ariel.

Boersma, Paul & Weenink, David (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.09, recuperado el 2 de marzo de 2023, http://www.praat.org/

Broś, K. (2012). "Chilean Spanish /s/ Weakening as an Example of Phonological Opacity". Generative Linguistics in Wroclaw, 2, pp. 192-209

Colina, S. (2016). "Sílaba". Enciclopedia de lingüística hispánica. Routledge Taylor & Francis Group, Londres y Nueva York, pp. 25-35

Cutillas, J. A. (2006). Universalidad y especificidad de las restricciones fonológicas: Acento y fonotaxis en inglés [Tesis de Doctorado, Universidad de Murcia]. Recuperado de https://digitum.um.es/digitum/bitstream/10201/181/1/CutillasEspinos.pdf

Erker, D. G. (2010). "A subsegmental approach to coda /s/ weakening in Dominican Spanish". Int'l. J. Soc. Lang. 203, pp. 9-26 doi: 10.1515/IJSL.2010.019

Gómez Ortín, F (2004). "El dialecto murciano y sus variedades". Tonos digital: Revista de estudios filológicos, 8, pp. 7-26. https://www.um.es/tonosdigital/znum8/portada/monotonos/03-GORTIN.pdf 49

Real Academia Española y Asociación de Academias de la Lengua Española (2009-2011): Nueva gramática de la lengua española. Fonética. Fonología. Madrid: Espasa, 3 vol. (NGLE).

Hernández-Campoy, J. M. & Monroy, R. (2015). "Murcian Spanish". Journal of the International Phonetic Association, 45, pp 229-240 doi: 10.1017/S0025100314000231

Hualde, J. I. (2013). "Consonantes y vocales". En Los sonidos del español: Spanish Language edition (pp. 27-43). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511719943.005

Kiparsky, P. (1968). "Linguistic universals and language change". En: Bach, E. y Harms, R.T. (eds.) Universals in linguistic theory. Nueva York: Holt, Rinehart y Winston.

Kiparsky, Paul (1973). "Abstractness, opacity and global rules (Part 2 of 'Phonological representations')". Fujimura, Osamu (Ed.). Three Dimensions of Linguistic Theory. Tokyo Institute for Advanced Studies of Language. pp. 57-86.

Naranjo Hayes, E. (2022). The sociophonetics and lyrical codemixing of hot latin songs: a comparison of the spontaneous speech and the artistic performance speech of the top two artists in the global market, Bad Bunny and J Balvin [Tesis de Doctorado, Universidad de Alabama]. Recovered from https://www.proquest.com/docview/2715461917

Núñez Cedeño, R. (2014). "Fonología autosegmental". Fonología generativa contemporánea de la lengua española. Georgetown University Press. pp. 47-81

# Implosive consonants productions in the speech of Córdoba (Spain). Early results of a phonetic analysis

Maria De Luca
*Universidad de Córdoba*
mariadelucaa8@gmail.com

**Abstract**
The multiple varieties of speech of the Andalusian Spanish have been deeply studied through the years either by a dialectological or sociolinguistic perspective. Nevertheless, the same does not happen in the case of the variety spoken in Córdoba. This is the reason why it has been chosen to develop a study that analyses its phonetical production and with the following paper we aim to provide with a part of the results achieved so far, by focusing on the realizations of the implosive /s/ and the word-final consonants.

**Keywords:** phonetics, implosive consonants, Andalusian Spanish, Córdoba

## 1. Introduction and background

Throughout the years, many scholars have been focusing on developing multiple kind of analysis of the Spanish linguistic varieties spoken in Andalusia. Starting with the early works of Navarro Tomás (1938), those linguistic varieties have always proved to be very appealing, mostly due to their phonetical variations. Later in time, thanks to the publishing of the ALEA, Atlas lingüístico y etnográfico de Andalucía, (Alvar et al., 1973) this interest kept growing, leading to the creation of different works, each one focusing on one or more varieties and the geographical variations that could be noticed in their phonetical traits. Despite that, as for the variety spoken in the city of Córdoba there are very few studies we can rely on. Indeed, as for that topic, the most remarkable works are those by Agustín Uruburu Bidaurrazaga. Starting with his PhD thesis (Uruburu Bidaurrazaga, 1988), in which he studied the speech variety of young students of Córdoba capital, he soon developed a series of studies regarding different areas of the speech variety of the city.

### 1.1. The example of Uruburu and the new study

Even if Uruburu's work can be useful as for being part of an example of this speech variety, it presents some problems that we can establish as the reasons behind the need of a new research. First, those works can relate to a period that is no longer close to the present, which means that it does not reflect the reality of the current society, due to changes that affect language throughout the time (Berruto & Cerruti, 2015, 144-145). Secondly, in a few cases, he interviewed people that lived in Córdoba, but they were originally from other parts of Andalusia or Spain; this can be a problem if the idea is to develop an analysis of a specific speech variety because we can perceive the influence of other varieties. Finally, he mostly focused on just one specific group of people: young men and women studying at university.

Those reasons moved us into creating a new study in which not only we will update the notions of the variety spoken in this city, but we are also going to interview only second-generation people from Córdoba, to avoid other varieties' influences. Moreover, we are going to include different groups of people.

## 2. Methodology

Regarding the methodology, first of all, there was the need to establish the characteristics of the subjects; to do that, different studies have been taken as an example and as the base of the method employed for this research. Amongst those studies there is the PhD thesis of Elena Férnandez de Molina Ortés (2014) and we also followed the guidelines of the studies used for

the PRESEEA project (Moreno Fernández, 2021, 14-17). Finally, it has been decided to consider three sociolinguistic variables: age (from 18 to 34 years old; from 35 to 54 years old; from 55 years old or more), sex (man or woman) and level of study (none or primary and secondary school; high school; university). Then we established which phonetic features were to be analysed. Both vowels and consonants have been considered, nevertheless, in this case we are only showing a part of the study in which we provide the analysis of the realizations of implosive /s/ and final-word consonants in one female and one male representative speaker from each group of age.

## 2.1. The questionnaire and the interviews

Once we crossed over these two parts, it was time for the questionnaire to take form. Again, Férnandez de Molina Ortés's PhD thesis (2014) served as the base for the creation of the questionnaire, together with the work presented in the article (Inter-) Fonología del Español Contemporáneo (Putska et. Al, 2018). As the idea is to study the phonetic production of the chosen features in different conversational context, the questionnaire has three different parts: semi-structured conversation, pronunciation of a list of words, reading of a short text. The point is to make sure that the subjects feel free to express themselves in the most spontaneous way as possible, trying to avoid a laboratory-constructed speech, that is why the questionnaire has been structured in this specific sequence.

Later, we started to analyse the collected materials thanks to the use of PRAAT, to verify the realizations of the multiple features tested during the interviews.

## 3. Results

As the task of the material's analysis is still in progress, we can only provide a first part of results, showing, as we previously said, just the examples of some representative interviewees. In the case of the implosive /s/, it is possible to conclude that there are three possible realizations: aspiration or omission of the phoneme, causing different changes in the previous phonemes depending on its nature, or there is the possibility to produce the phoneme as fricative alveolar. On the other hand, something very similar can happen in the case of final-word consonants, but not all consonants are affected by relaxation in the pronunciation or omission, indeed one or the other can be produced when the word ends with a dental, nasal or liquid consonant.

In both cases, so far, it has been noted that aspiration and omission of the implosive /s/ and the relaxation and omission of the final-word consonant are more common in men rather than in women either if they are younger or elder people. Also, considering the kind of interview, we always have to take into account the fact that at the beginning the subjects felt in a situation in which they had to control their pronunciation, then they gradually felt more and more comfortable, up to the point that they were focusing more on what they were saying rather than on how they were expressing themselves.

## 4. Discussion and conclusion

For now, we can conclude by saying that one of the reasons that pushed this study into its creation was the lack of analysis regarding the speech variety of Córdoba. That is why the hope is to open the way to future researches, mostly because in this case the focus relies on the phonological plan and yet, there are more fields that need to be considered, as this variety possess in itself characteristics that belong to multiple varieties, and that is what made the speech of Córdoba unique in its style.

## 5. References

Alvar, M., Llorente, A. & Salvador, G. (1973). Atlas lingüístico y etnográfico de Andalucía (Vol. Tomo VI). https://www.cervantesvirtual.com/nd/ark:/59851/bmc1212717

Berruto, G. & Cerruti, M. (2015). Manuale di Sociolinguistica. Novara: De Agostini SPA.

Fernández de Molina Ortés, E. (2014). El Habla de Mérida. Tesis doctoral. 115-153. Universidad de Extremadura

Jiménez Fernández, R. (1999). El andaluz. Arco Libros.

Moreno Fernández, F. (2021). Guía PRESEEA para la investigación lingüística. Editorial Universidad de Alcalá.

Navarro Tomás, T. (1938). Dédoublement de phonemes dans le dialecte andalou. Travaux du Cercle Linguistique de Prague (8), 184-186.

Putska, E., Gabriel, C., Meisenburg, T., Burkard, M. & Dziallas, K. (2018). (Inter-) Fonología del español contemporáneo (I) FEC: Metodología de un programa de investigación para la fonología de corpus. Loquens, 5(1). Obtenido de https://doi.org/10.3989/loquens.2018.046

Uruburu Bidaurrazaga, A. (1988). Niveles Sociolingüísticos del Habla Juvenil Cordobesa. Tesis doctoral. Universidad de Córdoba.

Uruburu Bidaurrazaga, A. (1994). Materiales para el estudio del habla urbana de Córdoba. Servicio de Publicaciones Universidad de Córdoba.

Uruburu Bidaurrazaga, A. (1996). La lengua hablada en Córdoba (España). Revista Española de Lingüística Aplicada (11), 225-250.

Uruburu, A., Aparicio, M. C., Serrano Fuensanta Delgado, P. & García, J. J. (1995). Sociolingüística y dialectología en Córdoba capital. Alfinge(8), 367-377. Obtenido de http://hdl.handle.net/10396/16764

# Eastern Andalusian Spanish laxing and harmony: The effects of -r/-s deletion in words with final stress

Maialen Casquete de la Puente
*University of Illinois at Urbana-Champaign*
maialen2@illinois.edu

**Abstract**

Eastern Andalusian Spanish (EAS) is characterized by the loss of most word-final consonants in casual speech. This consonant deletion leads to adjacent vowel laxing which, in turn, drives vowel harmony: the open quality of the final vowel spreads leftwards to the stressed syllable (beso [ˈbe.so] 'kiss, sg' vs. besos [ˈbɛ.sɔ] 'kiss, pl'. While it usually assumed that the process is caused by the loss of word-final -s and that targets are stressed syllables, other patterns remain underexplored. For this study, I recorded 19 EAS speakers' production of words containing two mid vowels, with and without underlying final consonants (-r or -s), in words with final stress. Target words included triplets such as bebé ('baby'), beber ('to drink'), and bebés ('babies'). First and second formants of the final, stressed vowels, as well as the preceding (harmonized) vowels, were extracted. The results show that both -r and -s produce comparable changes in the formant structure of the two preceding vowels, leading to harmony regardless of the deleted final consonant. These findings suggest that, in EAS, vowel harmony (a) operates independently of stress and (b) can be triggered by the deletion of both final fricatives and liquids. I propose allophonic opening and V-to-V coarticulation as potential origins for vowel laxing and harmony in EAS.

**Keywords:** vowel laxing, vowel harmony, coarticulation, Eastern Andalusian Spanish

## 1. Introduction and literature review

In Eastern Andalusian Spanish, final vowels lax when the adjacent coda is deleted. For instance, dices '(you) say' is realized as [ˈdiθɛ] (c.f. [ˈdiθes] in Standard Castilian). The laxness of the final vowel its transmitted to the preceding mid vowels, and perhaps optionally beyond (Navarro Tomás, 1939; Alarcos Llorach, 1983; Rodríguez Castellano & Palacio, 1948). There are a few aspects that remain understudied, including the effects of underlying -r in laxing and harmony, and the role of word-stress in harmony.

Much of the literature in Eastern Andalusian vowel laxing has concentrated on words ending in -s (Navarro Tomás, 1939; Jiménez & Lloret, 2009; Henriksen, 2017), and many have claimed that the deletion of final liquids does not result in the same phenomenon (Navarro Tomás, 1939; Rodríguez Castellano & Palacio, 1948; Jiménez & Lloret, 2009, 2019). On the other hand, Herrero de Haro (2019) found that deletion of -r also triggered vowel opening, though with slight differences compared to -s. This difference was not mirrored in perception: although EAS speakers could identify if a consonant had been deleted, they were unable to identify which.

In addition, only unstressed final vowels are said to trigger harmony. This process is therefore believed to be perceptually motivated: the quality of the accentually weak final vowel is improved by transmitting its laxness to a structurally stronger position (Jiménez & Lloret, 2009, 2018; Kaplan, 2020, Davis & Pollock, 2024). A coarticulatory origin is ruled out, as stressed vowels more frequently trigger anticipatory assimilation (Fowler, 1981; Magen, 1997; Majors, 2006; Recasens, 2016). Despite these assumptions, stress restrictions have often been assumed rather than experimentally tested.

In this study, I compare the effects of underlying -r, -s and bare final vowels (e.g. *beber, bebés, bebé*), in the two vowels within the word. I report in production and perception experiments, dealing with words with final stress.

## 2. Methodology

For production, 19 speakers from Granada were recorded. I designed an oral alternative of a sentence completion task, using stimuli produced by an EAS speaker. The stimuli prompted words with final stress, ending in -or, -er, -es, -e, -o. I segmented final and penultimate vowels of 994 words with final consonant deletion and extracted their F1 and F2 mean values. Two linear mixed-model regressions (one for F1, one for F2) were conducted in R, using the lme4 package (Bates et al, 2015), considering Position (V1, V2), Vowel (e, o) and Ending (-r, -s, V) + interactions. Random intercepts were Participant and Token.

The same participants (+20 control from Bilbao) participated in a perception experiment. Participants read a written word and heard two audios produced by an EAS speaker. They had to identify which of the two recordings corresponded to the written word. Stimuli presented contrasts between -s/V (e.g. bebés vs. bebé), -r/V (beber vs. bebé), and -s/-r (bebés vs beber). A generalized linear logistic mixed-effects model of the binomial family (glmer) was fitted, considering Type of contrast (-s/V, -r/V, -s/-r) City (Granada, Bilbao) + interactions. Random intercepts were Participant and Token.

## 3. Results

There were significant differences in F1 and F2 between words underlyingly ending in stressed -e (e.g. bebé) and -er or -es (e.g. beber, bebés) (p.<.001, in both regressions and comparisons, estimates around 160Hz for F1 and 350Hz for F2). No difference was found between -er and -es (Est. = 0.2, t = .02, p = .98). Underlying final -o (coló) was different from the same vowel in -or (color), but only F1 reached significance (Est. = -155.45, t = -2.73, p < .002). No significant differences were found between final and penultimate vowels in the -r/-s conditions, indicating harmonization. In perception, Granada participants were highly accurate in identifying whether consonant deletion had occurred (p.<.001 in both -r/V and -s/V). Participants could identify which consonant (-r or -s) had been deleted above chance (p. =.01); however, accuracy was very limited (58%).

## 4. Conclusion and discussion

Results indicate that mid vowels in words ending in -r and -s do not differ from each other (i.e, *beber* and *bebés* are identical in formant structure). With an underlying consonant, final vowels lax and penultimate vowels harmonize (/e/ lowers and centralizes, /o/ only lowers).

Finding that -r triggers vowel laxing in EAS supports the idea that allophonic vowel opening preceded the phonologized laxing found in the dialect today (Sanders, 1994; Corbin, 2006; Henriksen, 2017). Evidence of some opening next to -r in other Spanish varieties (Navarro Tomás, 1957; Lipski, 1987; González de Anda, 2013) suggests this process may be linked to articulatory factors such as pharyngeal constriction. This earlier diachronic stage could not be proposed for EAS if -r did not also cause vowel laxing.

Given that harmony is found in words with final stress, I would like to suggest that EAS harmony may not be (at least uniquely) motivated by perceptual constraints, but rather by V-to-V coarticulation. Coarticulation is widely recognized as the primary source of phonologized harmony cross-linguistically (Ohala, 1994).

Lastly, production results were partially contradicted in perception. Despite not finding a difference in formant structure between -s and -r, EAS speakers identified the deleted consonant above chance (albeit with very limited accuracy). There may be subtle cues such as remaining aspiration or vowel length, which were not considered in production.

## 5. References

Alarcos Llorach, Emilio (1983). Más sobre vocales andaluzas. Philologica hispaniensia: in honorem Manuel Alvar, 1, 49–56.

Davis, Stuart & Matthew Pollock (2024). A feature alignment approach to plural realization in Eastern Andalusian Spanish. Languages, 9(5), 166.

Fowler, Carol (1981). Production and perception of coarticulation among stressed and unstressed vowels. Journal of Speech and Hearing Research, 24(1), 127–139.

González de Anda, Gisela (2013). Acoustic analysis of the allophones of the Spanish mid-front vowel /e/. MA paper, The University of Texas at El Paso.

Henriksen, Nicholas (2017). Patterns of vowel laxing and harmony in Iberian Spanish: Data from production and perception. Journal of Phonetics, 63, 106–126.

Herrero de Haro, Alfredo (2019). Catorce vocales del andaluz oriental: Producción y percepción de /i/, /e/, /a/, /o/ y /u/ en posición final y ante /-s/, /-r/ y /-θ/ subyacentes en Almería. Nueva Revista de Filología Hispánica, 67(2), 411–446.

Hualde, José Ignacio & Brent P. Sanders (1995). A new hypothesis on the origin of the Eastern Andalusian vowel system. Annual Meeting of the Berkeley Linguistics Society, 21(1), 426–437.

Jiménez, Jesús & María-Rosa Lloret (2009). Un análisis óptimo de la armonía vocálica del andaluz. Verba, Anuario galego de filoloxia, 6, 293–325.

Jiménez, Jesús & Maria-Rosa Lloret (2018). Restricciones articulatorias y perceptivas en la armonía vocálica del andaluz oriental y del murciano. In Actas do XIII Congreso Internacional de Lingüística Xeral, 490–496. Vigo: Universidade de Vigo.

Jiménez, Jesús & Maria-Rosa Lloret (2019). Vowel harmony. In Silvia Colina and Fernando Martínez-Gil (eds.), The Routledge Handbook of Spanish Phonology, 100–128. London: Routledge.

Kaplan, Aaron (2020). Persistence and opacity in Eastern Andalusian harmony. In Ryan Bennett et al. (eds.), Proceedings of the 2020 Annual Meeting on Phonology. Washington DC: Linguistic Society of America.

Lipski, John M. (1987). Fonética y fonología del español de Honduras. Honduras: Guaymuras.

Magen, H. Susan (1997). The extent of vowel-to-vowel coarticulation in English. Journal of Phonetics, 25(2), 187–205.

Majors, Teresa (2006). The development of stress-dependent harmony. Southwest Journal of Linguistics, 25(1), 59–84.

Matluck, Joseph H. (1963). La /e/ trabada en la Ciudad de México: Estudio experimental. The University of Texas Institute of Latin American Studies Offprint Series, 9, 5–34.

Navarro Tomás, Tomás (1939). Desdoblamiento de fonemas vocálicos. Revista de Filología Hispánica, 1, 165–167.

Navarro Tomás, Tomás (1957). Manual de Pronunciación Española. Madrid: Consejo Superior de Investigaciones Científicas.

Ohala, John. J. (1994). Towards a universal, phonetically-based, theory of vowel harmony. Third international conference on spoken language processing, 3, 491-494.

Perissinotto, Giorgio S. A. (1975). Fonología del español hablado en la ciudad de México. México DF: El Colegio de México.

Recasens, Daniel (2016). Stressed vowel assimilation to palatal consonants in early Romance. Journal of Historical Linguistics, 6(2), 201–246.

Rodríguez Castellano, Lorenzo & Antonio Palacio (1948). El habla de Cabra. Revista de Dialectología y Tradiciones Populares, 4(3), 387.

# An acoustic analysis of implosives in Khmer

Nielson Sophann Hul
*University of Washington*
nhul@uw.edu

**Abstract**

This study examines the phonetic and phonological characteristics of Khmer implosives as produced by 22 native speakers in Phnom Penh. Using an automated method, we extracted acoustic measures of Voicing Proportion (VP) and Voicing Amplitude Slope (VA_Slope) to capture the gradient nature of implosivity. These measures were validated against aspirated, unaspirated, and nasal stops, then used to assess effects of Word-internal Position, Place of Articulation (POA), and speaker differences.

Results show that VP and VA_Slope are positively correlated and jointly define implosivity. Word-initial implosives exhibit greater slope and more voicelessness than medial tokens, while labials are more implosive than coronals. Female speakers tend to produce stronger implosives than males. Notably, Khmer implosives derive historically from voiceless stops, contrasting with the common cross-linguistic pathway from voiced stops. This highlights the importance of VP as a key diagnostic parameter.

Overall, the findings refine the acoustic characterization of Khmer implosives, reveal substantial interspeaker variation, and position Khmer as a critical case for understanding the typology of implosives and the phenomenon of implosivity cross-linguistically.

**Keywords:** Khmer, implosives, voicing, phonetics, Cambodia

## 1. Introduction and Background

Implosives constitute a rare but cross-linguistically significant class of stops. While often associated with voiced articulation, their phonetic realization and diachronic origins vary widely. This study examines Khmer implosives, focusing on how they pattern acoustically and phonologically within the language.

Khmer is particularly interesting because its implosives are thought to derive from a historically voiceless series of stops, contrasting with the more common pathway from voiced stops. This difference makes Khmer a valuable case study for refining theories of implosivity.

### 1.1 Research questions

This study addresses the following:

Which acoustic parameters best characterize Khmer implosives?

How do these parameters vary across phonological contexts, specifically Word-internal Position and Place of Articulation (POA)?

How do Khmer implosives compare to other stop classes (aspirated, unaspirated, nasal) in the language?

To what extent is there interspeaker variation in the production of implosives?

## 2. Methodology

### 2.1 Participants

Acoustic data were collected from 22 native speakers of Khmer from Phnom Penh. Participants represented both male and female speakers, allowing for examination of gender effects.

### 2.2 Data Collection and Stimuli

Speech materials were designed to elicit implosives in multiple phonological contexts, including word-initial, intervocalic medial (Vmed), and post-nasal medial (Nmed) positions.

Both labial and coronal places of articulation were targeted. Comparable tokens of aspirated, unaspirated, and nasal stops were also elicited.

## 2.3 Acoustic measures
To analyze implosive production, an automated method was used to extract several parameters from the stop closure interval. Two primary measures were emphasized:
Voicing Proportion (VP): the proportion of closure time with vocal fold vibration.
Voicing Amplitude Slope (VA_Slope): the rate of amplitude change in voicing over time.

## 2.4 Analytical Approach
These measures were validated by comparing implosives against aspirated, unaspirated, and nasal stops. Statistical modeling was then employed to test the effects of Word-internal Position, POA, and speaker-related factors on VP and VA_Slope. Furthermore, we modeled the interaction effect of POA and Position on VA_Slope and VP and finally the effect of VP on VA_Slope.

## 3. Results
The results demonstrate that implosivity in Khmer is gradient rather than categorical, with both Voicing Proportion (VP) and Voicing Amplitude Slope (VA_Slope) capturing different aspects of its variability. Word-internal position was a significant factor: word-initial implosives consistently displayed greater VA_Slope values and higher proportions of voicelessness compared to intervocalic medial (Vmed) and post-nasal medial (Nmed) contexts. Place of articulation also conditioned the results, with labial implosives exhibiting stronger implosivity than their coronal counterparts. Speaker-related effects emerged as well, as female speakers tended to produce more robust implosive gestures than male speakers. Importantly, VP and VA_Slope were found to be positively correlated, showing that these measures converge to jointly characterize the phenomenon of implosivity and providing empirical support for their use as complementary diagnostics.

## 4. Discussion and Conclusion
These findings highlight that implosivity in Khmer is both gradient and systematically conditioned by phonological structure. The strengthening of implosivity in word-initial position and at labial places of articulation demonstrates how prosodic and articulatory contexts interact to shape its realization. The fact that Khmer implosives derive historically from a series of voiceless stops underscores the relevance of VP, since variation in closure voicing in this case does not merely reflect speaker inconsistency but aligns with the language's diachronic trajectory. This contrasts with the more common cross-linguistic pattern of implosives evolving from voiced stops, making Khmer a valuable counterpoint in typological comparison.

At the same time, the observed interspeaker variation indicates the importance of analyzing larger and more diverse speaker samples when investigating rare sound classes. By integrating quantitative acoustic measures with phonological context, this study not only refines our understanding of Khmer implosives but also illustrates how standardized metrics such as VP and VA_Slope enable systematic cross-linguistic comparisons. Ultimately, these findings situate Khmer more clearly within the typology of implosives while contributing to a broader theoretical understanding of implosivity as a phonetic and phonological phenomenon.

## 5. References
Clements, G. N. & Osu, S. (2002). Explosives, implosives and nonexplosives: The linguistic function of air pressure differences in stops. Laboratory Phonology, 52.

Cohn, A. C. (2006). Is there gradient phonology? In G. Fanselow, C. Féry, R. Vogel & M. Schlesewsky (Eds.), Gradience in Grammar: Generative Perspectives (pp. 25–44). Oxford University Press.

Diffloth, G. (n.d.). A History of the Khmer Language [Article Manuscript - Kroch Library at Cornell University].

Ladefoged, P. & Maddieson, I. (1996). The Sounds of the World's Languages (1 edition). Wiley-Blackwell.

Lindau, M. (1984). Phonetic differences in glottalic consonants. Journal of Phonetics, 12, 147–155.

Maspong, S. (2023). Synchrony and Diachrony of Vowel Quality Differences Across Registers in Mon-Khmer Languages [PhD Thesis, Cornell University]. https://ecommons.cornell.edu/items/f77438cc-9fbb-49e5-861a-cf18f5f6bbf6/full

# Influence of Velars and Uvulars on Vowel Context in Four Typologically Unrelated Languages

Ekaterina Shepel[ab], Jalal Al-Tamimi[a]

[a]*Université Paris Cité, LLF, CNRS/* [b]*LST, Saarland University*

ekaterina.shepel@uni-saarland.de, jalal.al-tamimi@u-paris.fr

**Abstract**

We investigate the coarticulation between velars~uvulars and surrounding vowels in four languages from three unrelated families where velars and uvulars as phonemes: Alutor (Chukotko-Kamchatkan), Chukchi (Chukotko-Kamchatkan), Mehweb (East Caucasian), and Shughni (Southeastern Iranian < Indo-European). Following the Laryngeal Articulator Model (LAM), the articulation of uvulars requires tongue rising and retraction. As the articulation of velars does not require retraction of the tongue, we expect the quality of vowels to exhibit changes in the uvular environment compared to velar. Various acoustic properties of vowels in the context of velar and uvular stops are investigated including static and dynamic properties with the quality of vowels at the stationary and transition periods. To confirm that the coarticulatory processes are connected to the uvular environment, we compare the contribution of various properties of vowels to the differentiation between velars and uvulars using Random Forests (RF) classification tool. While vowel quality changes in the uvular environment in all four languages; the quality of vowels at the midpoints, onset/offset, dynamics of formants show different contribution to the differentiation between velars and uvulars across the four languages. The variability concerns the type of changes (retraction or centralization of vowels), direction of coarticulation, whether coarticulatory effects are observed only on transitions or affect the quality at the midpoint.

**Keywords:** velars, uvulars, gutturals, coarticulation, phonetic typology

## 1. Introduction and background

We follow the predictions of Laryngeal Articulator Model (LAM) (Esling 2005; Esling et al. 2019), which documents various types of articulation within the post-velar natural class. LAM explains that post-velar natural class induces retraction of the tongue, with a *back and down* gesture (Al-Tamimi 2023), leading to both more back and more open productions of the vowel in the environment of post-velars, respectively as one gesture.

In this research, we compare the productions of vowels in the environments of velar and uvular consonants expecting to find the clear differences connected to the properties of post-velar natural class: as a part of post-velar natural class, uvulars are expected to show influence on surrounding vowels (Sylak-Glassman 2014a), while velars are expected to pattern with oral consonants (Al-Tamimi 2023). Due to the retraction and raising of the tongue during the articulation of uvulars, uvulars do induce retraction of surrounding vowels (Al-Tamimi 2023) which patterns with the effects on vowels occurring in the pharyngeal environment (Sylak-Glassman 2014a). Velars do influence the vowel environment compared to the coronal consonants showing the backing of front vowels (Sylak-Glassman 2014b), however they do not cause backing of open vowels and thus this process cannot be described as retraction and does not pattern with the coarticulatory effects of uvulars.

Here, we focus only on environments of stops /k/ and /q/, in Shughni we include also voiced /g/ which shows the same properties as /k/ in terms of place of articulation.

## 2. Data

The data of four languages are used in this research: Alutor (Chukotko-Kamchatkan; Glottocode: alut1245), Chukchi (Chukotko-Kamchatkan; chuk1273), Mehweb (East

Caucasian; mege1234), and Shughni (Southeastern Iranian < Indo-European; shug1248). The vowel inventories of these languages show the oppositions between front, back, and open vowels which allows us to compare the behavior of velars and uvulars across these environments; in Mehweb, pharyngealized vowels are also present. The data comes from the controlled samples collected during linguistic fieldwork. In total, the data of 21 speakers were used in this research: 2 f and 1 m for Alutor, 5 f for Chukchi, 4 f and 5 m for Mehweb, 4 f for Shughni. Speakers produced a list of stimuli and for the purpose of our study, we look at a subset of /k/ and /q/ across various vowel contexts in word-initial, -medial, and -final positions. The sample analyzed here is composed of 837 velar and 270 uvular tokens for Alutor; 910 and 432 tokens for Chukchi, 239 and 1286 for Mehweb, 202 and 140 for Shughni. For Mehweb, we also include the data of vowels in non-velar and non-post-velar context for comparison.

## 3. Methodology

The data of the languages were annotated in Praat (Boersma & Weenink 2024) following similar procedures; the formant values were extracted using FastTrack (Barreda 2021). For vowels in the velar and uvular environment, the measures obtained included the values of F1, F2, F3 around the midpoints of the vowels and at the onset/offset, the distances between formants, and dynamic measures including rate of change of formants and formant distances. The duration of vowels was also extracted but it did not contribute to the analysis, so we do not discuss it further. We used Random Forests (RF), which was previously used in the post-velar natural class (Al-Tamimi, 2023; Arkhipov et al, 2023), to compare the contribution of various properties of vowels to the distinction between velars and uvulars (according to the position within a word) and thus confirm that particular properties signal uvular context.

## 4. Results

The results show that in all the languages investigated vowels show changes in the uvular environment compared to velar; however, the languages differ in the coarticulation patterns and direction. In Alutor and Chukchi, both previous and following vowels show retraction in the uvular context; in Mehweb, plain vowels in the uvular context show retraction; in Shughni, both retraction and centralization are observed depending on the type of the context: vowels preceding uvulars show centralization, while following vowels – retraction. While in Alutor coarticulation affects both previous and following vowels equally, in Chukchi the vowels following the uvular experience more changes, in Shughni the changes affect more the previous vowel. Pharyngealized vowels in Mehweb do not experience changes in the uvular context. The coarticulatory processes affect significantly front vowels in all four languages, causing their retraction, while back vowels are almost not affected.

   The results of RF suggest that across the languages, there is no single feature which signals that the environment is uvular and not velar. Formant values at the onsets/offsets and midpoints, formant distances contribute to the differentiation between velars and uvulars, but how the importance of the contribution is distributed between these parameters depends on the language. Dynamics of vowels contributed to the analysis only in Alutor and Chukchi.

## 5. Discussion and conclusion

The four languages investigated show similar trends whereby vowels preceding and following uvulars show changes compared to the vowels in the velar environment: during the coarticulation with uvulars openness and backness of vowels are affected. However, the quality of vowels show different changes in the languages of the sample: while front vowels are realized as retracted in all four languages, open vowels show different changes resulting in the retraction and centralization on the whole vowel space. Various properties of vowels

including quality on the transitions and stationary periods are affected but no single correlate with uvular environment was observed on our data. Our results partially confirm our hypothesis and point to language-specific differences in the dynamics of vowel changes.

**Figure 1**. Vowel spaces in velar and uvular context (normalized per language).

## 6. References

Al-Tamimi, J. (2023). *On the contribution of "fine phonetic detail" (FPD) to gradience in phonology: Acoustic, articulatory, perceptual and automatic methods*. [HDR in phonetics and phonology]. Université Paris Cité.

Arkhipov, A., Daniel, M., Shiryaev, A. & Shepel, E. (2023). Evaluating formant estimations and discrete cosine transform to differentiate between pharyngeal fricatives in Mehweb. In: R. Skarnitzl, J. Volín (eds.), *Proceedings of the 20th ICPhS*. Guarant International (pp. 3407–3411).

Barreda, S. (2021). Fast track: Fast (nearly) automatic formant-tracking using praat. *Linguistics Vanguard, 7*(1), 20200051.

Boersma, P. & Weenink, D. (2024). *Praat: doing phonetics by computer*. Computer program. Version 6.4.07, retrieved 17 March 2024.

Esling, J. H. (2005). There are no back vowels: The Laryngeal Articulator Model. *Canadian Journal of Linguistics/Revue canadienne de linguistique, 50*(1-4), 13–44.

Esling, J., Moisik, S., Benner, A. & Crevier-Buchman, L. (2019). *Voice Quality: The Laryngeal Articulator Model (1st ed.)*. Cambridge University Press.

Sylak-Glassman, J. C. (2014a). *Deriving natural classes: The phonology and typology of post-velar consonants.* University of California, Berkeley.

Sylak-Glassman, J. C. (2014b). The effects of post-velar consonants on vowels in Ditidaht. In *Papers for the international conference on Salish and neighbouring languages* (Vol. 49).

# Variable realization of ejectives in isolated vs connected speech in Northern Lushootseed

Ted Kye
*University of Washington*
kyeted20@gmail.com

**Abstract**
The goal of this study is to examine the effects of different speech styles (i.e., isolated vs. connected speech) on the production of ejectives in Northern Lushootseed. Archival recordings (dating to the 1950's) of an Indigenous female speaker were examined. An acoustic analysis of ejective and plain stops was conducted by examining VOT, burst duration, burst intensity, f0 perturbation, and jitter perturbation. The current findings reveal that there is a strong effect of speech style on the production of ejectives, where ejectives were produced with greater "strength" (as it's been traditionally defined) in isolated speech than in connected speech.

**Keywords:** ejectives, speech production, Lushootseed, connected speech, isolated speech

## 1. Introduction

Ejectives have traditionally been classified as falling under two types: strong vs. weak (Lindau 1984; Kingston 1985, 2005). These two types are based on a number of acoustic parameters used to distinguish the way they are produced across different languages. However, there have been several studies revealing that there is far more variability in the realization of ejectives across languages (Wright et al. 2002; Warner 1996; Hajek & Stevens 2005; Kye 2021, 2023; Percival 2015, 2019, 2024; Vicenik 2010), which suggests that many language's ejectives do not strictly fall under the strong vs. weak classification. The question of whether variability can be observed across different speech styles, however, have garnered far less attention. Most studies that examine the acoustic characteristics of ejectives examine how they are produced in isolated speech (such as wordlists and/or carrier phrases). There have been far fewer studies that examine the realization of ejectives in connected speech (such as conversational or storytelling) (Kye 2021, 2023; Vicenik 2010), and far fewer studies that compares its production between isolated vs. connected speech (Kye & Percival 2023). The goal of this study is to examine the realization of ejectives in isolated vs. connected speech in Northern Lushootseed (Central Salish language), and to address the following question: Are there differences in the realization of ejectives when we compare isolated vs. connected speech in Northern Lushootseed?

## 2. Materials and methods

In this study, recordings that date to the 1950's were examined. These recordings are reel-to-reel tapes that come from the Metcalf Collection and were digitized at 44.1kHz with a 32-bit depth. The speaker who was examined is Martha Lamont: An Indigenous elder of the Northern Lushootseed dialect. For the connected speech data, four recordings of storytelling and two recordings of private correspondence were examined (total duration of approximately 51 minutes). For the isolated speech data, one recording of a field work session (where words were produced in isolation) was examined (total duration of 32 minutes).

For data analysis, ejective stops [t' q'] were compared with plain stops [t q] to characterize the contrast between ejectives and plain (voiceless) stops. Each stop was examined in root-initial position. In the isolated speech data, there was a total of 43 ejectives (25 [q'] and 18 [t']) and 49 plain stops (22 [q] and 27 [t]). In the connected speech data, there was a total of 36 ejectives (14 [q'] and 22 [t']) and 92 plain stops (11 [q] and 81 [t]). Several acoustic measurements were examined using Praat (Boersma & Weenink 2024), which includes VOT, burst duration, burst intensity, f0 perturbation (difference between f0 at voice onset and f0 at

vowel midpoint), and jitter perturbation (difference between jitter at voice onset and jitter at vowel midpoint). Three independent variables were examined: Place of articulation (alveolar vs. uvular), laryngeal type (plain vs. ejective), and speech style (isolated vs. connected). A two-sample t-test was used to test the difference between each level of the independent variables for each acoustic measure.

## 3. Results

For VOT (see Figure 1a): (1) uvulars is significantly greater than alveolars; (2) ejectives is significantly greater than plain; and (3) isolated speech was significantly greater than connected speech for ejective stops only, while there was no difference for plain stops. For burst duration (see Figure 1b): (1) uvular's was significantly greater than alveolars; (2) ejectives have a significantly longer burst duration than plain stops; and (3) ejective uvulars in isolated speech was significantly greater than ejective uvulars in connected speech. For burst intensity (see Figure 1c): Ejectives is significantly greater than plain stops in connected speech, and isolated speech (overall) is greater than connected speech. For f0 perturbation (see Figure 2a): Ejectives are depressed, while plain stops are raised, but no significant difference in f0 perturbation between the two speech styles. For jitter (see Figure 2b): Jitter was significantly greater in ejectives than plain stops, and jitter was significantly greater in connected speech than isolated speech.
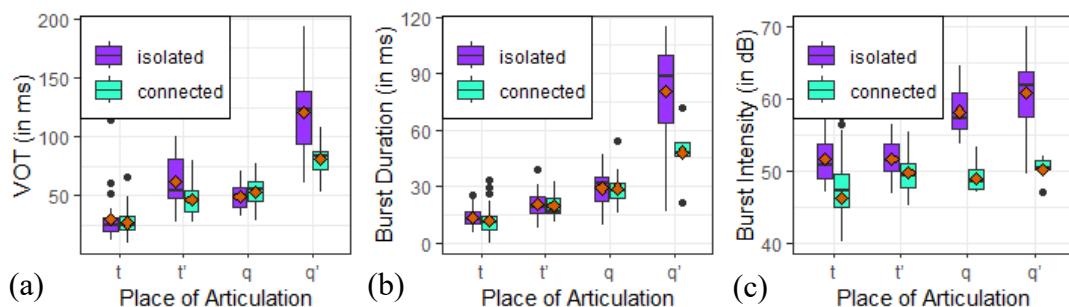


Figure 1. Box plots illustrating release parameters: (a) VOT, (b) burst duration, and (c) burst intensity.
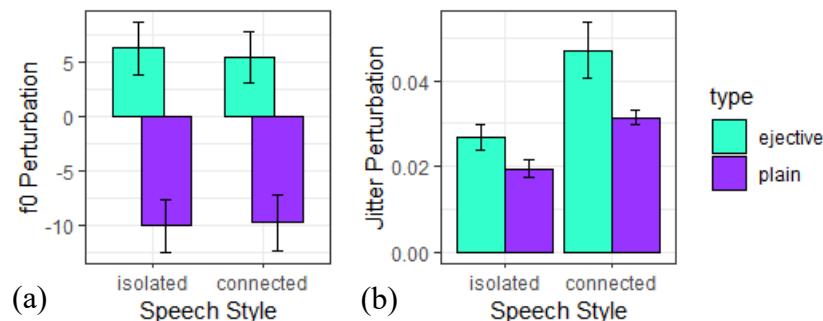


Figure 2. Bar graphs illustrating voice onset quality parameters: (a) f0 perturbation, and (b) jitter perturbation.

## 4. Discussion

The longer VOT, louder burst intensity, and lower jitter of ejectives in isolated speech suggests that these sounds are produced more closer to "strong" ejectives in this speech style. This may be due to greater hyperarticulation in isolated speech (Lindblom 1990), where speech sounds tend to be produced with increased articulatory and aerodynamic force when speaking carefully. However, while the burst intensity of ejectives in isolated speech is greater than ejectives in connected speech, it did not significantly differ from plain stops within the same speech style (i.e., isolated speech). These findings suggest that some of the variability attributed to language-specific differences is, instead (or additionally), attributable to differences in speech style.

## 5. References

Boersma, Paul & David Weenink (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.17, retrieved 10, September 2021 from http://www.praat.org/.

Hajek, John & Mary Stevens (2005). On the acoustic characterization of ejective stops in Waima'a. In Interspeech 2005, 2889-2892.

Kingston, John. (1985). The phonetics and phonology of the timing of oral and glottal events. Ph.D. dissertation, Department of Linguistics, University of California, Berkeley.

Kingston, John. (2005). The phonetics of Athabaskan tonogenesis. Amsterdam Studies in the Theory and History of Linguistic Science Series 4, 269, 137.

Kye, Ted (2021). Ejective typology: The case of Lushootseed. In Precedings of the 56th International Conference on Salish and Neighbouring Languages, 154-166.

Kye, Ted (2023). A grammar of Lushootseed: Phonetics, phonology, morphology. (Doctoral Dissertation, University of Washington).

Kye, Ted, & Maida Percival (2023). Cross-linguistic realization of lateral ejective affricates in connected vs isolated speech. In Journal of Acoustical Society of America, 154(4), A38-A38.

Lindau, Mona (1984). Phonetic differences in glottalic consonants. In Journal of Phonetics, 12(2), 147-155.

Lindblom, Björn (1990). Explaining phonetic variation, A sketch of the H&H theory. In W.J. Hardcastle & A. Marchal (Eds.), Speech production and speech modelling (pp 403-439). Dordrecht: Springer Netherlands.

Percival, Maida (2015). Dene stop contrasts: Data from Délįnę Slavey. In The Scottish Consortium of ICPhS 2015 (ed.), Proceedings of the 18th International Congress of Phonetic Sciences, 976-981.

Percival, Maida (2019). Contextual variation in the acoustics of Hul'q'umi'num' ejective stops. In Calhoun, Sasha, Escudero, Paola, Tabain, Marija, & Warren, Paul (eds.), Proceedings of the 19th International Congress of Phonetic Sciences, 19, 3270-3274.

Percival, Maida (2024). Production and perception of ejective stops across languages. (Doctoral Dissertation, University of Toronto).

Vicenik, Chad (2010). An acoustic study of Georgian stop consonants. Journal of the International Phonetic Association, 40(1), 59-92.

Warner, Natasha (1996). Acoustic characteristics of ejectives in Ingush. In Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96, 3, 1525-1528.

Wright, Richard, Sharon Hargus, & Katherine Davis (2002). On the categorization of ejectives: data from Witsuwit'en. Journal of the International Phonetic Association, 32(1), 43–77.

# Correlations of acoustic properties of ejective stops in Q'anjob'al

Maida Percival
*University of Oxford*, *University of Toronto*
maida.percival@phon.ox.ac.uk

**Abstract**
This goal of this paper is to determine whether there are correlations between the presence or absence of a period of silence within ejective stop releases in Q'anjob'al and acoustic dimensions that have been proposed to group typologically with the presence or absence of silence. Linear mixed regression analyses on a corpus of 2500 ejective stop tokens from 20 L1 speakers of Q'anjob'al found no correlations between silence in the release and following vowel F0, H1-H2, or jitter, but did find a correlation between silence and vowel intensity, and silence and burst intensity, at least for bilabials. The results support a model of ejective articulation where vocal fold stiffness alone cannot account for the extent of variation present in ejective production.

**Keywords:** ejective stops, Q'anjob'al, speech production, acoustic variation, phonetic typology

## 1. Introduction and background

Ejectives are produced with a wide range of variation, and it is still not entirely clear what underpins the variation. Early proposals (Lindau 1984; Kingston 1985) define two types of ejectives based on vocal fold stiffness and timing of the glottal closure: strong (or stiff) and weak (or slack). A strong version of this typology predicts that long releases (with a period of silence where the vocal folds remain closed following the oral burst), loud bursts, and modal (or tense) voicing onsets should cooccur in stiff ejectives and that short releases (no period of silence), quiet bursts, and creaky voice onsets should cooccur in slack ejectives.

Further research has suggested that a binary typology based on vocal fold stiffness cannot account for all variation present in ejectives. Wright et al. (2002) found that short VOT and lowered onset F0 do not correlate in ejectives in Witsuwit'en. This was replicated in Lushootseed (Kye 2021, 2023). To account for this, Wright et al. (2002) suggest that medial compression also plays a role in keeping the vocal folds closed instead of and/or in addition to longitudinal tension.

Understanding which acoustic properties of ejectives can and cannot cooccur will allow us to better model how articulation shapes acoustics for ejectives. As such, the goal of this paper is to contribute additional data to this question, and test whether the different values of the acoustic dimensions associated with each ejective type correlate within individual tokens in Q'anjob'al (Mayan). Q'anjob'al is particularly suitable for testing these correlations, because, in related research (Percival 2024), its ejectives were found to have about equal proportions of silence in the release (as in strong ejectives) and no silence in the release (just a burst, as in weak ejectives).

## 2. Methodology

### 2.1 Data

The data consisted of a corpus of 2500 Q'anjob'al ejective stop tokens from Percival (2024), which included [p', t', k', q] at word-initial, intervocalic, and word-final positions. Recordings were words produced in isolation by 20 L1 speakers (11F, 9M; mean age = 32) in Guatemala.

## 2.2 Analysis

The ejective stop tokens were annotated in Praat (Boersma & Weenink 2023). Counts of release type (silence/no silence) were made as were measurements of maximum burst intensity, and four vowel measurements: the differences between the point at 10% into a following vowel and the midpoints for F0 and H1-H2 (ΔF0 and ΔH1-H2), the vowel onset interval and middle interval for jitter (Δjitter), and vowel intensity at 10% into the vowel and maximum for intensity (Δintensity).

Five generalized linear mixed effects regression models were made in R (R Core Team) to examine the effect of each acoustic dimension on release type (no silence vs. silence). Random intercepts and slopes were included for participants and words. The data was subset to exclude ejectives produced as fricatives or implosives, and further subset to exclude null values such that between 1391-2163 tokens were used in the models.

## 3. Results

A negative correlation was found between release type and burst intensity in that tokens with no silence are significantly more likely to have higher burst intensity than tokens with silence. Since this result was contra any expected hypothesis based on ejective stop production, further exploration of the data was done, including a (admittedly non-converging) model which suggest that bilabial [p'] is the only place of articulation where greater intensity significantly co-occurs with no period of silence. A significant effect of Δintensity for the following vowel was also found, where vowel intensity increased less when preceded by a period of silence compared to when there was none. No significant effect of ΔF0, ΔH1-H2, nor Δjitter was found on release type.
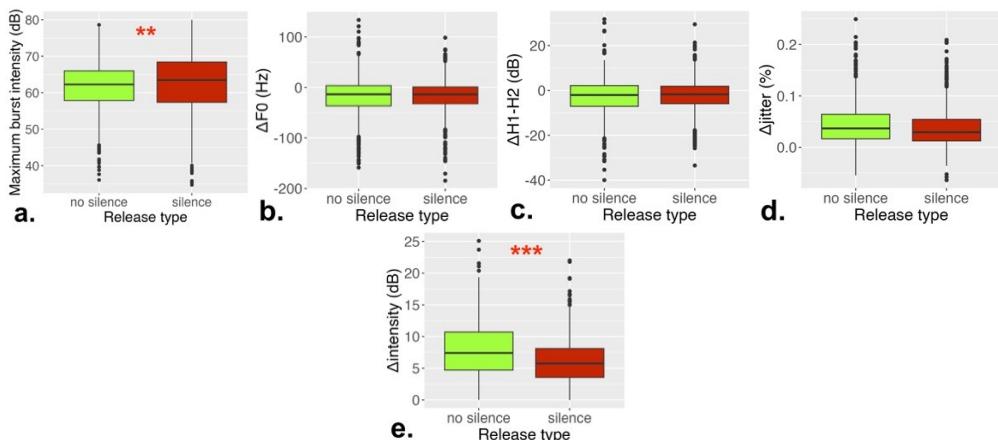


Figure 1. a. Burst intensity, b. ΔF0, c. ΔH1-H2, d. Δjitter, e. Δintensity by release type. *** = $p < .001$; ** = $p < .01$

## 4. Discussion and conclusion

Significant results were only found between release type and measurements of intensity. The correlation between release type and Δintensity on the following vowel suggests that there is more constriction of airflow at the vocal folds at the voicing onset when there is no silence (and regardless of whether other cues to creaky voicing which involves partially constricted vocal folds are present). The correlation between release and burst intensity where bilabial stops unexpectedly had lower intensity bursts when they occurred with silence may be due to weak oral airflow as the glottalized bilabial stop in Q'anjob'al (and Mayan languages more generally) is unique in that is often produced as an implosive or just glottalized (Kuang 2019). Additionally, bilabials are reported to be an uncommon place of articulation for ejectives (Maddieson 1984) presumably because it is harder to attain aerodynamic targets for ejectives with a larger intraoral cavity.

No correlation between release type with F0, H1-H2, or jitter seems to indicate that a period of silence, a strong ejective characteristic, can co-occur with creaky voicing at the onset of the following vowel, a weak ejective characteristic. A model where ejectives are classified as only either strong or weak could not account for this variation, and so this is further support that a view of ejectives with more than two types of realizations is needed. This result provides additional data in support of Wright et al. (2002) model that suggests medial compression must be able to act or be timed independently of longitudinal tension in ejective production. This study adds to the small but growing body of research aiming to eventually model how and whether laryngeal (or supralaryngeal) gestures in the production of ejective may have one or multiple acoustic consequences in production for ejectives.

## 5. References

Boersma, Paul & Weenink, David (2023). Praat. http://www.fon.hum.uva.nl/praat/. Accessed 2023/11/15.

Kingston, John (1985). *The Phonetics and Phonology of the Timing of Oral and Glottal Events* [Doctoral dissertation, University of California Berkeley].

Kuang, Jianjing (2019). Contextual variation of glottalic stops in Q'anjob'al. In Sasha Calhoun, Paola Escudero, Marija Tabain, & Paul Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 1114–1118). Australasian Speech Science; Technology Association Inc.

Kye, Ted (2021). Ejective typology: The case of Lushootseed. In D. K. E. Reisinger, Hannah Green, Laura Griffin, Marianne Huijsmans, Gloria Mellesmoen, & Bailey Trotter (Eds.), *Papers for the International Conference on Salish and Neighbo(u)ring Languages 56* (pp. 154–166). UBCWPL.

Kye, Ted (2023). *A grammar of Lushootseed: Phonetics, phonology, morphology* [Doctoral dissertation, University of Washington].

Maddieson, Ian (1984). *Patterns of sounds*. Cambridge University Press.

Percival, Maida (2024). Production and perception of ejective stops across languages. [Doctoral Dissertation, University of Toronto].

R Core Team. (2017-2024). R: A language and environment for statistical computing. http://www.R-project.org/. Accessed 2023/11/15.

Warner, Natasha (1996). Acoustic characteristics of ejectives in Ingush. *Proceedings of the International Conference on Spoken Language Processing*, 1525–1528.

Wright, Richard, Hargus, Sharon & Davis, Katharine (2002). On the categorization of ejectives: Data from Witsuwit'en. *Journal of the International Phonetic Association*, 32(01), 43–77.

# Hiatus resolution in sequences of identical vowels in Russian

Yury Makarov[1][2], Sofia Sedunova[3]

*[1]University of Cambridge, [2]Vinogradov Russian Language Institute, RAS /*
*[3]HSE University*

im562@cam.ac.uk, sonya@pamiri.online

**Abstract**

This study is dedicated to hiatus resolution in sequences of identical vowels in Russian. Using corpus data, we present our preliminary findings, showing that Russian speakers employ at least three strategies of hiatus resolution: (I) glottal stop insertion / glottalisation; (II) reduction to a short vowel; and (III) producing a long vowel. While Strategy II along with Strategy IV, the maintenance of the vowel sequence, have been described in literature, the latter is often illustrated with problematic transcriptions like [aa], which are not substantiated with any instrumental data and, as we show, are theoretically problematic. Strategies I and II, though, seem to have been overlooked by the accounts of Russian phonetics. In the proposed talk, we will discuss each of the mentioned strategies, illustrating them with examples and providing the corresponding statistical analysis. Furthermore, we examine how contextual factors such as stress and prosodic boundaries influence the choice of strategies.

**Keywords:** hiatus, glottal stop, glottalisation, transcription, Russian

## 1. Introduction

The scope of this study includes Russian words like оазис /oˈazʲisˠ/ 'oasis' or Валаам/vˠalˠaˈam/ 'Valaam (island),' containing sequences of identical1 vowels. Although the observed hiatus and strategies of its resolution seem to be an important subject for both phonological typology and Russian grammar, these topics have been largely overlooked. To address this gap, we selected 42 tokens with an identical vowel hiatus illustrating different phonetic contexts and grapheme sequences. Following the analysis of the compiled corpus of their usage, we identified three strategies of hiatus resolution: (I) glottal stop insertion / glottalisation, e.g., /vˠalˠaˈam/ [vˠəlˠaˈʔamˠ]; (II) reduction to a single vowel, e.g., [vˠəˈlˠamˠ]; and (III) a long vowel, e.g., [vˠəˈlˠaːmˠ]. Figure 1 illustrates these strategies. While Strategy II has been mentioned in the literature (Avanesov 1974: 107; Kalenchuk et al. 2017: 944; inter alia), the other two strategies seem to have remained undescribed.

Interestingly enough, another strategy (Strategy IV), involving the preservation of the identical vowel sequence, is found in some sources (ibid.; Panov 1967: 63; Kalenchuk et al. 2017: 943). Usually illustrated by transcriptions like [aˈa], it remains extremely ambiguous and theoretically problematic. If the vowels were indeed identical and acoustically unseparated, following the basic principles of phonetic transcription would suggest using [aː] instead. On the contrary, if the two vowels are identical but separated by some acoustic phenomenon, the phenomenon in question must be clearly indicated in the transcription.2 The only explicit statement regarding Strategy IV is found in Jones & Ward (1969: 213):

By saying that a sound is doubled we mean that it is repeated without any pause in its continuity. The effect of repetition is obtained not by stopping the sound altogether, but by diminishing the force of the breath in the

---

1 In some cases, the vowels become identical as a result of unstressed vowel alternation rules ("reduction"), cf. /oˈazʲisˠ/ 'oasis'.

2 One could suggest that [aˈa] designates the syllabic boundary (words like Valaam can indeed be regarded as trisyllabic), but this is not explicitly stated in the referenced works and is still

an extremely ambiguous way of transcribing the contexts in question.middle of it. Thus a doubled a (written phonetically aa) is distinct from a long a:; also from two a's separated by a pause (a a), and from two a's separated by a glottal stop (aʔa), etc., etc. The difference between long sounds and doubled sounds is only important in the case of vowels. Doubled vowels constitute two syllables, and may therefore be treated differently from long vowels in the matters of stress and intonation.

The nature of 'the force of the breath,' however, remains unexplored, and it is unclear what it correlates with when it comes to acoustics. According to our results, [aˈa] (Strategy IV) is in fact a monophthong.

## 2. Methodology

We compiled and manually annotated a corpus of 805 usages of 42 tokens containing the hiatus in question. Each token is illustrated by several examples, randomly selected using subtitle search services on YouTube. While this method does not ensure sample representativeness, it allows us to demonstrate the existence of a strategy and impressionistically assess its frequency. In descriptions of the studied vowel clusters, stress and the presence of morpheme/word boundaries are considered additional sources of variability (Kalenchuk et al. 2017; Avanesov 1974). Such a classification of contexts appears primarily motivated by an interest in the quality of the allophones used. In contrast, our study examines the relationship between context type and the choice of strategy for resolving vowel hiatus.

## 3. Results

Preliminary findings suggest that the distribution of the strategies depends on stress placement and the type of boundary. For instance, Strategy I is more common when the second vowel is stressed and a morpheme boundary is present. In contrast, Strategy II is rarely observed under these conditions and is almost never found across word boundaries. Strategy III has been described as a normative way to produce sequences of identical vowels in Russian (see references for Strategy IV), which stands true in our data.
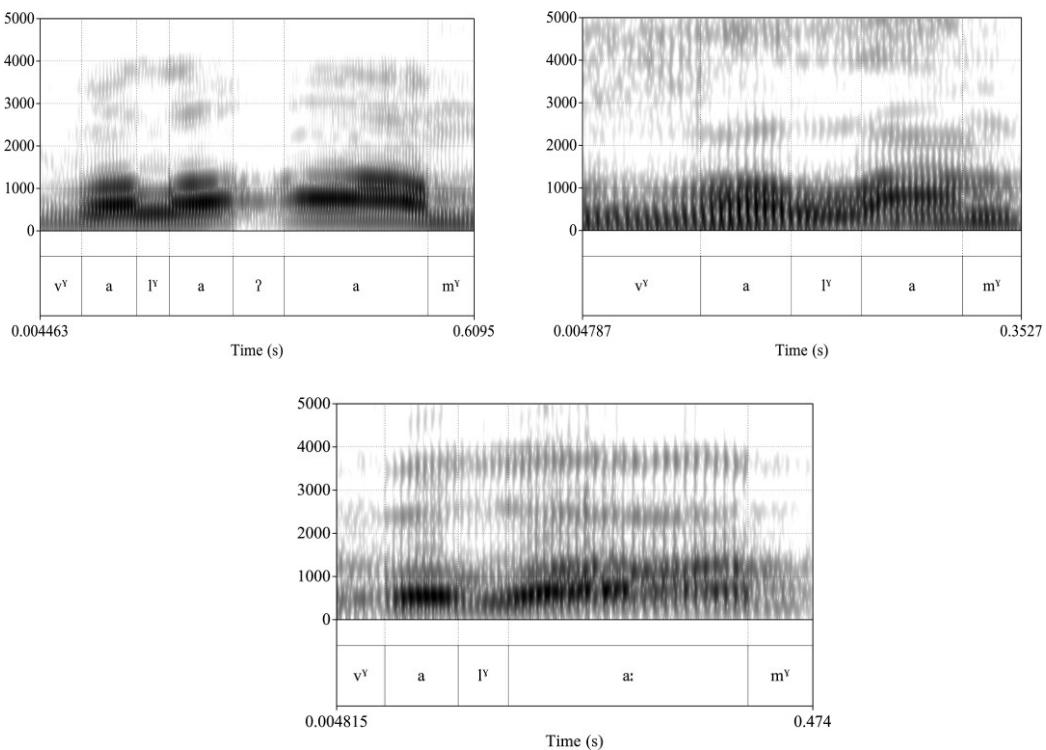


Figure 1. Russian /valaˈam/ 'Valaam (island)' produced using Strategies I–III. The vertical axis indicates frequency in Hz.

## 4. References

Avanesov, R. I. (1974). Russkaya literaturnaya i dialektnaya fonetika [Russian Literary and Dialectal Phonetics]. Prosveshchenie.

Kalenchuk, M. L., Kasatkin, L. L. & Kasatkina, R. F. (2017). Bol'shoy orfoepicheskiy slovar' russkogo yazyka [Great Orthoepic Dictionary of the Russian Language]. AST-PRESS KNIGA.

Panov, M. V. (1967). Russkaya fonetika [Russian Phonetics]. Prosveshchenie.

Jones, D., & Ward, D. (1969). The Phonetics of Russian. Cambridge University Press.

# Beyond interdentalization: Coda /k/ realizations in Basque Spanish

Carolina González, Gabrielle Isgar
Florida State University
cgonzalez3@fsu.edu, gisgar@fsu.edu

**Abstract**

In North-Central Spain, one of the most frequent variants of voiceless stops in coda are interdental fricatives, except in the Basque Country, where they are attested only in 15% of realizations. The present study investigates variation in the realization of coda /k/ in a large dataset from the Basque Country. It also examines the impact of segmental factors reported to be significant in the literature, particularly in connection to potential manner dissimilation and place assimilation to the following consonant.

Reading and interview data was obtained from 18 native speakers of Basque Spanish. Manner, place and voicing of coda /k/ (C1) were analysed acoustically using waveforms, spectrograms, and spectra. Chi-square tests, z-scores, and Bonferroni post-hoc tests were used to test the effect of stress, word length, vowel height and advancement, and manner and place of the following consonant (C2). Preliminary results of 167 words from 10 participants' interviews show that most C1 realizations involve stops and deletion, with pervasive voicelessness. The most common places of articulation are velar, interdental, and glottal. Word length impacted manner with fricatives significantly more frequent in short words (p = 0.005). C2 manner also influenced C1 manner, with deleted outcomes significantly more likely before fricatives, and stops before other stops and nasals (p < 0.001). Vowel height significantly affected place, with velars more frequent after non-high vowels and interdentals after high vowels (p = 0.001). In addition, C2 manner significantly affected voicing, with voiceless realizations more frequent before stops and voiced ones before nasals (p = 0.018). Some of the variation observed is durational: deletion occurs in unstressed positions and long words, which have shorter syllables; while fricatives, which are long, pattern with shorter words. Our presentation will include data from 8 additional participants and from the reading task, and will compare our results to those reported previously for Basque Spanish, and to those in geographically adjacent areas, including Logroño and Santander.

**Keywords:** interdentalization, coda, stops, fricatives, Basque Spanish

## 1. Introduction and background

Across Spanish dialects, voiceless stops in coda position tend to be realized as stops or deleted (Navarro Tomás 1918, Quilis 1993, Hualde 2014). In North-Central Spain, including Madrid, Castilla León, Asturias, Cantabria and Navarra, their pronunciation as interdental fricatives ('interdentalization'), is frequent, particularly for coda /k/, encompassing 43% to 95% of realizations (Quilis 1963; Martínez Martín 1983; Antón 1994, 1998; Peña Arce 2020 a, b; Pérez Castillejo 2022; cf. García de León 2015). In the Spanish spoken in the Basque Country, interdentalization of coda /k/ is also attested, but it appears to be less frequent (~15% of cases) according to a small study based on a reading task (Barbero & González 2015). The present study investigates the realization of coda /k/ in a much larger dataset from the Basque Country that includes read and interview data. It also examines the impact of segmental factors that are reported to be significant in the literature, including stress, word length, vowel height and advancement, and manner and place of the following consonant, particularly in connection to potential manner dissimilation and place assimilation to the following consonant (González 2008).

## 2. Methodology

The study examines words with word-medial coda /k/, such as doctor, acne, or electricidad in reading and interview data from 18 native speakers of Spanish from the Basque Country. The manner, place and voicing of coda /k/ were analysed acoustically in Praat using waveforms, spectrograms, and LPC spectra. The analysis of acoustic cues draws from Asensi et al. (1997), Barreiro Bilbao (2000) and Martínez Celdrán & Fernández-Planas (2013).

The effect of stress, word length, vowel height and advancement, and manner and place of the following consonant (C2) were analysed statistically with chi-square tests, z-scores, and Bonferroni post-hoc tests. After Barbero & González (2015), it was hypothesized that coda /k/ would show ample variation, with deletion and stops being the primary realizations. It was also hypothesized that place assimilation would take place, and that deletion and fricatives would be more likely before stops than in other contexts, showing manner dissimilation (González 2008).

## 3. Results

Preliminary results of 167 words from 10 participants' interviews show that most realizations involve stops and deletion, with voiceless productions being pervasive. The most common place was velar, followed by interdental and glottal (Figure 1). Interdental fricatives are attested in only 5% of instances.
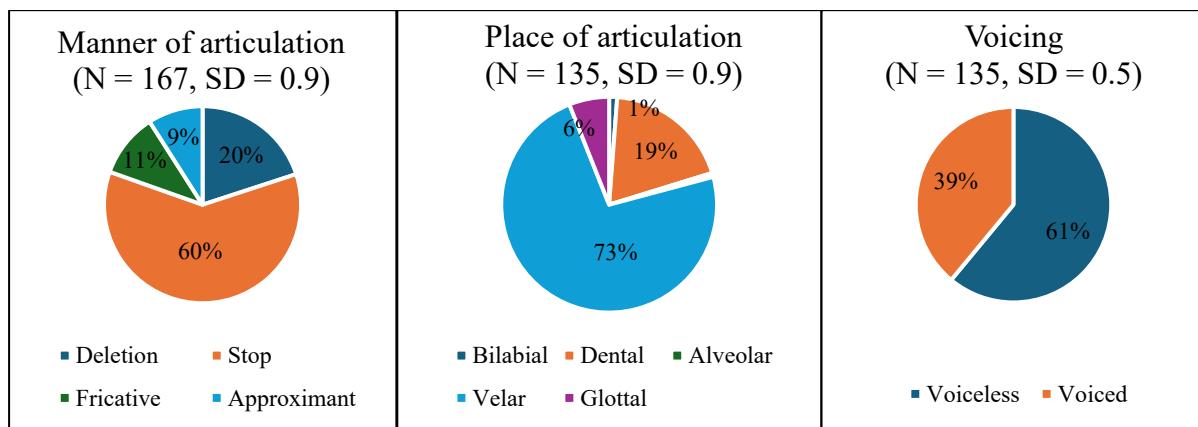


Figure 1. Distribution (%) of /k/ realizations in terms of manner, place, and voicing: interview data.

Word length impacted manner (p = 0.005), with fricatives significantly more frequent in short words (27%, z = 3.1). C2 manner also influenced C1 manner (p < 0.001), with deleted outcomes significantly more likely before fricatives (57%, z = 6.4) and stops before other stops (68%, z = 2.4) and nasals (68%, z = 0.8).

Vowel height significantly affected place (p = 0.001), with velars more frequent after non-high vowels (77%, z = 3.2), and interdentals after high vowels (70%, z = 4.4). In addition, C2 manner significantly affected voicing (p = 0.018), with voiceless realizations more frequent before stops (67%, z = 2.2), and voiced ones before nasals (62%, z = 2.8).

## 4. Discussion and conclusion

Our data supports a lower incidence of interdentalization for coda /k/ in Basque Spanish than reported in Barbero & González (2015). The interview data analysed so far tentatively supports manner assimilation rather than dissimilation, since stop realizations of coda /k/ are significantly more frequent before other stops. Our data also supports assimilation from the preceding vowel rather than to the following consonant, contra González (2008). Some of the variation observed for coda /k/ is durational: deletion occurs in unstressed positions and long words, which tend to have shorter syllables. On the other hand, fricatives, which are long,

pattern with shorter words. Our presentation will include data from 8 additional participants and from the reading task, and will compare our results to those reported previously for Basque Spanish and to those in geographically adjacent areas, including Logroño and Santander.

## 5. References

Antón, M. M. (1994). Sociolinguistic Aspects of Post-nuclear Phonological Phenomena in Asturian. Ph.D. thesis, University of Massachussets, Amherst.

Antón, M. M. (1998). Del uso sociolingüístico de las oclusivas posnucleares en el español peninsular norteño. Hispania, 81(4), 949-958.

Asensi, L., Portolés, S., & del Río, A. (1997). Barra de explosión, VOT y frecuencia de las oclusivas sordas del castellano. Journal of Experimental Phonetics, 9, 221-242.

Barbero, N. & González, C. (2015). Acoustic analysis of syllable-final /k/ in Northern Peninsular Spanish. In J. Romero & M. Rieira (eds.), The phonetics-phonology interface. Representations and methodology. Amsterdam: John Benjamins. 151-169.

Barreiro Bilbao, S. (2000). Análisis acústico comparado de las fricativas castellanas no sibilantes en realizaciones aisladas. Contextos XVII-XVIII. 243-260.

García de León, C. L. (2015). Variación en la secuencia /kt/ en español centropeninsular: Estudio sincrónico y aplicaciones diacrónicas. Loquens, 2(1), e016.

González, C. (2008). Assimilation and dissimilation of syllable-final /k/ in North-Central Spanish. In J. Bruhn de Garavito & E. Valenzuela (eds.), Selected proceedings of the 10th Hispanic Linguistics Symposium. Cascadilla Proceedings Project. 170–183.

Hualde, J. I. (2014). Los sonidos del español. Cambridge: CUP.

Martínez-Celdrán, E. & Fernández-Planas (2013). Manual de fonética española: Articulaciones y sonidos del español. Barcelona: Ariel (2nd ed, 1st ed. 2007).

Martínez Martín, F. M. (1983). Fonética y sociolingüística en la ciudad de Burgos. Madrid: CSIC.

Navarro Tomás, T. (1996 [1918]). Manual de pronunciación española. Madrid: CSIC.

Peña Arce, J. (2020a). Estudio sociolingüístico de la interdentalización del /k/, dentro del grupo /kt/, en el español hablado en la ciudad de Logroño. Cuadernos De Investigación Filológica 47. 141–161.

Peña Arce, J. (2020b). Estudio sociolingüístico de la interdentalización del /k/, dentro del grupo /kt/, en el español hablado en la ciudad de Santander (España). Philologica Canariensia 26: 20-34.

Pérez Castillejo, S. (2022). A usage-based analysis of the variable production of /k/ and /d/ as interdental fricatives. In M. Díaz-Campos (ed.), The Routledge Handbook of Variationist Approaches to Spanish. 66-79.

Quilis, A. (1964). Description phonétique du parler madrilène actuel. Phonetica 12, 19-24.

Quilis, A. (1993). Tratado de fonología y fonética españolas. Madrid: Gredos.

# Aspirated stops in Andalusian Spanish: a pan-Andalusian perspective

Paul O'Neill
*Ludwig-Maximilians-Universität, Munich*
p.oneill@lmu.de

**Abstract**

This study examines the realisation of coda /s/ before voiceless stops in Andalusian Spanish from a pan-Andalusian perspective, focusing particularly on both word internal and external sandhi contexts. Previous research has shown that aspiration of coda /s/ may undergo metathesis, producing post-aspiration and, in some areas, lengthened stop closures. However, most existing accounts focus exclusively on /s/ before voiceless stops in word-internal position and draw primarily on data from Western Andalusia. As a result, our understanding of the full range of gestural configurations involved in these processes remains incomplete. This paper presents an analysis based on comparative data from four Andalusian cities—Seville, Cádiz, Granada and Almería—covering coda /s/ in both word-internal and external sandhi contexts and with different morphological values of /s/. Across the dataset, different findings emerge depending on (a) word-internal vs external sandhi contexts and (b) the morphological value of /s/. Word-internally, the different values of coda /s/ typically show overlap between the glottal gesture associated with aspiration and the oral gesture of the following consonant. What differs across tokens, speakers and varieties is not the presence of overlap but rather the trade-off between pre-aspiration, post-aspiration and stop-closure duration. The analysis sheds new light on the phonetic motivation of metathesised or post-aspirated realisations and the phonological status of /s/. Instead of a single developmental pathway, the data suggest multiple co-existing gestural strategies that speakers employ with differing degrees of stability. By integrating cross-varietal evidence and by considering contexts beyond word-internal position, this study offers a more comprehensive account of the gestural dynamics involved in coda /s/ aspiration and contributes to our understanding of how phonetic variation may lead to divergent phonologisation pathways across a single dialect region

**Keywords:** metathesis, phonetic variation, Andalusian Spanish, aspirated stops

## 1. Introduction and background

Recently, aspiration of coda /s/ before voiceless occlusives in Andalusian Spanish has attracted substantial academic attention due to its various phonetic realisations. The most notable of these is the lack of phonetic cues of aspiration after the vowel and, in its stead, the most frequent pronunciation is a voiceless aspirated stop, which can also be elongated. There has been much debate on the phonetic motivation and phonological status of such sound changes. Regarding the latter, Torrerira (2007b, 2007a, 2012) has argued that these post-aspirated pronunciations are not intended by speakers but are the result of purely phonetic processes relating to the synchronisation and overlap of articulatory gestures (see also Gilbert (2022, 2023) for defense of underlying /s/). However, others have argued that at least some, if not the majority, of post-aspirated pronunciations are intended by speakers (O'Neill 2010, 2009; Parrell 2012). All scholars agree, however, that there is a sound change in progress in Andalusian Spanish, resulting in a process whereby the historic effect of /s/ aspiration is undergoing a metathesis and producing post-aspirated stops, which also have a tendency to be elongated in EAS ((Gerfen 2002; Moya Corral 2007; O'Neill 2010, 2009; Parrell 2012; Ruch 2012; Ruch and Harrington 2014; Torrerira 2007b, 2007a, 2012; Torrerira and Ernestus 2011). Regarding the phonetic motivation of the sound changes, issues have centered around whether an elongated consonant [paʰt·a] [pat:a] (due to the overlap between the glottal gesture for the aspiration and the occlusive gesture) is a necessary initial step, or not, for the triggering of post aspiration via

increased air-pressure during the closure and concomitant early release of the stop (O'Neill 2010, 2009; Torrerira 2007b, 2007a, 2012; Torrerira and Ernestus 2011; Ruch 2012; Ruch and Harrington 2014). Alternatively, it has been defended that post-aspirated stops could have emerged via the reorganization of the glottal spreading gesture for /s/ and the oral closure gesture for the stop, whereby instead of the gestures ocurring sequentially, they come to be synchronised with each other (Parrell, 2012). The different explanations make different predictions regarding the trade-offs between durations of pre-aspiration, post-aspiration and the occlusive closure gesture. However, different experiments produce different and, at times, contrasting results (see Ruch & Harrington 2014 and Ruch & Peters 2016 for comprehensive overviews). Moreover, a recent study (O'Neill 2024) has noted that, for the speech of Seville, whilst the effects of aspiration on voiceless occlusives is robust for word-internal coda /s/, when /s/ is a marker of plural on nouns or 2SG on verbs there are no phonetic cues for /s/ suggesting the absence of /s/ in these context. However, in the speech of Granada the cues for /s/ were similar across words as in word-internal position.

## 2. Methodology

The data for this experiment came from a corpus of recordings collected by the author in 2005. The recordings were of 4 subjects (2 male and 2 female) from the capitals of Almería, Cádiz, Granada and Sevilla. The corpus contained words containing the sequences V[owel]S[ibilant]O[cclusive]V[owel] (VSOV) in both word internal position and across words (VS#OV) and their minimal or near minimal pairs (VOV and V#OV sequences respectively). Subjects were presented with photographs depicting the words in the corpus, along with a number of control words that did not contain orthographic <s>; participants were requested to name what they saw[1]. In word internal position the following carrier phrase was used *dame un ____ pa[ra] mí* 'give me a ____ for me', whilst in the external sandhi contexts there were three different types of carrier phrase depending on the type of <s>. The first type of <s> was the marker of 2SG (S$_{2SG}$); its pronunciations were contrasted with phrases with 3SG verb forms in which there was no <s> (VS$_{2SG}$#OV vs. V#OV contexts). Speakers were required to name the images they saw in the photographs in the phrase *tu siempre comes ____* 'you always eat ____' and then again in the phrase *él siempre come ____* 'you always eat ____'. The second type of <s> tested was that in the word *dos* (S$_{DOS}$); speakers were required to name what they saw in the photographs in the phrase *digo dos ____ pa[ra] ti'* 'I say two ____ for you' and then again in the phrase *digo un/una ____ pa[ra] ti'* 'I say a ____ for you' (VS$_{dos}$#OV vs. V#OV contexts)[2]. These same carrier phrases were also used to examine the effects of the presence/absence of the third type of <s>, the nominal plural marker (S$_{PLURAL}$) and its effects in external sandhi exclusively on the following /p/ consonant in the phrase *digo dos ____ pa[ra] ti* (VS$_{PLURAL}$#C$_{/p/}$V vs. V#C$_{/p/}$V contexts)

## 3. Results

The results, which still need to be fully analysed, reveal differences with respect to the existence of trade-off relationships between glottal and occlusive gestures (e.g. these are present in Granada and Cádiz but not present in Almería and Seville). Moreover, when /s/ is a marker of plural on nouns and 2SG on verbs there is a tendency in Seville and Cádiz for there to be no phonetic cues for /s/, suggesting the loss of /s/. However, there is inter-individual variation, with certain individuals in both cities showing retention of /s/.

---

1 For those words in the corpus which it was impossible to solicit via the naming of images, subjects read the words embedded in a carrier phrase, from a card.

2 Note that speakers were presented with two images which corresponded to the word in first phrase and just one image in the second phrase to ensure that they produced the singular and plural forms

## 4. Discussion and conclusion

The analysis sheds new light on the phonetic motivation of metathesised realisations and the phonological status of /s/. Regarding the former, instead of a single developmental pathway towards metathesis, the data suggest multiple co-existing gestural strategies that speakers employ with differing degrees of stability. Regarding the phonological status of /s/, the lack of phonetic cues for aspiration in external-sandhi contexts in the cities of Seville and Cádiz vs. the presence in word-internal contexts suggests that /s/ is being lost in specific external-sandhi contexts in these cities. However, the robust phonetic cues for /s/ both word externally and across word-boundaries in Granada and Almería suggest that the metathesised realisations due to /s/ aspiration were, in origin, a neogrammarian sound change.

## 5. References

Gerfen, C. (2002). ' Andalusian codas. ', *Probus*, 14: 247–77.

Gilbert, M. B. (2022). 'An experimental and formal investigation of Sevillian Spanish metathesis', New York University.

Gilbert, M. B. (2023). 'Testing for underlying representations: Segments and clusters in Sevillian Spanish', *Natural Language & Linguistic Theory*, 2: 493-531.

Moya Corral, J. A. (2007). 'Noticia de un sonido emergente: la africada dental procedente del grupo -st- en Andalucía. ', *Revista de Filología de la Universidad de La Laguna*, 25: 457–65.

O'Neill, P. (2009). 'The effect of s-aspiration on occlusives in Andalusian Spanish', *Oxford University Working Papers in Linguistics, Philology & Phonetics*, 12: 73-86.

O'Neill, P. (2010). 'Variación y cambio en las consonantes oclusivas del español de Andalucía. ', *Estudios de Fonética Experimental*, XIX: 11–41.

O'Neill, P. (2024). 'La aspiración de /s/ implosiva en el interior de palabra y en la fonética sintáctica: un estudio experimental de cuatro ciudades andaluzas', *Variación*, 1.

Parrell, B. (2012). 'The role of gestural phasing in Western Andalusian Spanish aspiration. ', *Journal of Phonetics*, 40: 37–45.

Ruch, H. (2012). 'Affrication of /st/ in Western Andalusian Spanish: Variation and change from a sociophonetic point of view.' in S Calamai, C Celata and L  Ciucci (eds.), *Proceedings of "Sociophonetics, at the crossroads of speech variation, processing and communication"* (Edizioni della Normale: Pisa).

Ruch, H. & J. Harrington (2014). 'Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish', *Journal of Phonetics*, 45: 12-25.

Ruch, H. & S. Peters (2016). 'On the Origin of Post-Aspirated Stops: Production and Perception of /s/ + Voiceless Stop Sequences in Andalusian Spanish', *Laboratory Phonology: Journal Of The Association For Laboratory Phonology*, 7.

Torrerira, F. (2007a). 'Coarticulation between aspirated-s and voiceless stops in Spanish: An interdialectal comparison.' in N Sagarra and A. J.  Toribio (eds.), *Selected proceedings of the 9th hispanic linguistics symposium* (Cascadilla Proceedings: Somerville, MA).

Torreira, F. (2007b). 'Pre- and postaspirated stops in Andalusian Spanish.' in P. Prieto, P Mascaró and M. J.  Solé (eds.), *Segmental and prosodic issues in Romance Phonology* (Benjamins: Amsterdam).

Torreira, F. (2012). 'Investigating the nature of aspirated stops in Western Andalusian Spanish', *Journal of the International Phonetic Association*, 42: 49–63.

Torrerira, F. & M. Ernestus (2011). 'Realization of voiceless stops and vowels in conversational French and Spanish', *Laboratory Phonology*, 2: 331–53.

# Using lip tracking to disentangle gradient articulatory changes from segmental deletion in American English

Karolina Broś, Peter A. Krause

*University of Warsaw, CSU Channel Islands*

k.bros@uw.edu.pl

## Abstract

American English is well-known for optional /t d/ deletion in final clusters. The nature of the process is still not fully accounted for despite numerous studies because deletion can be incomplete, as evidenced by articulatory studies. At the same time heritage Spanish speakers are reported to have high rates of deletion, yet their articulatory patterns have not been studied. In this paper, we use motion capture to look at /t d/ deletion and weakening in sequences of sounds containing flanking bilabials. The results show higher rates of deletion in bilinguals and the existence of fully deleted segments, as well as weakened variants of different types.

**Keywords:** heritage Spanish, American English, weakening, motion capture

## 1. Introduction and background

American English is well-known for optional /t d/ deletion in clusters, as in best [bes] man, kept [khep] going or whipped [wɪp] butter (Coetzee 2004, Walker 2012). However, different studies have led to varying conclusions concerning the conditioning of the process and its categoricity. Even in the absence of an acoustic trace of /t d/, articulatory gestures related to the alveolar constriction may still be there (Browman & Goldstein 1990; Purse & Turk 2019), although they may be of a smaller magnitude, which questions the completeness of deletion from a phonological standpoint. The disadvantage of articulatory studies on the subject is that they only elicit a small number of tokens in a few speakers. Furthermore, research has shown that separate dialectal features have formed among bilingual, fully English-proficient representatives of the Latine population in the US. This includes Chicano English (Santa Ana, 1991), which has a particularly high rate of /t d/ deletion. As Spanish syllable structure does not allow complex codas at the end of a word, and many Spanish varieties weaken or fully delete even simple codas in this position (Hualde, 2005), it is only natural for Spanish speakers to simplify final clusters. Thus, if there is an influence or interference of Spanish on the English of bilingual Spanish-English speakers, we expect more frequent /t d/ deletion and/or different articulatory patterns than in American monolinguals.

The aim of the present study was to look at /t d/ segments trapped in between consonants vs a neutral context to look at the interplay of acoustics with articulation. To elicit a larger data sample, the motion capture technique was employed following previous work (e.g. Broś & Krause 2024). The overarching goal was to determine whether there are differences in the articulatory dynamics of /t d/ deletion between heritage Spanish and monolingual American English speakers and if so, whether we can reliably distinguish between the different variants or the observed differences are gradient in nature.

## 2. Methodology

Twenty American English speakers (10 monoliguals and 10 heritage Spanish speakers) from California were recorded while producing 320 sentences each, with regular verbs ending in a bilabial sound embedded in them, produced either in the present or the past tense and followed by a bilabial or a vowel (e.g. stop making, stop eating, stopped making, stopped eating). Recordings were made using AAA software. This enabled the analysis of pronounced vs. unpronounced /t d/ segments under the assumption that the initial lip closure for the bilabial

sound must be interrupted to produce an alveolar stop and will not be interrupted or will only be weakened in the case of lenited or unpronounced /t d/.

## 2.1. Data Analysis

Relevant sound sequences were transcribed and /t d/ acoustic deletions were annotated in Praat. Lip video data were preprocessed in AAA using the DLC algorithm tracking lip contours over the upper and lower lip (Fig. 1). Sequence trajectories were then extracted using two metrics: lip aperture (distance between the inner lower and inner upper lip) and lip compression (distance between the inner and outer lower lip). Measurements were taken over 11 equally spaced timepoints in each sound sequence and time-normalized. More specifically, bilabial-alveolar-bilabial, bilabial-bilabial (control), bilabial-alveolar-vowel (vowel control) and bilabial-vowel (second vowel control). The expectation was that a greater value of lip aperture means lip opening for the production of the alveolar stop in between other sounds, while a greater value of lip compression means that the lips are less tense and less pressed against each other. Greater lip compression and smaller lip aperture mark no lip opening/softening between flanking sounds: /t d/ was either not pronounced or weakly pronounced. This was confronted with the acoustics (audible/inaudible /t d/) and compared with tongue gestures from simultaneously recorded ultrasound data (separate publication).
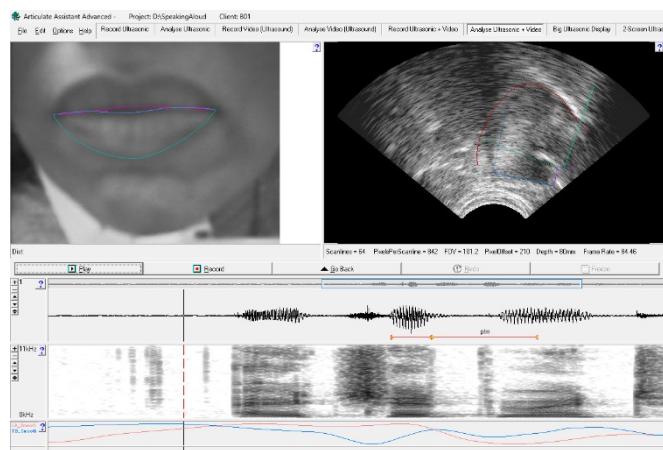


Fig. 1. Tongue and lip trajectories for the sentence containing the sequence /ptm/.

## 3. Results

The preliminary analysis of the data shows differences between acoustically audible sounds, inaudible sounds accompanied by lip aperture (incomplete deletions) and inaudible sounds without lip aperture (complete deletions). It also shows effects of the way both stops in the cluster are released. Heritage Spanish speakers show much higher rates of deletion than monolinguals, which is in line with the literature on Chicano English and may be due to the phonological properties of Spanish.
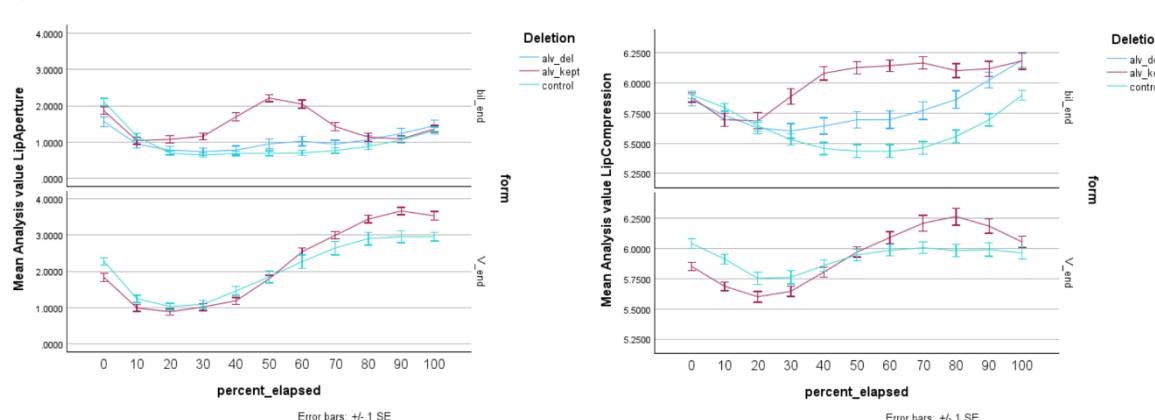


Fig 2. Lip trajectories over the sequence of interest showing lip aperture (left) and lip compression (right) on a subset of speakers

## 4. Discussion and conclusion

Lip aperture was a less consistent marker of /t d/ weakening, showing effects in some speakers but not others. Lip compression, in turn, captured weakened pronunciations across speakers, despite slightly different individual strategies. Additionally, bilabial stop weakening and cluster simplification observed especially in the heritage Spanish group reveals that there are more articulatory patterns than simply deleted, non-deleted and weakened /t d/ in these contexts. The results of the study have important implications for the way phonetic and phonological deletion should be analyzed and interpreted and for the role played by the linguistic background of a particular population.

## 5. References

Broś, Karolina & Krause, Peter A. (2024). Stop lenition in Canary Islands Spanish – a motion capture study. Laboratory Phonology 15(1). doi: https://doi.org/10.16995/labphon.9934.

Browman, Catherine P. & Goldstein, Louis. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston and M. Beckman (Eds.), Papers in laboratory phonology i: Between the grammar and physics of speech (pp. 341–397). Cambridge: Cambridge University Press.

Coetzee, Andries. (2004). What it means to be a loser: non-optimal candidates in optimality theory (Doctoral dissertation). University of Massachusetts, Amherst.

Hualde, J. I. (2005). The sounds of Spanish. Cambridge University Press.

Purse, Ruaridh & Turk, Alice. (2016). '/t d/ deletion': Articulatory gradience in variable phonology. Paper presented at LabPhon 2016.

Santa Ana, Otto (1991). Phonetic Simplification Processes in the English of the Barrio: A Cross-generational study of the Chicanos of Los Angeles (Doctoral dissertation). University of Pennsylvania.

Walker, James A. (2012). Form, function, and frequency in phonological variation. Language Variation and Change, 24, 397–415.

# Phonetic outcomes of contrast preservation in Andalusian Spanish: Focus on /s/ + voiced stop sequences

Kaitlyn Sabb, Nicholas Henriksen
*University of Michigan*
ksabb@umich.edu, nhenriks@umich.edu

## Abstract

Maintaining contrast is crucial for effective communication, yet phonological processes such as lenition can undermine these distinctions. Dispersion Theory (Flemming, 2017) suggests that languages employ various strategies to preserve contrasts that might otherwise be neutralized. In Andalusian Spanish, while studies show that post-aspiration in /s/ + /p t k/ sequences help distinguish them from phonological /p t k/, far fewer examine whether similar strategies apply to phonological /s/ + /b d g/ versus /b d g/ across its subvarieties. Here, we collected speech data from 21 speakers of Western Andalusian Spanish and 26 of Eastern Andalusian Spanish. Speakers of both subvarieties preserve the contrast between phonological intervocalic /b d g/ and /s/ + /b d g/ by strengthening stops following /s/ lenition, indicated by increased C-Duration, decreased C/V-Intensity-Ratio, and decreased closure voicing. The results support the notion that Spanish resists contrast neutralization, even amid phonological processes like lenition.

**Keywords:** lenition, contrast preservation, Andalusian Spanish, voiced stops

## 1. Introduction and background

This study examines how Western (WAS) and Eastern Andalusian Spanish (EAS) speakers produce the phonological voiced stops /b d g/ in two contexts: between vowels (i.e., phonological /b d g/) and following a lenited /s/ (i.e., phonological /s/ + /b d g/). Regarding Andalusian Spanish, spoken in southern Spain, recent literature focuses on contrast preservation in phonological /p t k/ versus /s/ + /p t k/, the latter resulting in post-aspirated voiceless stops following /s/ lenition (Henriksen et al., 2023; Ruch & Harrington, 2014). Fewer studies offer a systematic analysis of contrast preservation between phonological /b d g/ and /s/ + /b d g/ (Broś et al., 2021). Intervocalic phonological /b d g/ undergo systematic lenition in Spanish, realized as [β] [ð] [ɣ]. If this lenition pattern applies to all intervocalic surface stops, deleting /s/ in /s/ + /b d g/ sequences could obscure the contrast between underlying /b d g/ and /s/ + /b d g/—this is the focus of our study. Dispersion Theory (Flemming, 2017) provides a theoretical framework for understanding contrast maintenance: phonological systems operate under constraints that favor perceptually distinct contrasts over less distinct ones. When phonological processes threaten to neutralize contrasts by making them insufficiently distinct, languages can either neutralize the contrast or enhance the contrast through modification of secondary acoustic cues.

## 2. Methodology

Speech data were collected from 47 native speakers of Andalusian Spanish, including 21 WAS speakers from the city of Seville and 26 EAS speakers from Granada. The participants ranged from 18 to 65 with a mean age of 33.30 years (SD = 13.81). Each participant produced 120 target items embedded in carrier phrases, balanced for singular /b d g/ forms (e.g., mi bata 'my robe') and plural /s/ + /b d g/ forms (e.g., mis batas 'my robes'). We performed acoustic analysis in Praat and extracted consonant duration (C-Duration), consonant-to-vowel intensity ratio (C/V-Intensity-Ratio), and Percent Voicing of each closure. We conducted statistical analysis in R using linear mixed-effects and Beta regression models.

## 3. Results

The results indicate systematic differences between phonological contexts for C-Duration, C/V Intensity Ratio, and Percent Voicing. C-Duration values are consistently longer in the phonological /s/ + /b d g/ context compared to intervocalic phonological /b d g/, across both EAS and WAS (Fig. 1). Regarding C/V-Intensity-Ratio, we find higher values for intervocalic /b d g/ compared to /s/ + /b d g/ across all phonemes and subvarieties (Fig. 2). Exploratory cluster analysis further revealed two distinct speaker groups based on Percent Voicing, with one group (Cluster 1, Figs. 3 & 4) showing closure voicing up to 60%, and the second group (Cluster 2, Figs. 3 & 4) showing voicing values between 60% and 100%.

## 4. Discussion

The acoustic findings indicate that EAS and WAS speakers strengthen voiced stops following /s/ lenition by increasing C-Duration and decreasing C/V-Intensity Ratio. This suggests that voiced stops are longer and produced with greater oral constriction in phonological /s/ + /b d g/. Some speakers further emphasize the contrast by overall decreased voicing of the closure. Altogether, the results align with Dispersion Theory (Flemming, 2017)—when /s/ undergoes lenition before voiced stops, rather than allowing neutralization, speakers enhance secondary acoustic cues by strengthening the articulation of the voiced stops. This reflects "contextual enhancement," where a language compensates for the weakening of one contrast dimension (the presence of /s/) by enhancing another dimension (the manner of articulation of the following stop). Our findings show that these enhancement strategies occur across both WAS and EAS, suggesting that contrast maintenance is a fundamental property of Spanish phonology.
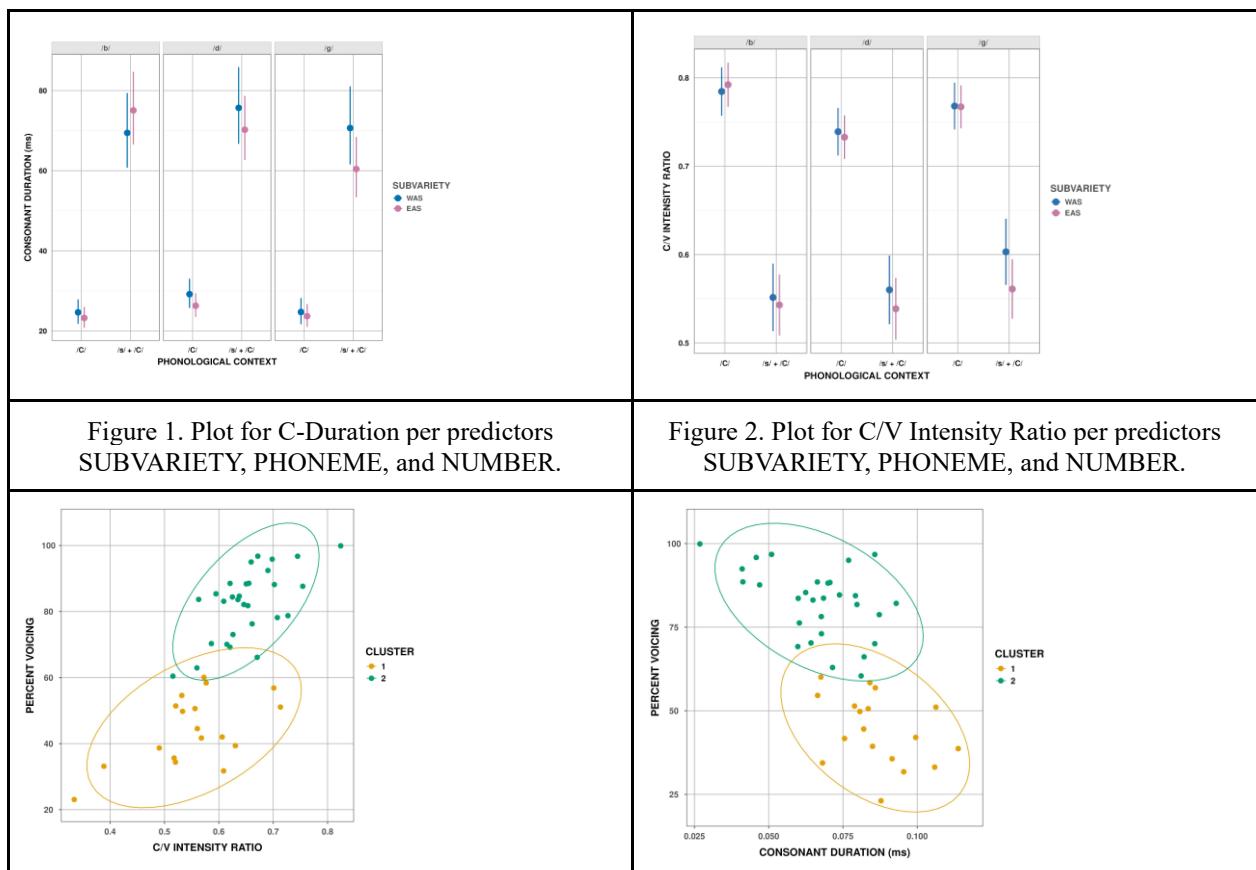


| Figure 1. Plot for C-Duration per predictors SUBVARIETY, PHONEME, and NUMBER. | Figure 2. Plot for C/V Intensity Ratio per predictors SUBVARIETY, PHONEME, and NUMBER. |

| Figure 3. Scatterplot of speaker-level mean values for the outcome measures Percent Voicing and C/V-Intensity Ratio, color-coded by cluster assignment. | Figure 4. Scatterplot of speaker-level mean values for the outcome measures Percent Voicing and C-Duration, color-coded by cluster assignment. |
|---|---|

## 5. References

Broś, K., Zygis, M, Sikorski, A. & Wołłejko, J. (2021). Phonological contrasts and gradient effects in ongoing lenition in the Spanish of Gran Canaria. Phonology, 38(1), 1–40.

Flemming, E. (2017). Dispersion Theory and Phonology. Oxford Research Encyclopedia of Linguistics.

Henriksen, N., Galvano, A. & Fischer, M. (2023). Sound change in Western Andalusian Spanish: Investigation into the actuation and propagation of post-aspiration. Journal of Phonetics, 98, 1–21.

Ruch, J. & J. Harrington (2014). Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. Journal of Phonetics, 45, 12–25.

# *Do You Hear What They Hear?* Approximants, Fricatives, and Dialectal Bias in Spanish Stop Perception

Santiago Arróniz
*University of Nevada, Reno*
sarroniz@unr.edu

## Abstract

This study examines how dialectal experience shapes the perception of fricated variants of voiced stops in Spanish. In Western Andalusian Spanish (WAS), the weakening of coda /s/ often gives rise to Compensatory Voiced Fricatives (CVFs), which differ acoustically from canonical approximants, serving as phonological compensation for an elided /s/. To assess whether this distinction is perceptually accessible across dialects, a forced-choice AX discrimination task was administered to WAS, Northern Castilian, and Mexican Spanish listeners. Stimuli included approximants, fricatives, and overt tokens of /s/ + /b d g/ in minimal pair contexts. Results showed that while all groups reliably identified overt forms, only WAS listeners consistently distinguished fricatives from approximants. Northern and Mexican listeners exhibited lower accuracy for these contrasts, especially for /d/ and /g/. These findings suggest that fine-grained perceptual distinctions are shaped by dialect-specific phonetic exposure and support the view that CVFs are undergoing phonologization within WAS. More broadly, the study highlights the role of experience in shaping phonological categories and the perception of gradient variation.

**Keywords:** phonetic variation, compensatory fricativization, Western Andalusian Spanish, speech perception, cross-dialectal confusion

## 1. Introduction and background

In Western Andalusian Spanish (WAS), the weakening of coda /s/ often leads to the emergence of fricated variants of /b d g/, known as Compensatory Voiced Fricatives (CVFs) (Romero Gallego 1995, Arróniz & Willis 2023, Arróniz 2025). These segments differ acoustically from canonical approximants and may serve as perceptual cues to plural morphology. Previous findings indicate that WAS listeners are sensitive to these distinctions, but it remains unclear whether this sensitivity reflects a unique consequence of local exposure or a more general perceptual accessibility shared across Spanish dialects.

This study compares listeners from three dialectal backgrounds—WAS, Northern Castilian, and Mexican Spanish—using an AX discrimination task. AX paradigms are designed to probe fine-grained auditory sensitivity without invoking lexical judgments, making them ideal for testing emergent contrasts such as CVFs. By examining cross-dialectal differences in the perception of approximant–fricative contrasts, the study addresses broader questions in phonological theory: how phonetic gradience becomes phonologically relevant, how perceptual categories adapt to distributional input, and how dialect-specific experience shapes phonological representations.

## 2. Methodology

To evaluate dialectal differences in the perception of CVFs, we administered the online AX discrimination task to 127 native Spanish speakers. Participants were grouped based on regional background: 92 from Western Andalusia, 23 from Northern Spain, and 12 from Central Mexico. All participants confirmed the use of headphones and a quiet environment. Ages ranged from 18 to 62 (M = 26.81).

Stimuli consisted of 54 AX pairs derived from 18 triplets representing three phonetic realizations of /b d g/: approximants, fricatives, and overt clusters with /s/. Tokens were

extracted from controlled recordings of minimal pairs (e.g., *una bota* vs. *unas botas*) produced by two Andalusian speakers.

Trials were presented in 21 randomized blocks, with each AX pair separated by 200 ms of silence. Participants judged whether the two tokens sounded the "same" or "different," focusing on phonetic form. A forced-response protocol ensured full participation. Responses exceeding 20 minutes or reported under suboptimal listening conditions were excluded from analysis.

## 3. Results

The AX discrimination task revealed systematic differences across dialect groups in how listeners perceived approximants, fricatives, and overt productions of /b d g/. While all groups consistently distinguished overt tokens from the other variants, notable variation emerged in their ability to discriminate fricatives from approximants, particularly for /d/ and /g/.

For /b/, Andalusian and Northern Castilian listeners performed well across all pairings, with high accuracy for both fricative–approximant and overt–approximant contrasts. Mexican listeners identified overt tokens with near-perfect accuracy but showed markedly reduced performance (25%) in the fricative–approximant condition, suggesting a perceptual merging of these variants in this group.

For /d/, Andalusian listeners again showed high accuracy across contrasts (76.8%). Northern listeners performed moderately on approximant–fricative pairs (58.3%), while Mexican listeners again displayed low discrimination (37.5%), indicating less perceptual separation between these variants.

The /g/ condition yielded the greatest variability. WAS listeners showed moderate success with fricative–approximant contrasts (55%), while Northern Castilian listeners and Mexican listeners struggled with the such token pairs (28.1% and 12.5% respectively).

In sum, all groups treated overt tokens as perceptually distinct, but only Andalusian listeners consistently distinguished fricative and approximant variants. These results suggest dialect-specific phonological encoding of gradient fricativization, with WAS listeners showing heightened sensitivity likely rooted in local phonetic exposure.

## 4. Discussion and conclusion

The AX task revealed clear perceptual asymmetries across dialects in the treatment of fricated variants of /b d g/ following coda /s/ weakening. While all groups consistently distinguished overt tokens, only WAS listeners reliably differentiated approximants from fricatives across all consonants. This suggests that fine-grained phonetic contrasts are not equally accessible across dialects, but shaped by exposure to local phonetic distributions.

For Northern Castilian and Mexican listeners, reduced accuracy in distinguishing fricatives from approximants—particularly for /d/ and /g/—suggests that these variants are not perceptually encoded as distinct categories in the absence of phonological relevance. In contrast, WAS listeners' consistent sensitivity supports the hypothesis that CVFs are becoming phonologized within this variety's system.

These findings align with models in which category structure emerges from distributional regularities rather than fixed acoustic boundaries. The WAS case illustrates how phonetic compensation for segmental loss can give rise to perceptual reorganization and potential category emergence. More broadly, this study highlights how dialect-specific experience guides the interpretation of gradient variation, offering insight into how phonologization unfolds in real time. Future work will explore how these perceptual patterns interact with lexical processing and morphological interpretation.

## 5. References

Arróniz, S. (2025). *Compensatory cues of /s/ reduction in voiced stop consonants in Western Andalusian Spanish*. Bloomington, IN: Indiana University dissertation.

Arróniz, S. & Willis, E.W. (2023). From one fricative to another: A perception experiment of coda /s/ and voiced phonemic stops as fricatives in Western Andalusian Spanish. *Internacional de Fonética Experimental CIFE 2023*.

Bishop, J. B. (2007). Incomplete neutralization in Eastern Andalusian Spanish: Perceptual consequences of durational differences involved in s-aspiration. *Proceedings of the International Congress of Phonetic Sciences*, 16, 1765–1768.

Boomershine, A. & Hall, K. C. & Hume, E. & Johnson, K. (2008). The impact of allophony versus contrast on speech perception. In Peter Avery, Elizabeth Hume & Keith Johnson (eds.), *Contrast in phonology* (143–172). Berlin: Mouton de Gruyter.

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.

Galarza, I. & Delgado-Díaz, G. & Willis, E. W. (2014). ¿Nuevamente /s/? Una nueva mirada a la elisión de /s/ implosiva en el español de Puerto Rico. *7th Workshop on Spanish Sociolinguistics*, Madison, WI.

Martínez Celdrán, E.  (1991). Sobre la naturaleza fonética de los alófonos de "b, d, g" en español y sus distintas denominaciones. *Verba: Anuario galego de filoloxía*, 18, 235–253.

Romero Gallego, J. (1995). *Gestural organization in Spanish: An experimental study of spirantization and aspiration*. Storrs, CT: University of Connecticut dissertation.

Romero Gallego, J. (1995). *Gestural organization in Spanish: An experimental study of spirantization and aspiration* (Doctoral dissertation). University of Connecticut.

Torreira, F. (2007a). Coarticulation between aspirated-s and voiceless stops in Spanish: An interdialectal comparison. In H. Campos et al. (eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium* (113–120). Somerville, MA: Cascadilla Proceedings Project.

Torreira, F. (2007b). Pre- and postaspirated stops in Andalusian Spanish. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, 282, 67.

Torreira, F. (2012). Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association*, 42(1), 49–63.

Willis, E. W. (2006). Trill variation in Dominican Spanish: An acoustic examination and comparative analysis. In Nuria Sagarra & Almeida Jacqueline Toribio (Eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium* (121–131). Somerville, MA: Cascadilla Proceedings Project.
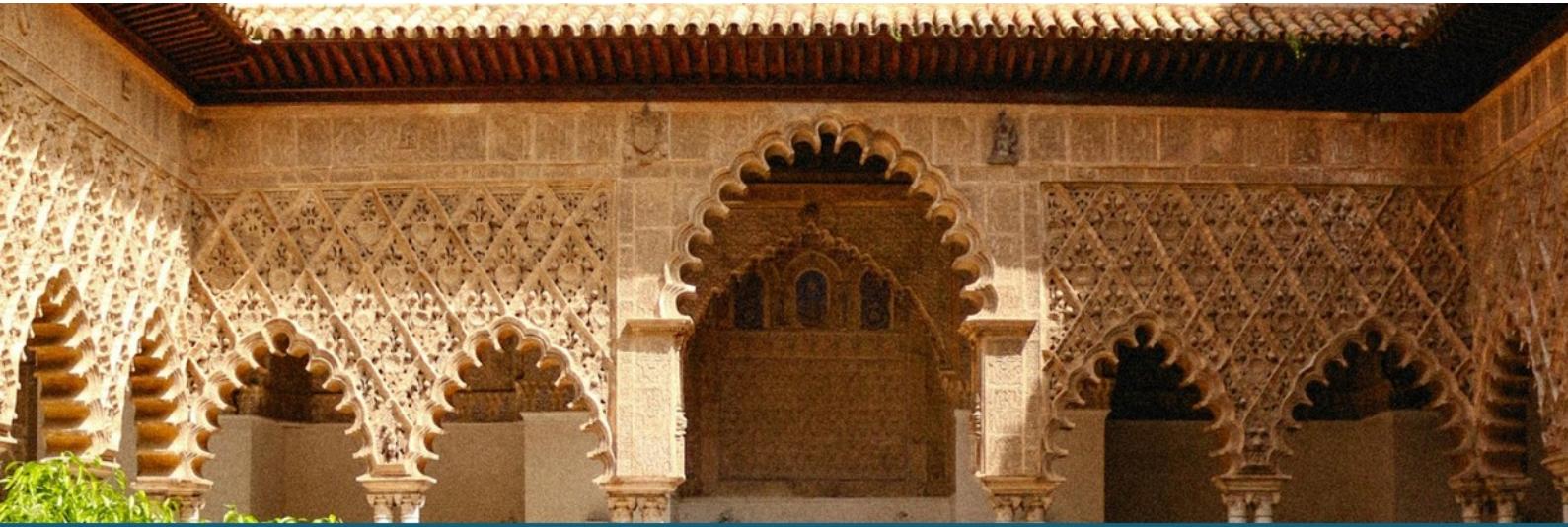
# SPEECH PERCEPTION

# Perception of Spanish Intonation Contours of Declarative Sentences by Spanish Speakers of Different Varieties

Izaro Bedialauneta Txurruka
*University of Illinois at Urbana-Champaign*
izarob2@illinois.edu

**Abstract**

Spanish declarative questions and statements often differ in prosody, though the specific intonational cues and contours vary across dialects. This study investigates how native Spanish listeners from three regional varieties, Castilian Spanish (CS), Buenos Aires Spanish (BAS), and Cibao Spanish (DRS), perceive interrogative intonation across dialects. While CS and BAS questions typically align with the universal tendencies by ending in a rising or circumflex contour, DRS questions may end in either L% or H% and not following completely the patterns predicted by the Strong Universalist Hypothesis (SUH). This variation offers the opportunity to test the perception of question intonation by same language speakers but whose language variant is different. A total of 24 native Spanish speakers (from Spain, Argentina, and the Dominican Republic) were presented with 60 auditory stimuli, consisting of full sentences as well as partial utterances containing either the first accentual phrase (high initial tone), or the second accentual phrase (the boundary tone). Participants were asked to decide whether each stimulus was a question or not. We hypothesized that listeners would rely primarily on their native prosodic cues and, in their absence, would use the universal tendencies, predicting DRS sentences as the most challenging ones. Logistic regression analysis showed no significant effect of listener group, but significant effects of variant and the accentual phrase. DRS contours were more difficult to interpret. Boundary tone served as the primary cue to interrogativity in CS and BAS, while the initial pitch height played a more prominent role in DRS. These results indicate that when perceiving nonnative intonational patterns, cue interpretation was influenced by the prosodic features most salient in each dialect and that they were not necessarily relying on their own native phonology or universal tendencies.

**Keywords:** intonation, perception, universality, Spanish variation

## 1. Introduction and background

Spanish declarative questions and statements often differ only in their prosody. For example, ¿Bebe agua? "Does he/she drink water?" and bebe agua "He/she drinks water" are lexically and syntactically identical but differ in intonation. However, intonational contours vary across Spanish-speaking regions. In Castilian Spanish (CS), neutral questions usually end with a rise (Sosa, 1999, Hualde, 2003; Estebas-Vilaplana et al., 2010), while in Buenos Aires Spanish (BAS) they tend to end with a circumflex contour (Gabriel et al. 2010). In both varieties, declaratives end with a fall. Dominican Spanish (DRS), by contrast, shows less patterns that are cross-linguistically less common: statements end in H%, and questions end in either L% or H% (Willis, 2010, Hualde et al. 2015). Thus, CS and BAS follow universal interrogation patterns (Ladd, 1981; Ohala, 1984) more closely than DRS. To date, no study has explored how native speakers perceive primary and secondary cues in the intonation of different Spanish varieties.

## 2. Methodology

We report on an experiment where participants were auditorily presented with statements and questions produced by speakers of the three Spanish varieties just described. Listeners heard 60 audio stimuli: 20 stimuli were sentences containing two accentual phrases and 40 stimuli were partial sentences that contained either the first or the second accentual phrase of the original sentences (20 each). After each stimulus, participants indicated whether it was a

question. Participants included 24 native Spanish speakers: 10 from Buenos Aires, 7 from the Dominican Republic, and 9 from Spain. We hypothesized that listeners would rely primarily on their L1, and when encountering unfamiliar contours, they would fall back on universal tendencies. We predicted DRS contours would be particularly difficult to identify. For partial stimuli, we hypothesized that the final contour would be more informative, although a high initial pitch might also cue interrogativity (Face, 2005, Face, 2007; Van Heuven, 2000).

## 3. Results
A binomial logistic regression model showed no significant differences among listener groups. However, a significant difference was found in the interpretation of complete and incomplete sentences and in the Spanish dialects of the stimulus.

## 4. Discussion and conclusion
As predicted, DRS sentences were the most challenging, whereas CS and BAS sentences were accurately identified. Listeners may have relied on universal tendencies for the identification of these contours. However, while the final part of intonational contours generally was perceived as the primary cue to interrogation, a relatively high beginning was also found to convey interrogativity, especially in the case of DRS stimuli. This suggests that participants cue preference may not be static. Instead, listeners can identify the part of the utterance in which questions diverge the most from the statements.
Keywords: intonation, perception, universality
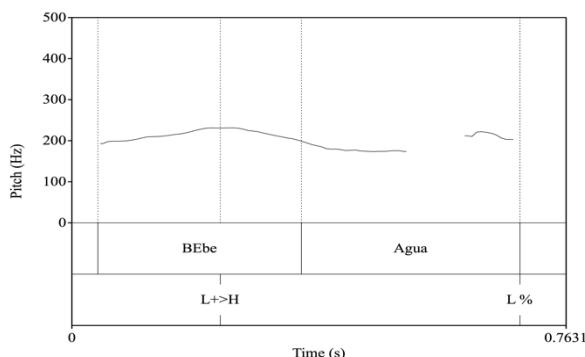

Figure 1: A statement contour in CS
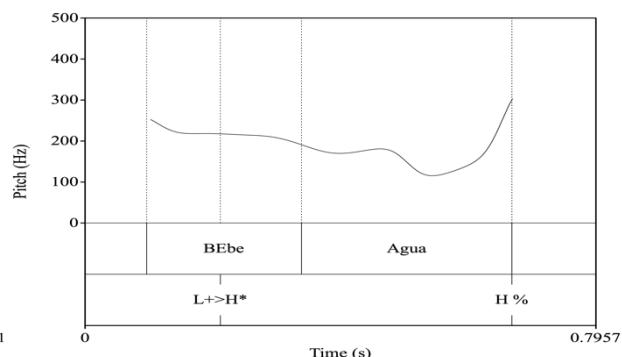

Figure 2: A question contour in CS
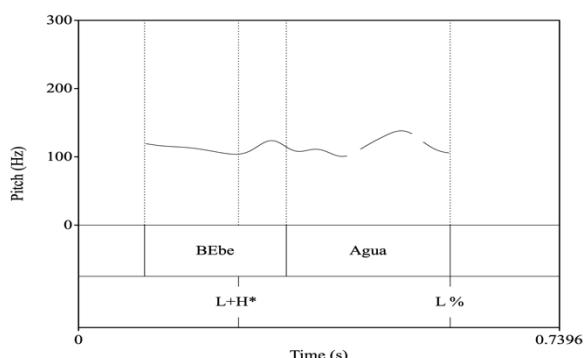

Figure 3: A statement contour in BAS
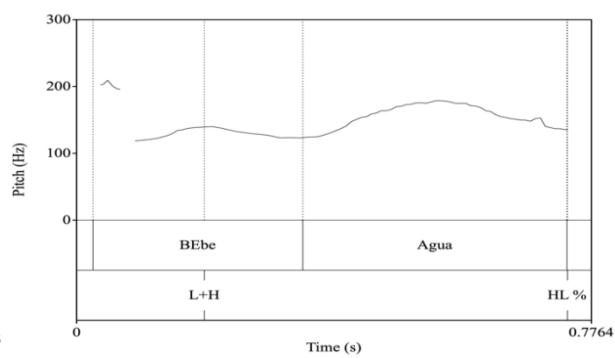

Figure 4: A question contour in BAS
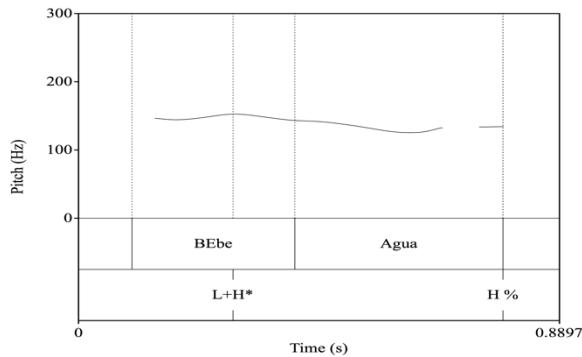
Figure 5: A statement contour in DRS
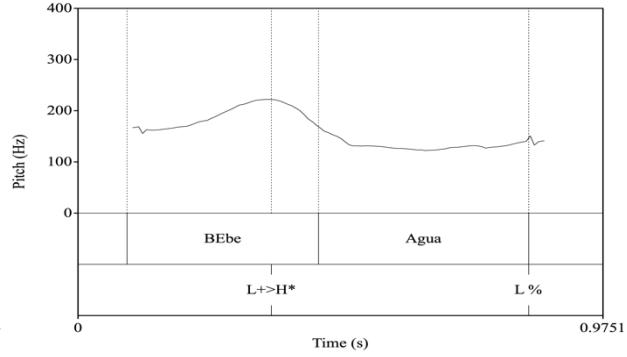

Figure 6: A question contour in DRS

## 5. References

Estebas-Vilaplana, E. & Prieto, P. (2010). Castilian Spanish intonation. Transcription of intonation of the Spanish language, 17-48.

Face, T. L. (2004). The intonation of absolute interrogatives in Castilian Spanish. Southwest Journal of Linguistics, 23(2).

Face, T.L. (2007). The role of intonational cues in the perception of declaratives and absolute interrogatives in Castilian Spanish. Estudios de fonética experimental, 186-225.

Gabriel, C., Feldhausen, I., Pešková, A., Colantoni, L., Lee, S. A., Arana, V. & Labastía, L. (2010). Argentinian spanish intonation. Transcription of intonation of the Spanish language, 285-317.

Hualde, J. I. (2003). El modelo métrico y autosegmental. Teorías de la entonación, 155, 184.

Hualde, J. I. & Prieto, P. (2015). Intonational variation in Spanish: European and American varieties. In Intonation in romance. Oxford University Press.

Ladd, D. R. (1981). On intonational universals. In Advances in psychology (Vol. 7, pp. 389-397). North-Holland.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of $F_0$ of voice. Phonetica, 41(1), 1-16.

Sosa, J. M. (1999). La entonacion del español: Su estructura fónica, variabilidad y dialectología. Madrid: Catedra.

Van Heuven, V. J. & Haan, J. (2000). When and how do we hear whether a Dutch speech utterance is a statement or a declarative question? Linguistics in the Netherlands, 17(1), 93-104.

Willis, E. (2010). Dominican Spanish intonation. Transcription of intonation of the Spanish language, 123-153

# Dialectal effects on the perception of Swedish vowel category boundaries

Renata Kochančikaitė, Mikael Roll
*Lund University*
renata.kochancikaite@ling.lu.se, mikael.roll@ling.lu.se

## Abstract
We investigated whether the dialectal area of native Swedish speakers has an effect on their perception of vowel category boundaries. Individual category boundaries were measured along six inter-category continua with the help of an identification task. A generalised additive mixed effects model with spatial smoothing was used to evaluate areal effect on category boundaries. The effect was statistically significant. Northern dialectal areas were associated with a clockwise shift in category boundary along the periphery of the F1/F2 space, while Southern areas—with a counter-clockwise shift. Part of the variance in boundary shifts was explained by formal musical training and frequent use of other languages than Swedish.

**Keywords:** category boundaries, vowel perception, dialectal variation, Swedish, GAMM

## 1. Background
Variation in speech perception is often explored in the context of second language acquisition or native contrasts across languages. Within a language, dialectal variation is usually studied in terms of speech production. Studies of dialectal variation in perception of speech sounds remain relatively few. We chose to investigate variation in vowel perception, in terms of category boundaries. Although vowel boundaries are typically fuzzier than consonant boundaries (Fry et al., 1962; Stevens et al., 1969), the acoustic space of vowels, too, is perceptually warped to accommodate native phoneme categories (Kuhl 1991; Dehaene-Lambertz 1997; Cheour et al 1998). Swedish language has 9 long vowels (Riad 2014) competing for the limited F1/F2 space, which offers interesting material for investigations of vowel perception variation. In terms of dialect, varieties spoken in mainland Sweden can be grouped into main types: Western and Central Swedish (closest to standard variety), Northern Swedish, and South Swedish. It is established that pronunciation of vowels differs across and within these dialectal groups (Leinonen, 2010) but the perceptual variation is, to our knowledge, not mapped yet.

## 2. Methodology
### 2.1. Perception experiment
Sixty healthy native Swedish speakers from 15 different regions (see section 6.), completed a 2AFC identification task. The stimuli were based on Central Swedish vowel prototypes and covered six inter-category continua between nine long vowel phonemes arranged in the round: /uː/-/oː/, /oː/-/ɑː/, /ɑː/-/ɛː/, /ɛː/-/eː/, /yː/-/ʉː/, /ʉː/-/øː/. Large number of inter-category continua was used to separate non-phonological effects of acoustic variation, e.g. increased salience of merged formants (Chistovich 1985) or peripheral location in the F1/F2 space (Polka & Bohn 2011), which are non-uniformly present in the Swedish vowel inventory (Kochančikaitė & Roll 2022). The morphed inter-category vowels were embedded in minimal CVC word pairs, the choice alternatives on the left or right of the screen were counterbalanced. The step size was kept constant at 0.15 Bark, computed in the F1/F2 acoustic space using Manhattan distance, $(\Delta F1+\Delta F2)/2$. Due to constant step size, the number of steps along each continuum varied depending on phoneme's prototypical position on F1/F2 vowel space. Before fitting sigmoid functions to participants' responses, the steps along the inter-category continua were scaled to range from 0 to 1, so that 0.5 would represent the middle between two Central Swedish prototypes, a likely position for a standard Swedish category boundary.

Logistic mixed effects regression was used to estimate individual categorisation sigmoids and measure the random slopes and intercepts for each participant. Random slopes and intercepts for each continuum were defined to separate the non-phonological effects from the categorisation function. Individual sigmoid thresholds at 50/50 likelihood were taken as the location of the individual category boundaries along the intercategory continuum (Fig. 1).



**Figure 3:** Distribution of estimated category boundary locations (sigmoid thresholds at 50% likelihood of either choice) along the common scaled inter-category continuum.

## 2.2. Dialect analysis

The participants completed a survey, indicating their dialectal area (city or town where they grew up), other languages they frequently use, and whether they have any formal musical education. 38 specific dialectal areas were represented in the dataset (see section 6.). Each location's coordinates in WGS84 standard were obtained. To measure the effect of dialectal area on vowel category boundaries, a generalized additive model (Wood 2011) was fitted, also featuring exposure to other languages and musical education as predictors.

## 3. Results

The effect of dialectal area was statistically significant ($p = 0.035$), suggesting that dialectal variation affects the perceptual warp of the acoustic space of Swedish vowels (Fig. 2). Exposure to other languages ($p < 0.001$) pushed the category boundary leftwards by 19% of the full range of boundaries estimated in the data (i.e. all sigmoid thresholds obtained when fitting a logistic regression model). Musical education ($p = 0.013$) pushed the boundary in the opposite direction by 13% of the data range.



**Figure 4:** Spatial representation of the estimated effect of dialectal area on vowel category boundaries. Positive effect (warm hues) means a rightward shift from the neutral 0.5 boundary. Negative effect (cold hues) means a leftward shift. An effect of 0.04 corresponds to ca. 15% of the entire boundary range observed in the data, from 0.4 to 0.66.

## 4. Discussion and conclusion

Significant variation in vowel category boundaries was attributed to dialectal area. Since prototype placement on experiment screen was counterbalanced, and non-phonological effects separated away, the boundary shifts cannot be attributed to visual, motor, acoustic or lexical bias. More likely, the results could indicate two opposite chain shifts in the vowel systems of the dialects, as Northern listeners preferred the prototype that is adjacent clockwise along the vowel space edge, while Southern listeners showed the opposite preference. It remains to confirm specifically, which of the adjacent phoneme pairs are affected, whether these shifts

can be a drag or a push, and to what extent these shifts match the production patterns across dialects.

## 5. References

Bruce, Gösta (2004). An intonational typology of Swedish. In Proceedings of Conference on speech prosody 2004 (175–178). Nara.

Cheour, Marie, Ceponiene, Rita, Lehtokoski, Anne, Luuk, Aavo, Allik, Jüri, Alho, Kimmo and Näätänen, Risto (1998). Development of language-specific phoneme representations in the infant brain. Nature neuroscience, 1(5), 351-353.

Chistovich, Ludmilla A. (1985). Central auditory processing of peripheral vowel spectra. Journal of the Acoustical Society of America, 77(3), 789–805.

Dehaene-Lambertz, Ghislaine (1997). Electrophysiological correlates of categorical phoneme perception in adults. NeuroReport, 8(4), 919-924.

Fry, Dennis B., Abramson, Arthur S., Eimas, Peter D. & Liberman, Alvin M. (1962). The Identification and Discrimination of Synthetic Vowels. Language and Speech, 5(4), 171–189.

Kochančikaitė, Renata & Roll, Mikael (2022). Phonetic and Phonological Variation in Vowel Discrimination Performance: Effect of Swedish Vowel Categories and Dialects. Preprint: Proceedings of Fonetik 2022—the XXXIIIrd Swedish Phonetics Conference (1-5). Stockholm: KTH Royal Institute of Technology.

Kuhl, Patricia K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. Perception & Psychophysics, 50(2), 93–107.

Leinonen, Therese (2011). Aggregate analysis of vowel pronunciation in Swedish dialects. Oslo Studies in Language, 3(2).

Riad, Tomas (2014). The phonology of Swedish. Oxford: Oxford University Press.

Stevens, Kenneth, Liberman, Alvin M., Studdert-Kennedy, Michael & Öhman, S. E. G. (1969). Crosslanguage Study of Vowel Perception. Language and Speech, 12(1), 1–23.

Wood, Simon N. (2011). Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. Journal of the Royal Statistical Society Series B: Statistical Methodology, Volume 73(1), 3–36.

## 6. Appendix

Country regions (regular type) and dialectal areas (italic type) represented in the data:

1. Dalarna (*Falun, Mora, Orsa*),
2. Gästrikland (*Gävle*),
3. Jämtland (*Östersund*),
4. Medelpad (*Sundsvall*),
5. Norrbotten (*Luleå*),
6. Östergötland (*Linköping, Motala*),
7. Skåne (*Ängelholm, Höör, Lund, Malmö*),
8. Småland (*Aneby, Jönköping, Kalmar, Tingsryd, Tranås, Växjö*),
9. Södermanland (*Nyköping*),
10. Stockholm (*Nacka, Stockholm*),
11. Uppland (*Knivsta, Kungsängen, Uppsala*),
12. Värmland (*Arvika, Kristinehamn, Munkfors*),
13. Västerbotten (*Umeå*),
14. Västergötland (*Bohus, Borås, Göteborg, Kungälv, Mölndal, Skövde, Trollhättan*),
15. Västmanland (*Kolsva, Kungsör, Västerås*).

# How non-binary visual gender cues prime German sibilant perception

Miriam Oschkinat[1], Melanie Weirich[2], Stefanie Jannedy[1]

*[1]Leibniz-Centre General Linguistics (ZAS), [2]Friedrich Schiller University Jena*

oschkinat@leibniz-zas.de

**Abstract**

This study tests speech perception in context by accompanying a synthetic gender-ambiguous voice with synthetic faces that vary in perceived gender. Previous research on American English has shown that a phoneme boundary can shift when co-presented with information on the speaker's gender. When a gender-ambiguous voice is seemingly produced by a male speaker, listeners hear more /s/- responses along a /s/ - /ʃ/ continuum, while when paired with a female face, more /ʃ/- responses are perceived. The current study aims to extend our knowledge of gender cues on speech perception by testing for priming effects in German, and with faces that are prototypical or non-prototypical in perceived gender. We supplemented the same gender-ambiguous voice with four different pictures of faces (a feminine looking woman, a masculine looking woman, a feminine looking man, and a masculine looking man). A total of 173 listeners identified stimuli on an acoustic continuum as either belonging to the German words Tasse or Tasche while exposed to one of the faces as priming cues. Our findings show different perception patterns in dependence of the priming cue. More /s/-responses were perceived for the feminine looking man and the feminine looking woman, and more /ʃ/ responses for the masculine looking woman and the masculine looking man, suggesting that the concept of femininity and masculinity triggered the responses and not the binary gender category: more femininity let listeners expect higher frequencies and more masculinity lower frequencies. The findings suggest that priming mechanisms affect speech perception in German real-words, however, in the opposite direction than found for American English listeners. Our data further supports an impact of gender as a non-binary social category on speech perception.

**Keywords:** categorical perception, gender, sibilants, morphing, German

## 1. Introduction and background

Speech production varies tremendously between speakers and can largely be attributed to social influences. The centre of gravity of the fricatives /s/ and /ʃ/ produced by men is typically lower than when produced by women, due to both physiological factors and socially constructed gender roles (May, 2005; Strand, 1999). In speech perception, information about the speaker's gender is incorporated into processing to account for such differences. For example, Strand and Johnson (1996) matched synthetic stimuli of non-prototypical male and female voices along a continuum between sod and shod each with videos of a male and a female speaker producing these stimuli. Their results showed that the female speaker shifted the perceived boundary between /ʃ/ and /s/ upward in frequency (more /ʃ/ responses), while the male speaker induced a perceptual downward shift (more /s/ responses). Similar effects were found when listeners were explicitly primed with pictures of a man or a woman while listening to ambiguous stimuli (Munson et al., 2017). However, further studies on priming have shown that visual cues can shift perceptual boundaries towards the expected category of the visual cue (Hay et al., 2006; Jannedy & Weirich, 2014), thus in the opposite direction to the findings by Strand and Johnson (1996) and Munson et al. (2017) for American English. The current study aims at shedding light on the influence of prototypical and non-prototypical gender cues on speech perception in German by priming listeners with visual gender cues. Acoustic stimuli on a gender-ambiguous Tasse (/ta.sə /, cup) – Tasche (/ta.ʃə /, bag) continuum were supplemented with faces that differed in perceived femininity and masculinity.

## 2. Methodology

The gender-ambiguous carrier word (Ta*e) was created by morphing acoustic instances of a female and a male Tasche stimulus. For the sibilant continuum, female Tasse- and Tasche-stimuli were morphed in 11 steps (Kawahara et al., 2008), and the 11 resulting sibilants implemented into the gender-ambiguous carrier word. The visual stimuli (Figure 1b) were created as similar faces that exclusively differ in their perceived femininity/masculinity by using the browser-based MetaHuman creator (https://metahuman.unrealengine.com/). We created a feminine woman, a masculine woman, a feminine man, and a masculine man, as confirmed by picture ratings (Oschkinat et al., in press). Four stimulus sets were created with the acoustic continuum accompanied by one of the four faces. A total of 173 participants between 16 and 91 years (mean 45y, 47 male, 120 female, 5 non-binary / genderqueer, one without indication) participated via the Plapper App (Jannedy, 2022) and listened to at least two of the four priming sets (one of the women and one of the men). Auditory stimuli presentation within a set was randomized, and participants identified the stimuli as either Tasse or Tasche (two-alternative forced choice identification task) by means of pictures.

## 3. Results

Comparisons across the different sets reveal that for the masculine man (Figure 1a, dark blue) and the masculine woman (Figure 1a, light magenta) listeners hear longer Tasche (cf. stimulus 6), while when primed with the feminine man (light blue), listeners hear longer Tasse (cf. stimulus 6). The feminine woman (dark magenta) is at chance level at stimulus 6, however closer to the perception curve of the feminine man. She also has less /s/ ratings for stimuli 7 and 8 than the others. Thus, the visual priming cue shifts perception along the /s/-/ʃ/ continuum.
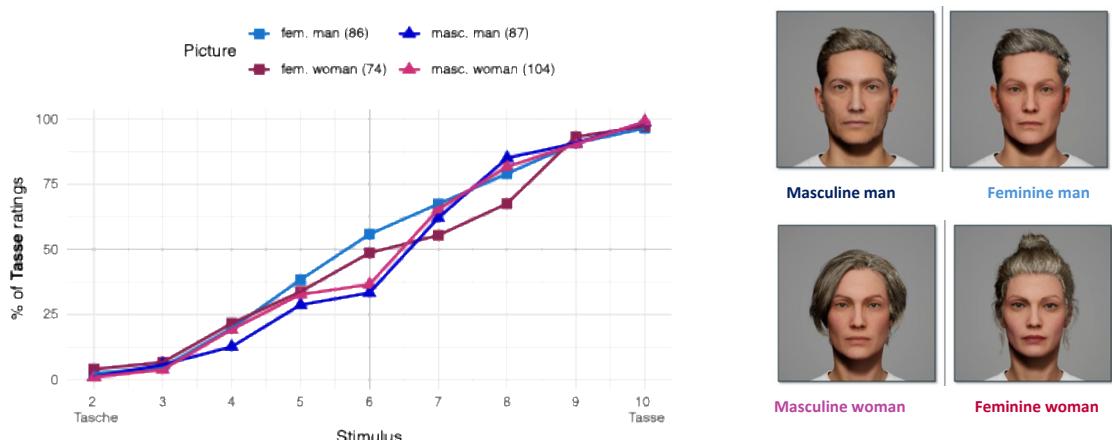


Figure 1. a) Percent /s/-responses along the continuum (steps 2 – 10 shown). b) Faces created in MetaHuman.

## 4. Discussion and conclusion

The first insights into the data indicate that the context of a person' gender and gender typicality affects the category boundary between German sibilants. The visual priming with the masculine man and the masculine woman cluster together and trigger more /ʃ/ responses - the sibilant category with energy in lower frequencies - as would be expected from a male speaker. On the other hand, visual priming with the feminine woman and the feminine man trigger more /s/-responses, the sibilant category with energy in higher frequencies, as would be expected from a female speaker. This pattern suggests that it is not explicitly the attributed gender of the speaker (male or female) that alters perception, but rather the concepts of femininity and masculinity that crucially affect the perception of the acoustic stimuli. Further interpretation of the priming effects as well as effects of listener gender on perception must await further statistical analyses of the perception curves, which will be presented at the conference.

## 5. References

Hay, J., Nolan, A. & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. Linguistic Review, 23(3).

Jannedy, S. (2022). Plappern für die Wissenschaft – Eine App gegen die Datenlücke. Bericht über das Forschungsjahr 2020/2021: ZAS; Impressionen, 115-126.

Jannedy, S. & Weirich, M. (2014). Sound change in an urban setting: Category instability of the palatal fricative in Berlin. Laboratory Phonology, 5(1), 91-122.

Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T. & Banno, H. (2008). Tandem-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 3933–3936.

May, J. (2005). Vocal tract normalization for /s/ and /š/. The Journal of the Acoustical Society of America, 59(S1), S25. https://doi.org/10.1121/1.2002554

Munson, B., Ryherd, K. & Kemper, S. (2017). Implicit and explicit gender priming in English lingual sibilant fricative perception. Linguistics, 55(5), 1073-1107.

Oschkinat, M., Weirich, M. & Jannedy, S. (in press). MetaHumans as holistic Personae for implementation into Virtual Reality. Proceedings of the P&P 20, Halle Saale, Germany.

Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. Journal of language and social psychology, 18(1), 86-100.

Strand, E. A. & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. KONVENS, 14-26.

# The Perception-Production Link of Taiwan Mandarin /s/ and /ʂ/ through Identification and Imitation Tasks

Yung-Ning Jhuang
*Graduate Institute of Linguistics, National Taiwan University*
r12142004@ntu.edu.tw

## Abstract

This study explores how phonological categorization of merging sibilants can be reflected in imitation of Taiwan Mandarin (TM) /s/ and /ʂ/. Previous research found that the productions of TM sibilants are merging, while the perception remains categorical. Categorization and imitation tasks using s-ʂ followed by /a ɨ u/ items on a continuum were performed by nine participants. Results showed categorical imitation for all speakers. Participants' imitation of an ambiguous signal was associated with their perception of the signal, showing that the imitation paradigm can reveal phonological organization, highlighting a perception-production link.

**Keywords:** sibilant merger, phonetic imitation, Taiwan Mandarin

## 1. Introduction and background

Previous studies on TM sibilant production showed that /s/ and /ʂ/ demonstrate large overlaps both acoustically (Chang, 2012; Chuang & Fon, 2010) and articulatorily (Chiu et al., 2020). The perception, on the other hand, remains categorical even for speakers with merged sibilants in production. As Chang (2012) explained, compared with Beijing Mandarin, TM had a higher cutoff frequency for /ʂ/ perception, indicating a broader perceptual boundary for /ʂ/. However, Lee-Kim and Tung (2023) highlighted that although merged speakers were outperformed by their unmerged counterparts, they could perceptually distinguish /s/ and /ʂ/ above the chance level.

The link between perception and production is also observed through the design of identification and imitation tasks since they provide a direct mapping between low-level acoustics and phonological awareness. Hwang and Lu (2025) employed the design of imitation to investigate how nasality is imitated differently depending on whether it is a primary cue. Categorical imitation was found only in the medial position, where nasality was used as a primary cue, even when ambiguous targets were presented. Following this, categorical imitation would be expected when an ambiguous stimulus is not identified as either category; gradient imitation would be anticipated when the stimulus continuum is identified less categorical. The current study seeks to test this hypothesis on the imitation of the Taiwan Mandarin s–ʂ continuum, aiming to examine the link between acoustic merger and phonological representations. It is expected that categorical s-ʂ phonological contrast will lead to categorical imitation, while merged perception will be manifested as gradient imitation.

## 2. Methodology

Nine (6 female) TM native speakers participated in two tasks: identification and imitation. In the identification task, the stimuli were an 11-step s–ʂ continuum across 3 vowel contexts: /a ɨ u/. The recording was done by a female native speaker who distinguished the two phonemes and synthesized by Tandem-Straight (Kawahara et al., 2008). Participants were instructed to identify each stimulus as /s/ or /ʂ/. For analyses, a mixed-effects logistic regression model was fit for the identification task using R (R Core Team, 2020). In the imitation task, subjects imitated stimuli step 1 (prototypical /ʂ/), step 6 (ambiguous) and step 11 (prototypical /s/) from the previous task. Acoustic analysis of the imitation productions was conducted by measuring the center of gravity (CoG). A mixed-effects linear regression for the imitation using R. Identification results served as a perceptual baseline for each subject and were considered as a

factor in the model, and the emmeans package in R was used to examine how different perception baselines contributed to the CoG values (Lenth, 2025).

## 3. Results

As Figure 1a shows, the identification results found substantial individual variations in the proportion of /ʂ/ response to the ambiguous step 6, showing a continuum of consistent /ʂ/ classification to /s/ classification. Random effects analysis showed a variance of 6.26 (SD = 2.50) in the subject-level intercepts, reflecting considerable variation across subjects. Participants were thus classified into two groups: one that tended to identify the ambiguous signals as retroflex /ʂ/ (i.e., the ʂ-favored group) and the other as dental /s/ (i.e., the s-favored group).

As shown in Figure 1b, different imitation results were found for these two groups. For the ʂ-favored group, the CoG of step 6 imitation was not significantly different from their step 1 /ʂ/ imitation, but was significantly different from their step 11 /s/ imitation. For the s-favored group, on the other hand, the imitation of the ambiguous signal was significantly different from their step 1 /ʂ/ imitation, but not their step 11 /s/ imitation. In other words, both groups exhibited categorial imitation, aligning their productions of the ambiguous signal with their perceptual classifications.
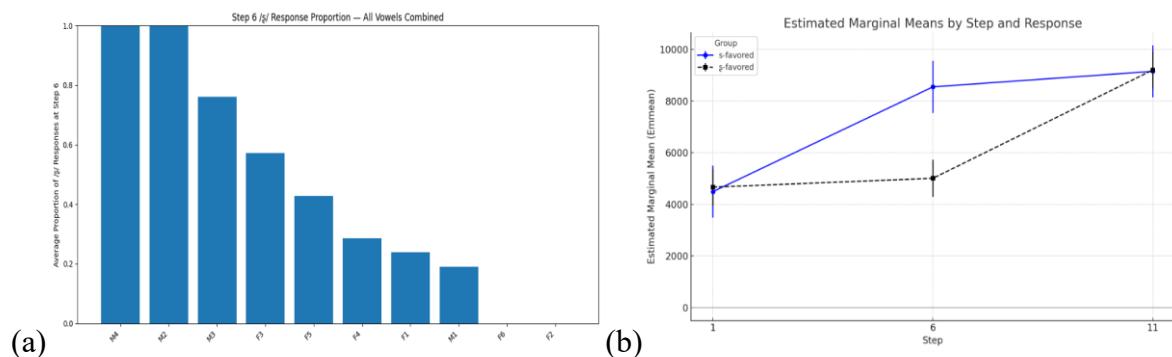


Fig. 1. (a) By-subject step 6 /ʂ/ response proportion from the identification task. (b) Estimated Center of Gravity (CoG) across steps 1, 6, and 11 as a function of listeners' response type.

## 4. Discussion and conclusion

The results are consistent with previous imitation tasks, showing that phonetic imitation reflects the abstract phonological organization. Both groups in the experiment showed categorical behavior in the imitation task, and their categorization in the imitation task followed their pattern in perception task. The results from the current experiment thus conformed with previous findings in Lee-Kim and Tung (2023), providing further evidence that phonological contrasts of /s/ and /ʂ/ in Taiwan Mandarin remained categorical.

## 5. References

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

Chiu, C., Wei, P.-C., Noguchi, M. & Yamane, N. (2020). Sibilant Fricative Merging in Taiwan Mandarin: An Investigation of Tongue Postures using Ultrasound Imaging. Language and Speech, 63(4), 877-897. https://doi.org/10.1177/0023830919896386

Chuang, Y.-Y. & Fon, J. (2010). The effect of prosodic prominence on the realizations of voiceless dental and retroflex sibilants in Taiwan Mandarin spontaneous speech. Proceedings of Speech Prosody 2010.

Hwang, J. & Lu, Y.-A. (2025). Perception-production link mediated by position in the imitation of Korean nasal stops. JASA Express Letters, 5(3). https://doi.org/10.1121/10.0036057

Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T. & Banno, H. (2008). Tandem-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. 2008 IEEE international conference on acoustics, speech and signal processing,

Lee-Kim, S.-I. & Tung, H.-Y. (2023). Sibilant Perception By Merged Speakers: The Case Of Taiwan Mandarin. Proceedings of the 20th International Congress of Phonetic Sciences – ICPhS 2023 (pp. 122-126),

Lenth, R. V. (2025). emmeans: Estimated Marginal Means, aka Least-Squares Means. https://github.com/rvlenth/emmeans.

R Core Team. (2020). R: A language and environment for statistical computing, R Foundation for Statistical. Computing.

# Tonal Training and Cue Salience in Naïve Perception of Northern Vietnamese Tones

Hồng Quang Châu, Ioana Chitoran, Hiyon Yoo
*Université Paris Cité*
hong-quang.chau@etu.u-paris.fr

**Abstract**

Non-tonal listeners typically rely on pitch height when perceiving lexical tones, but it is unclear how they approach languages like Northern Vietnamese (NV), where phonation is also phonemic. This study investigated how naïve listeners perceive NV tones and whether brief, cue-specific, language-mismatched training can improve performance. Fifty native French listeners performed a pre-test/post-test AXB discrimination task on four NV tone pairs contrasting mainly in pitch height, contour, and/or creaky voice. Between tests, participants received brief feedback-based training in NV, Cantonese, Thai, Mandarin, or a French segmental control. At pre-test, the rising creaky vs. modal tone pair was discriminated most easily, suggesting listeners rely on the most salient combination of acoustic cues rather than defaulting to pitch alone. Post-training, all groups improved modestly in accuracy and significantly in speed, but gains were uniform across all conditions. There was no advantage for any tonal training language over the control and no evidence of cue-specific transfer. These results suggest that a single, low-variability training session provides a general attentional boost but is insufficient to restructure perceptual cue-weighting for non-native tones.

**Keywords:** lexical tone, Northern Vietnamese, cross-language perception, perceptual training, cue weighting

## 1. Introduction and background

Listeners whose native language is non-tonal have been shown to rely on absolute $F_0$ height while giving less weight to dynamic contour (Gandour 1983; Huang & Johnson 2011). This arguably reflects either language-general psychoacoustic salience (Hallé et al. 2004), or the assimilation of new tones to the listener's native intonation system (So & Best 2014). Such understanding, however, is based primarily on languages where pitch is the sole or primary phonetic dimension of tone. The present study extends this inquiry to Northern Vietnamese (NV), where the phonation type is contrastive (Brunelle & Kirby 2016). While non-tonal listeners can be sensitive to phonation (Cao et al. 2012), it remains underexplored how they navigate a system where both $F_0$ and voice quality systematically contrast.

Furthermore, although studies show that extended, high-variability training can shift listeners' weighting of acoustic cues (Francis et al. 2008; Chandrasekaran et al. 2010), it remains unclear whether brief, cue-specific training can be effective, especially when the training language differs from the target language. This study therefore asks: (RQ1) What is the baseline perceptual hierarchy for NV tone contrasts? (RQ2) Does brief training in a tonal language improve overall discrimination of NV tones? (RQ3) Does such training enhance discrimination selectively for cues emphasized in the training language?

## 2. Methodology

Fifty native Metropolitan French speakers ($20.6 \pm 2.4$ years) with no tonal language experience were assigned to five training groups: Vietnamese (NV), Cantonese (CAN), Thai (TH), Mandarin (MAN), and a French segmental control (CTRL). The experiment used a three-phase AXB discrimination task (pre-test, training, post-test) with a 250 ms ISI. Stimuli were CV syllables produced by native speakers. The pre- and post-tests used four NV tone pairs selected to highlight contrasts in pitch height, pitch contour, and two creaky vs. modal phonation pairs (one rising, one falling). During training, groups heard analogous tone contrasts from their

assigned language with feedback, while the CTRL discriminated French segments. Dependent measures of accuracy and log-transformed RTs were analyzed with mixed-design ANOVAs, using non-parametric equivalents where model assumptions were violated.

## 3. Results

Baseline perception (RQ1): At pre-test, discrimination accuracy varied significantly by tone pair (Friedman $\chi^2(3) = 62.8$, $p < .001$). As shown in Figure 1, the rising creaky vs. modal pair was discriminated most accurately (M=0.90), while the falling creaky vs. modal pair was the most difficult (M=0.71). The height and contour pairs were intermediate. RTs revealed the same pattern of difficulty ($p < .001$). There were no baseline differences between groups.
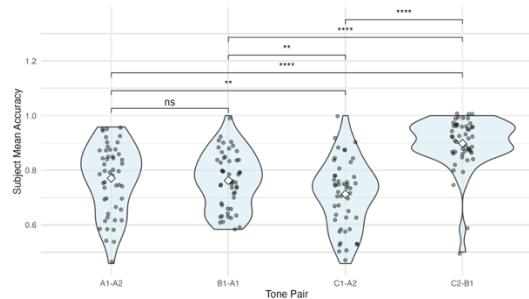


Figure 1: Pre-test accuracy by Vietnamese tone pair
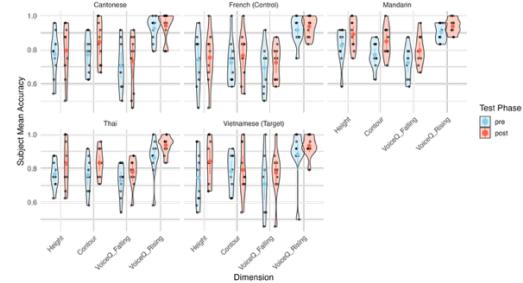


Figure 2: Pre- and post-test accuracy by training group and tone pair

Training effects (RQ2 & RQ3): After training, all groups improved significantly in both accuracy ($F(1, 45) = 31.42$, $p < .001$) and speed ($F(1, 45) = 75.85$, $p < .001$). These gains, however, were uniform across all training conditions, with no significant main effect of Training Language and no Time x Language interaction ($p > .50$). There was also no evidence of the predicted cue-specific transfer, as no interactions involving the training language were significant. A significant Time x Dimension interaction was found for reaction times ($F(3, 135) = 4.42$, $p = .005$), indicating that the amount of speed-up differed by cue, with the easiest pairs showing the largest RT improvement. The difficulty hierarchy, however, remained intact for all groups post-training (see Figure 2).

## 4. Discussion and conclusion

This study examined how naïve French listeners perceive Vietnamese tones and whether brief, cue-specific, language-mismatched training could reshape their perception. The exceptional ease of discriminating the rising creaky vs. rising modal pair suggests that performance is guided by overall psychoacoustic salience, where a combination of robust cues (a large $F_0$ movement and a distinct phonation change) is more effective than a single-dimension contrast. This aligns with findings that non-tonal listeners can exploit salient phonation cues (Cao et al. 2012; Kuang & Liberman 2018), nuancing models based on phonological assimilation (Best 1995). The training results demonstrate the limits of short-term, low-variability exposure. The uniform improvement across all groups, including the non-tonal control, is best explained as a general attentional boost or the effect of "active listening" (Heald & Nusbaum 2014). The training was insufficient to induce cue-reweighting, arguably because within-category variance was not manipulated to signal to the listener that their default perceptual strategy, or their preferred tone cue, is unreliable (Holt & Lotto 2006). By keeping the tokens for each tone category fixed, our design increased exposure frequency but did not provide the necessary variance to force a shift in cue weighting, a finding that contrasts with the successes of longer-term, high-variability phonetic training (Wayland & Guion 2004).

In sum, the results suggest that naïve perception of complex tones is guided by the salience of bundled acoustic cues. While a single training session can sharpen attention and improve response speed, it is insufficient to restructure listeners' underlying perceptual hierarchy. Inducing genuine cue re-weighting for non-native tones likely requires longer, high-variability phonetic training that systematically manipulates the reliability of competing acoustic dimensions.

**5. References**
Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), Speech perception and linguistic experience: Issues in cross-language research (171–204). Baltimore: York Press.
Brunelle, M. & Kirby, J. (2016). Tone and phonation in Southeast Asian languages. Language and Linguistics Compass, 10(4), 191–207.
Cao, R., Wayland, R. & Kaan, E. (2012). The role of creaky voice in Mandarin tone 2 and tone 3 perception. Proceedings of Interspeech 2012, 426–429.
Chandrasekaran, B., Sampath, P. D. & Wong, P. C. M. (2010). Individual variability in cue-weighting and lexical tone learning. The Journal of the Acoustical Society of America, 128(1), 456–465.
Francis, A. L., Ciocca, V., Ma, L. & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. Journal of Phonetics, 36(2), 268–294.
Gandour, J. (1983). Tone perception in Far Eastern languages. Journal of Phonetics, 11(2), 149–175.
Hallé, P. A., Chang, Y.-C. & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. Journal of Phonetics, 32(3), 395–421.
Heald, S. & Nusbaum, H. C. (2014). Speech perception as an active cognitive process. Frontiers in Systems Neuroscience, 8, 35.
Holt, L. L. & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. The Journal of the Acoustical Society of America, 119(5), 3059–3071.
Huang, T. & Johnson, K. (2011). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. Phonetica, 67(4), 243–267.
Kuang, J. & Liberman, M. (2018). Integrating voice quality cues in the pitch perception of speech and non-speech utterances. Frontiers in Psychology, 9, 2147.
So, C. K. & Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. Studies in Second Language Acquisition, 36(2), 195–221.
Wayland, R. P. & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. Language Learning, 54(4), 681–712.

# Perception of flapped /t/ in British English: a generational shift?

Eric Ferrari, Coline Caillol, and Emmanuel Ferragne
*Université Paris Cité*
E.Ferrari@sms.ed.ac.uk, coline.caillol@etu.u-paris.fr, emmanuel.ferragne@u-paris.fr

## Abstract

This study investigates the perception of the flapped allophone of /t/ in British English, a feature traditionally associated with American English. We explore whether this perception has evolved among British English speakers. Using an identification perception experiment, participants from two age groups (18-25 and 55+) had to identify the orthographic representation of the audio stimuli of nonce words containing a flap /t/, produced by speakers of SBE and GA. Preliminary analysis revealed that while the effect of generation alone was not significant, the interaction between generation and accent was significant. Younger participants were more likely to correctly identify a flap in British English compared to older participants, suggesting a generational shift in the perception of this phonetic feature. These findings indicate a potential evolution in the acceptance of flapped /t/ as a variant in British English, which may be due to the prevalent use of American English in popular media.

**Keywords:** perception, varieties of English, flapped /t/, media

## 1. Introduction and background

The phoneme /t/ has various possible allophonic realizations in intervocalic position in English, the most common being an alveolar stop [t], a glottal stop [ʔ] or an alveolar flap [ɾ]. This study focuses on the evolution of the perception of flap3 /t/ in word-internal intervocalic positions by British English speakers, on the basis that it has been described as "one of the most striking characteristics of American pronunciation to the ears of a non-American" (Wells, 1982: 248). More than a markedly American pronunciation trait, phoneticians in the 1970s and 1980s have predicted it may in time spread to varieties of British English (O'Connor, 1973) and beyond, potentially making this trait "the first distinctively American phonetic innovation likely to spread in time to all accents of English" (Wells, 1982: 250). Though flapped allophones of intervocalic /t/ do rarely occur in some specific varieties of British English (Cockney, for example) it is extremely marginal (Tollfree, 1999) and most speakers typically use stops, whether glottal or alveolar. Recently, a study by Alderton (2022) looked into the use of flapping in young speakers of Standard Southern British English from high socio-economic backgrounds, and found that it indexes both authority and informality. He further posits that that link may come from the association of flapping with American English and its widespread use in popular media. Though his results show that flapping is still the least favored realization of /t/ over both alveolar and glottal stops, they also reveal that it is increasingly considered as an optional contemporary variant, used for specific social-indexical purposes. This study proposes to examine the perceptual aspect of flap /t/, namely if it is still considered as non-British feature, as it was described as markedly American in the 1980s, or if it has evolved into an allophone of /t/ that is more generally accepted by British English speakers and listeners today. In other words, is there a generational difference in the perception of flaps as allophones that are not as unexpected as they used to be in British English?

---

3 There is debate over the terminology used to refer to this allophone (i.e. *flap* or *tap*). We follow De Jong (2011)'s logic in calling in a flap, as it is the term used in a majority of works to describe this phenomenon.

## 2. Methodology

The study takes the form of an identification perception experiment. Stimuli are nonce words, the structure of which is inspired by Braver (2014) to reduce reliance on lexical frequency and knowledge in determining the participants perception of the phoneme. All stimuli are three syllables: the first is an unstressed CV syllable with [ə], the second is the nuclear syllable with either the vowel [i], [e], or [a] and a final consonant, and the third syllable is '-ing', to mimic the structure of a gerund, and to ensure that the target environment is VCV, such as buhPEETing, puhPATing, and tuhPETing. The stimuli were recorded by two native English speakers: one Southern British English speaker, and one General American speaker. The speakers read 72 words each, 36 target words with either a stop or a or flap /t/, and 36 fillers. PsychoPy was used to construct the experiment online (Pierce et al. 2022). In part one, participants are presented with one audio stimulus and then presented with two possible orthographic representations of the stimulus. For example, they might hear /pəˈkiːɾɪŋ/, then have to select between the orthographic representations, such as puhKEETing or puhKEEDing. 51 participants took part in the online study. Participants all hail from England, excluding the rest of the UK and Ireland. They were filtered into two age groups: 18-25 (n = 24) and 55+ (n = 27) to account for generational differences in the perception of flapped /t/ in British English. In part two, participants are presented with a questionnaire inquiring about their attitude towards American and British accents and the potential influence of the media.

## 3. Analysis, results and discussion

As a preliminary look into the data, only the answers to the identification of flap audio stimuli were considered (18 flapped tokens in SBE and 18 flapped tokens in GA per participant). If participants correctly identified the flap as an orthographic <t> as opposed to a <d>, the answer was considered correct. The proportion of correct answers by accent and generation can be seen in Figure 1 below. A generalized linear mixed effects model was fitted to the data including fixed effects for generation (18-25 vs. 55+) and accent of the stimuli (SBE vs. GA), as well as their interaction, and a random intercept for participants. While the effect of generation on correctly identifying a flap as a /t/ was not statistically significant (p = 0.754), there was a significant effect of accent, where hearing a flap in an SBE accent strongly decreased the likelihood of getting a correct answer ($\beta$ = -1.86, SE = 0.18, p < 0.001). Interestingly, the interaction of generation and accent was also significant ($\beta$ = 0.75, SE = 0.24, p < 0.01), suggesting that younger participants have an easier time recognizing a flap in British English compared to older participants. These results may therefore suggest that younger British people might be more accustomed to hearing this sound in British English. The data will be further exploited by looking at reaction times in a second half of the analysis, which will be carried out in time for the conference. We believe reaction times will help provide a more finely detailed picture of the potential evolution of the perception of flaps for British speakers, which the preliminary analysis appears to point to already. Qualitative results from the survey also show an interesting glimpse into the attitudes of participants towards the media as a possible source of influence for this evolution, as was suggested by Alderton (2022). Both groups in the survey largely agree that media, particularly social media and TV, has influenced pronunciation in England, such as a [ɾ] in place of a [t].

## 4. References

Alderton, Roy (2022). T-tapping in Standard Southern British English: An 'elite' sociolinguistic variant? Journal of Sociolinguistics, 26(2), 287-298.

Braver, Aaron (2014). Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in American English flapping. Lingua, 152, 24-44.

De Jong, Kenneth (2011). Flapping in American English: Flapping in American English. In M. van Oostendorp, C. J. Ewen, E. Hume & K. Rice (Eds.), The Blackwell Companion to Phonology (1-19). Oxford: John Wiley & Sons.

O'Connor, Joseph Desmond (1973). Phonetics. Middlesex: Penguin Books.

Peirce, Jonathan, Hirst, Rebecca & MacAskill, Michael (2022). Building Experiments in PsychoPy. 2nd Edn. London: Sage.

Tollfree, Lisa (1999). South East London English: Discrete versus continuous modelling of consonantal reduction. In P. Foulkes & G. J. Docherty (Eds.), Urban Voices: Accent studies across the British Isles (163-184). London: Routledge.

Wells, John C. (1982). Accents of English. Cambridge: Cambridge University Press.
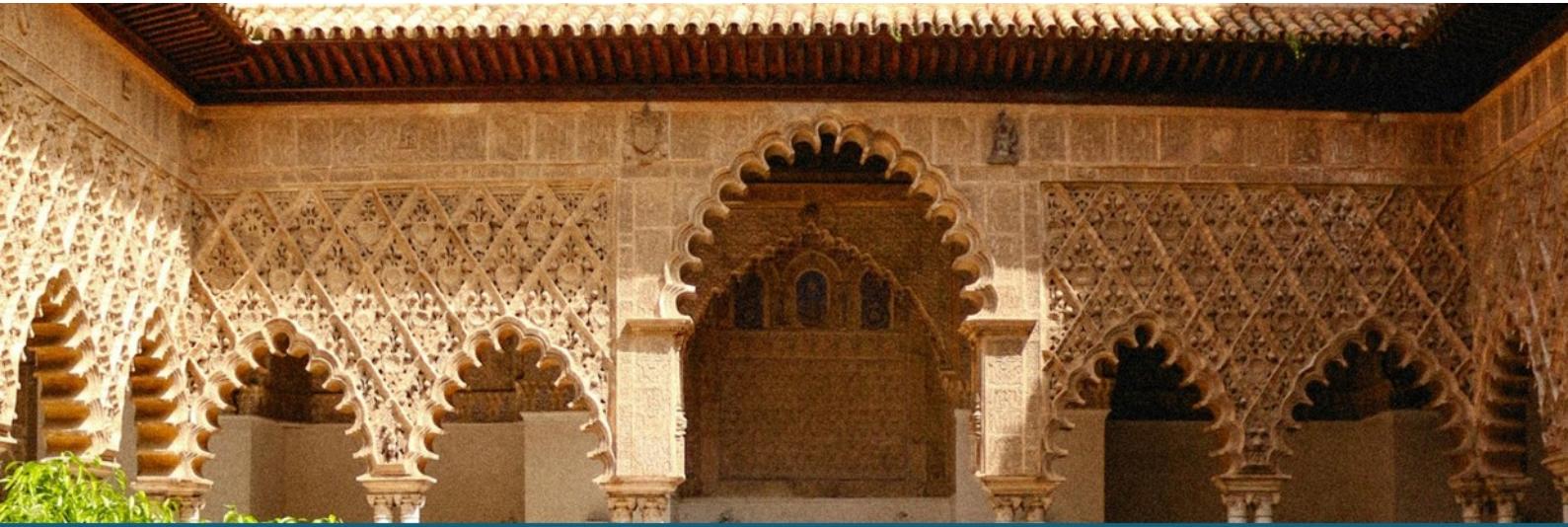
# L2 ACQUISITION

# Vowels and Consonants in Contact: A Comparative Study of the Spanish Spoken by Natives Speakers and the Spanish of Polish Residents in Granada

Edyta Waluch de la Torre, Marcin Sosinski
*University of Granada*
ewaluch@ugr.es, sosinski@ugr.es

## Abstract

This study compares vowels and selected consonants in the Spanish spoken in the North of Spain, Spanish spoken in Granada, and Spanish of Polish immigrants. Using PRAAT analysis and diverse corpora, it examines phonetic variation, assimilation, and transfer. Findings highlight mother tongue influence, prescriptive pressure, and orthography as key factors shaping immigrant pronunciation.

**Keywords:** phonetic assimilation, acoustic analysis, Spanish spoken in Granada, Polish immigrants

## 1. Introduction and background

This presentation analyzes the phonetic characteristics of vowels and some consonants in the context of Spanish spoken in the North of Spain, Spanish spoken in Granada, and the Spanish used by a group of Polish immigrants residing in the city. This study has been carried out within the framework of the research project "Study of the social determinants of current Spanish in Granada: new identities, new challenges, new solutions" (PID2023-148371NB- C41).

## 2. Methodology

The research is based on data collected from the vernacular population of Granada, as well as from a sample of adult Polish residents who have lived in Spain for more than ten years and were late learners of Spanish. The study offers a phonetic and acoustic description and comparison—using PRAAT software—of the vowels and consonants identified as the most difficult for these speakers to acquire. The corpora used include recordings of the local population of Granada (PRESEEA-Granada corpus [Moya, 2007, 2009, 2010], ECOS-Granada corpus) and of Polish residents in the province (POL- GRAN corpus). The results are also compared with Spanish spoken in the north of Spain.

## 3. Results

The presentation describes the different realizations of vowels in Spanish spoken in the North of Spain and in Spanish spoken in Granada, and then examines the most complex aspects of vowel and consonant production in Spanish by Polish immigrants, along with the degree of phonetic dissimilation in specific contexts. Levels of adjustment and assimilation to native vowel and consonant pronunciation are also detailed (Sosinski, 2022; Waluch de la Torre, 2022).

The findings show that the phenomena under study are mainly influenced by mother tongue transfer, prescriptive pressure (Sosinski & de la Torre, 2021), and the impact of orthography. The conclusions reveal varying degrees of phonetic assimilation in concrete cases, as well as the parameters affecting whether adjustment is absent, partial, or fully aligned with native pronunciation.

In the consonantal domain, the qualitative analysis reveals three main factors that account for the phonetic mismatches: transfer from the mother tongue, prescriptive pressures, and the influence of orthography. Transfer is evident both in the persistence of Polish articulations in supposedly "analogous" phonemes (/b d g/ as stops) and in the substitution of phonemes that do not exist in Polish (/θ/, /r/). Prescriptive pressures lead speakers to avoid local Andalusian

features (such as aspiration or the lenition of /s/ and /x/), even when these variants might actually be easier for Polish speakers. Finally, orthography plays a significant role, particularly in the labiodental pronunciation of <v> and in the effort to distinguish between <y> and <ll>.

As for vowels, the results of the acoustic analysis using F1 and F2 indicate a strong accommodation of Polish speakers to the vowel system in Granada, with relatively small discrepancies. Polish speakers tend to produce local vowels with greater openness than in Polish and with somewhat higher fronting/advancement, especially among men. Nevertheless, the deviations from the local phonological pattern are minor and confirm an effective phonetic assimilation to the model of Spanish spoken in Granada.

## 4. References

Moya, J. A. (Coord.) (2007). El español hablado en Granada: Corpus oral para su estudio sociolingüístico (PRESEEA-Granada; I, nivel de estudios alto). Editorial Universidad de Granada.

Moya, J. A. (Coord.) (2009). El español hablado en Granada II: Corpus oral para su estudio sociolingüístico (PRESEEA-Granada; II, nivel de estudios medio). Editorial Universidad de Granada.

Moya, J. A. (Coord.) (2010). El español hablado en Granada III: Corpus oral para su estudio sociolingüístico (PRESEEA-Granada; III, nivel de estudios bajo). Editorial Universidad de Granada.

Sosinski, M. (2022). Realización de consonantes españolas por inmigrantes polacos en Granada. Estudio contrastivo. Círculo de Lingüística Aplicada a la Comunicación, 92, 63-86.

Sosinski, M., & de la Torre, E. (2021). Creencias y actitudes de los jóvenes universitarios de la Universidad de Varsovia hacia la variedad andaluza del español. Philologia Hispalensis, 35(1), 193-214.

Waluch de la Torre, E. (2022). Realización de vocales españolas por hablantes nativos e inmigrantes polacos. Círculo de Lingüística Aplicada a la Comunicación, 92, 109–125.

# The realization of /b d g/ by Czech learners of Spanish

Radek Skarnitzl, Petr Čermák
*Charles University*
radek.skarnitzl@ff.cuni.cz, petr.cermak@ff.cuni.cz

**Abstract**
The study focuses on the production of the Spanish /b d g/ – and particularly their approximant variants [β̞ ð̞ ɣ̞] – by more and less experienced Czech learners of Spanish and relates their realizations to the Spanish pronunciation norm, as well as to native speakers of Spanish. The data are analyzed using logistic mixed-effects regression modelling and exploratory analyses. The results point to an interesting variability at the level of phonetic detail not only in the LX speakers, but also in the native Spanish group.

**Keywords:** stops, approximants, Spanish as L2, Czech

## 1. Introduction and background

The Spanish phonemes /b d g/ may be realized as stops [b d g] or approximants [β̞ ð̞ ɣ̞], depending on the context (RAE 2011). The primary stop variants are pronounced after a pause (bueno [ˈbu̯eno]) and when preceded by a nasal sound both within a word and across the word boundary (dónde [ˈdo̞n̪d̪e], vienen de [ˈbi̯en̪en̪ d̪e]); with /d/, a stop is also produced when preceded by the lateral liquid /l/ (el día [el̪ ˈd̪ia]). The approximant variants are pronounced in all other contexts, for example intervocalically (la bebida [la βe̞ˈβi̞ð̞a]) or in consonantal contexts (hablar [aβ̞ˈlaɾ], los dos [los ˈð̞os]).

The acquisition of the stop vs. approximant distinction in Spanish has been examined especially in L1 speakers of English (see González Bueno 2019 for a review). The approximant realizations are reported to be particularly difficult for LX learners to master, as they tend to be rare the learners' L1 phonological systems (Face & Menke 2009; Wayland et al. 2024). Elliott (1997, cited in Face & Menke 2009) regards these sounds among those which most affect the accentedness of learners of Spanish.

Given the existing knowledge about the challenges faced by LX learners of Spanish when producing the Spanish phonemes /b d g/ and their allophonic variants, this study focuses on Czech speakers of Spanish. Apart from the basic stop vs. approximant distinction, we will be interested in the effect of specific phonetic patterns of Czech on the learners' performance: the asymmetry in the "alveolar" stops, with [t] being denti-alveolar and [d] alveolar (and, in some speakers, even postalveolar; see Skarnitzl 2013); and the effect of spelling due to the existence of /v/ in Czech, which may be pronounced as a labiodental fricative [v] or approximant [ʋ] (Skarnitzl & Volín 2005).

## 2. Methodology
### 2.1. Material and speakers

The Araneum corpus of written Spanish (Benko, 2014) was used to prepare texts which comprised all three phonemes (or, more precisely, all four graphemes – <b>, <v>, <d> and <g>) in all possible phonetic contexts. The final text contained 41 sentences and 338 words, and 122 target sounds. In total, the analyses are based on 3,280 items of /b d g/.

A total of 18 Czech university-level learners of Spanish (aged 19–25, 15 females and 3 males, level B2 or C1 as per CEFR) were recorded in a sound-treated recording studio. As a proxy for the learners' language proficiency, we determined an LX experience score which tapped into their length of Spanish study and use of Spanish for various purposes. They were then divided into the less experienced group (n = 8) and the more experienced group (n = 10).

For the sake of comparison, we used recordings of 9 native speakers of Spanish (aged 29–56, 4 females and 5 males, all university-educated, from Spain).

## 2.2. Analyses

The stop vs. approximant realization of the target sounds was analyzed using a combination of detailed listening and visual inspection of the acoustic signal in Praat (Boersma & Weenink, 2024). In addition, we coded alveolar stop realizations of the Spanish dental /d̪/, as well as labiodental fricative or approximant realizations of the spelled <v> (see section 1).

First, logistic regression was used to fit the data to predict the likelihood of a correct phonetic realization (i.e., one in agreement with the Spanish pronunciation norm) as a function of speaker group (less experienced, more experienced, native), grapheme (<b v d g>) and phonetic context. Random effect structure included by-speaker and by-item intercepts. Subsequent exploratory analyses focused on the finer distinctions in phonetic realizations.

## 3. Results and discussion

The logistic regression model confirmed a significant effect of speaker group, with native speakers and more proficient learners being significantly more likely to produce target-like realizations than less experienced learners. The effect of grapheme was also significant, with realizations of <d> being significantly more frequently noncanonical. As for phonetic context, realizations in post-consonantal (e.g., algunos) and word-final (e.g., club) contexts were substantially less likely to be correct, while post-pausal contexts increased the likelihood of a correct realization. Figure 1 shows a sample of the context-specific probabilities.
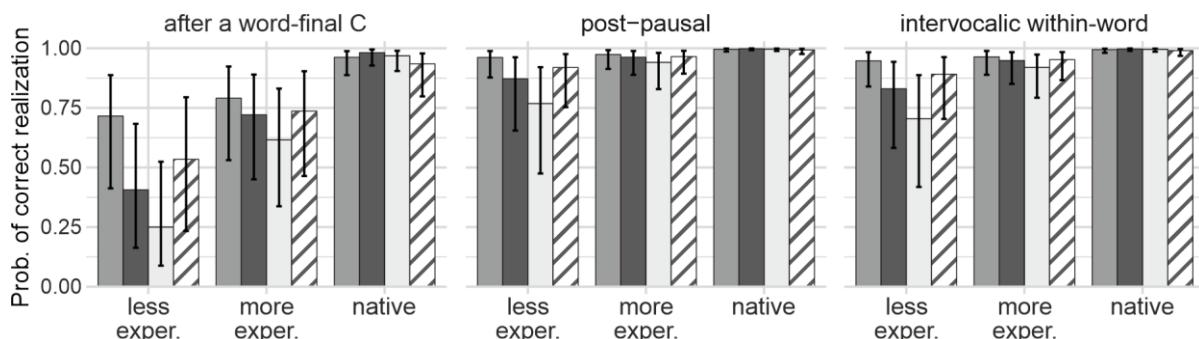


Figure 1. Probability of correct realization of <b>, <v>, <d> and <g> as a function of speaker group (less experienced, more experienced, native) in three phonetic contexts.

The exploratory analyses confirmed a strong effect of spelling of <v>, in both Czech speaker groups, but with a clear effect of language experience: [v] realizations accounted for nearly 63% of all <v> items in the less and for 26% in the more experienced group. Similarly, the prevalence of the alveolar stop [d] was considerably higher in the less experienced (nearly 66% of all /d/ items) than in the more experienced group (40% of all items).

In addition, our analyses revealed interesting and by no means marginal realizations of the native speakers of Spanish. Naturally, we did not expect all native productions to necessarily be in line with the pronunciation norm, but a closer inspection revealed that it was especially those contexts where the target sound followed a nasal (or, with /d/, a lateral) within a word (and, to a lower extent, also word-initially). These nasal–stop sequences, where we would predict pronunciation with no audible release (e.g., cambiar [kam̚ˈbi̯aɾ] or tengo [ˈteŋ̚go]), were often lenited and, through progressive assimilation of manner, realized as geminated nasals [kam̚ˈmi̯aɾ ˈteŋŋo] (see, e.g., Martínez Gil, 2020: 35–37 for more on such processes in Spanish). Returning to the Czech LX users of Spanish, it was interesting to observe

that this lenition pattern could be observed in both our groups, but considerably more often in the more experienced speakers, indicating their awareness of this process in Spanish.

## 4. General discussion

Although noncanonical realization of the variants of Spanish /b d g/ is related to accentedness rather than intelligibility, the results of the study do have pedagogical implications.

## 5. References

Benko, Vladimír (2014). Aranea: Yet another family of (comparable) web corpora. In P. Sojka et al. (eds.), Proceedings of the 17th International Conference on Text, Speech and Dialogue (257-264). Cham: Springer International Publishing.

Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer Version 6.4.20. www.praat.org. Accessed 20/09/2024.

Face, Timothy L. & Menke, Mandy R. (2009). Acquisition of the Spanish voiced spirants by second language learners. In J. Collentine et al. (eds.), Selected Proceedings of the 11th Hispanic Linguistics Symposium (39-52). Somerville, MA: Cascadilla Proceedings Project.

González Bueno, Manuela (2019). Suggestions for teaching Spanish voiced stops /b, d, g/ and their lenited allophones [β, ð, ɣ]. In R. Rao (ed.), Key issues in the teaching of Spanish pronunciation: From description to pedagogy (60-83). London: Routledge.

Martínez Gil, Fernando (2020). Spirantization and the phonology of Spanish voiced obstruents. In S. Colina & F. Martínez Gil (eds.), The Routledge handbook of Spanish phonology (34-83). London: Routledge.

RAE. (2011). Nueva gramática de la lengua española. Fonética y fonología. Madrid: Espasa Libros.

Skarnitzl, Radek (2013). Asymmetry in the Czech alveolar stops: An EPG study. AUC Philologica 1/2014, Phonetica Pragensia XIII, 101-112.

Skarnitzl, Radek & Volín, Jan (2005). Czech voiced labiodental continuant discrimination from basic acoustic data. In I. Trancoso (ed.), Proceedings of Interspeech 2005 (2921-2924). Lisbon: ISCA.

Wayland, Ratree, Meyer, Rachel, Vellozzi, Sophia & Tang, Kevin (2024). Lenition in L2 Spanish: The impact of study abroad on phonological acquisition. Brain Sciences, 14(9), 946.

# Clipping of English /iː/ triggered by /p, t, k/ codas in the speech of Polish learners of English

Michał Baran
*University of Szczecin*
michal.baran@usz.edu.pl

**Abstract**
This paper aims to contribute to the acoustic description of Polish English by studying pre-fortis clipping of /iː/ triggered by /p, t, k/ in the speech of primary school, high school and university learners. The study of recordings of 18 participants (primary school n=6, high school n=6 and university n=6) suggests that pre-fortis clipping is not exhibited consistently by Polish learners. As for the level differences, university students generated the closest realisations to the native ones, while primary and high school learners' pronunciations were not consistent regarding vowel length.

**Keywords:** pre-fortis clipping, Polish English

## 1. Introduction and background

The English language exhibits pre-fortis clipping which is defined as shortening the length of a vowel happening when the vowel is followed by a fortis consonant (see, for instance, Roach 2013). Pre-fortis clipping may be key in telling the difference between minimal pairs as English word-final obstruents undergo devoicing at the end of syllables which contributes to reducing any voicing-based differences. Pre-fortis clipping, seen in a broader context of phonemic vowel length in English, constitutes a serious challenge for Polish learners of English as in Polish vowel length is not phonemic.

Pre-fortis clipping of English vowels produced by Polish learners of English has been a subject of a number of phonetic studies (see, for instance, Waniek-Klimczak 1998; Zając 2013), but few of them compare the speech of participants representing various age groups. In order to fill this gap, the present study looks at clipping of English /iː/ followed by /p, t, k, b, d, g/ codas in the speech of Polish primary school, high school and university learners. The main research goal is to check whether the length patterns are similar to the native ones and whether they are influenced by the age group or coda type.

## 2. Methodology

This study is based on recordings of 18 participants representing three different age groups (primary school n=6, high school n=6 and university n=6) reading sentences from a reading list. All of the participants, regardless of the age group, are native speakers of Polish who were born and raised in Poland and have never lived in an English-speaking country.

The participants were asked to read a list of sentences while sitting 20 centimetres away from Røde NT-USB+ microphone connected to a laptop. The sentences were created by putting target lexemes in a carrier phrase "Say X again.". The sentences were repeated so that each of the plosive codas had 20 repetitions by a given speaker.

The recordings were analysed with Praat software (Boersma & Weenink, 2024). Vowel length was measured on the basis of visual identification of the beginning and end of modal voicing and the appearance and disappearance of vowel formants. The collected data was analysed using MS Excel.

## 3. Results

| | /p/ | /b/ | Difference | /t/ | d/ | Difference | /k/ | /g/ | Difference |
|---|---|---|---|---|---|---|---|---|---|
| **Primary School** | 87,7 (±30) | 90,6 (±26) | 2,9 | 91,7 (±25) | 110,2 (±39) | 18,5 | 83,9 (±27) | 89,1 (±31) | 5,2 |
| **High School** | 79,3 (±18) | 77,6 (±18) | -1,7 | 80 (±18) | 89,5 (±23) | 9,5 | 81,4 (±22) | 77,2 (±22) | -4,1 |
| **University** | 90,3 (±29) | 100 (±25) | 9,8 | 98,5 (±21) | 127,8 (±29) | 29,3 | 84,5 (±21) | 97,4 (±25) | 12,9 |

Table 1. Mean vowel length for syllables with /p, t, k, b, d, g/ codas in the three age groups

As shown in Table 1. above, vowels produced by the participants of the study do not exhibit clipping patterns identical to the ones of native English. Primary school students produced slightly longer vowels (i) before /b/ than before /p/ and (ii) before /g/ than before /k/, which, taking into consideration a relatively high standard deviation suggests that the length is not consistently influenced by the coda. The difference between the effect of fortis and lenis coda is higher in the case of the alveolar consonants, but still there is a possibility of overlap of the vowel length values before /t/ and /d/ (see Figure 1. below). Interestingly, the results of high school students are even further away from the native tendencies than the ones of primary school students. In the cases of /p/ vs /b/ and /k/ vs /g/, the difference is a negative number which means that, on average, the vowels followed by fortis /p/ and /k/ are longer than those followed by lenis /b/ and /g/. This is exactly the opposite of what happens in native English. The difference in the case of /t/ vs /d/ is positive, but it does not seem to be big enough to claim that this contrast is consistently marked with vowel length. Finally, the results of the university students seem to be the closest to the ones expected in native English, however, it should be noted that in the cases of /p/ vs /b/ and /k/ vs /g/, the differences remain relatively small. Figure 1. below present box-and-whisker plots of the vowel length to show that in all the contrasts under review here, there is a high chance of value overlap.
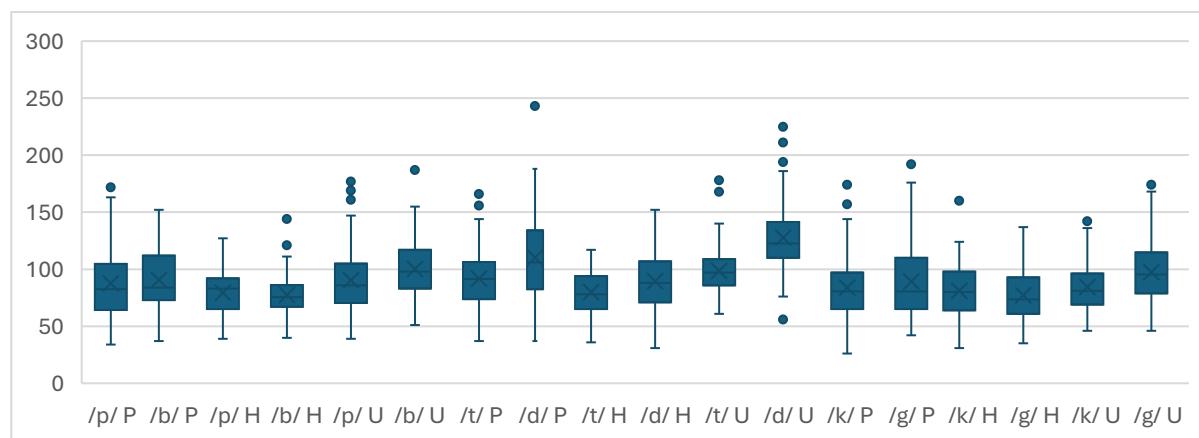


Figure 5. Box-and-whisker plots of vowel length before /p, t, k, b, d, g/ codas

## 4. Conclusions

All in all, the study reveals that the university students produced vowels with the lengths closest to the native ones, i.e. they seem to exhibit some degree of pre-fortis clipping. Nevertheless, even in this case, the differences are not always big enough to claim that the contrast is consistent. It may, therefore, be concluded that for all three age groups analysed here, vowel clipping in English remains problematic. As for the differences between the three coda contrasts studied here, the /t/ vs /d/ clearly generated results closest to the native realisations with the average vowel length being higher for /d/ coda than for its fortis counterpart.

## 5. References

Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer. Version 6.4.34. https://www.fon.hum.uva.nl/praat/ Accessed on 20.05.2025.

Roach, Peter (2013). English Phonetics and Phonology. Cambridge: Cambridge University Press.

Waniek-Klimczak, Ewa (1998). Sources of variability in L2 phonology: vowel duration in 'Polish English'. In: C. Gruber et al. (eds.) CLS 34/2: The Panels (387-401). Chicago: Chicago Lingustic Society.

Zając, Magdalena (2013). Phonetic Imitation of Vowel Duration in L2 Speech. Research in Language. 11 (1), 19-29.

# Prosodic Variation in the Context of a Public Speaking University Course in English

Raúl Jiménez Vilches
*Universidad Autónoma de Madrid*
raul.jimenez@uam.es

**Abstract**
An increasing number of Spanish universities use English in their bilingual and international programmes. Many Spanish speakers do not seem to have problems at expressing themselves when it comes to writing in English. However, they seem to have trouble when reporting their ideas orally, especially when delivering oral presentations. Some universities offer public speaking courses to first-year students, but they still seem far from obtaining a high level of oral mastery. The present study proposes a new programme for teaching public speaking in which two crucial prosodic parameters such as key and unfilled pause are developed. Discourse intonation approach is chosen as the framework for the proposal due to its direct application to language learning. Public speeches before and after the intervention are acoustically analysed. Results seem to reveal that the proposed programme causes improvement when it comes to key or unfilled pause, but not simultaneously (both parameters can be used to mark new topics within an oral presentation more effectively).

**Keywords:** prosody, public speaking, discourse intonation, Spanish learners of English

## 1. Introduction and background

Spanish speakers seem to encounter difficulties when they have to deliver oral presentations in English (Bang-Rouhet 2020). An effective public speaker focuses not only on the message but also the delivery. Prosody is mostly responsible for the musical aspects of speech and its discourse and pragmatic values. This musicality is triggered by prosodic elements such as pitch, pauses, or voice quality (Wichmann 2000). Native speakers acquire prosody at very early stages in lide, and those sounds are internalised subconsciously and automatically (Ramírez Verdugo 2003). Contrarily, L2 speakers do not learn prosody in the same way. Specific forms and functions of the second language are not internalised as they are in the first language since L2 speakers are not aware of the multiple prosodic patterns that they can choose (Wells 2006). The present study aims to design, adapt, and implement a new course to develop the speaker's communicative value through prosody in the context of teaching public speaking at university level. It also aims to check whether the subjects are able to improve their prosodic usage because of the prosodic training course.

## 2. Methodology

This study proposes an innovative methodology implying the idea of not only creating something new but also improving the existing ideas and practices (Mitchell 2003). Discourse intonation (Brazil 1997) was chosen as the framework for the innovation due to its direct application to language learning. The population for the present study includes Spanish learners of English who are enrolled at a state university that offers a course of public speaking in English. The pre- task includes the same topic for all subjects, and they are also asked to give an oral presentation following a set structure – introduction, main body and conclusion - and a set of phrases which are commonly used in a public speech delivered by non-native speakers of English. The presentations before and after the intervention are recorded in the classroom. All the oral presentations are acoustically analysed using Praat (Boersma and Weenik 2010). The central part of the course teaches the two prosodic elements (key and unfilled pause) included in this study.

## 3. Results and discussion

The following table is based on DeVito's (2012) pausal framework so we can measure not only the length of pauses but also the use of key in order to mark new ideas or topics. The table quantifies the pitch height of the key (measured acoustically in Hertz) and long pauses lasting between one to two seconds that have been proved to be prosodic markers of a new topic. The speaker's use of key at new topics is compared before (BPT) and after (APT) the training course. The bottom row quantifies the mean key and pauses of all the new topics landmarks both before and after the instructed discourse intonation approach.

| DISTRIBUTION OF KEY (Hz) AND UNFILLED PAUSE (sec) | FS1 | | FS2 | | FS3 | | FS4 | |
|---|---|---|---|---|---|---|---|---|
| | BPT | APT | BPT | APT | BPT | APT | BPT | APT |
| Before beginning speech > greeting | 239.3 Hz 2.29 sec | 322.7 Hz 2.57 sec | 323.9 Hz 0.99 sec | 325.5 Hz 1.67 sec | 301.3 Hz 2.04 sec | 404.6 Hz 2.13 sec | 257.6 Hz 1.72 sec | 269.3 Hz 2.25 sec |
| Introduction > body point 1 | 386 Hz 1.16 sec | 354.9 Hz 3.81 sec | 296.2 Hz 1.14 sec | 430.6 Hz 1.59 sec | 326.5 Hz 1.26 sec | 395.2 Hz 0.83 sec | 296.1 Hz 1.01 sec | 223.8 Hz 1.59 sec |
| Body point 1 > body point 2 | 333.4 Hz 1.25 sec | 330.8 Hz 1.02 sec | 288 Hz 1.06 sec | 298.5 Hz 0.72 sec | 236.9 Hz 0.34 sec | 367.2 Hz 0.51 sec | 248.9 Hz 0.91 sec | 272 Hz 1.66 sec |
| Body point 2 > conclusion | 369 Hz 1.59 sec | 412.6 Hz 1.36 sec | 254.1 Hz 0.50 sec | 262.4 Hz 0.68 sec | 317.5 Hz 0.82 sec | 337.6 Hz 0.96 sec | 238.5 Hz 1.44 sec | 245 Hz 2.09 sec |
| Conclusion > ask for questions | 243.4 Hz 1.60 sec | 338.6 Hz 1.47 sec | 210.6 Hz 0.92 sec | 282.7 Hz 0.73 sec | 221.7 Hz 0.65 sec | 303.5 Hz 0.91 sec | 192.4 Hz 0.96 sec | 211.1 Hz 0.73 sec |
| Mean key Mean unfilled pause | 314.2 Hz 1.44 sec | 351.8 Hz (+37.7 Hz) 1.78 sec (+0.34 sec) | 274.6 Hz 1.18 sec | 319.9 Hz (+45.3 Hz) 1.08 sec (-0.17 sec) | 280.8 Hz 1.01 sec | 361.2 Hz (+80.8 Hz) 1.14 sec (+0.13 sec) | 246.7 Hz 1.11 sec | 244.2 Hz (-02.5 Hz) 1.42 sec (+0.31 sec) |

FS1 = Female Subject 1; FS2 = Female Subject 2; FS3 = Female Subject 3; FS4 = Female Subject 4; BPT = Before Prosodic Training; APT = After Prosodic Training.
Table 1. Speaker's selection of key and unfilled pauses to mark new topics in an oral presentation.

Table 1 reveals that most subjects choose a higher key in the first tone unit to cue the beginning of a new topic in their after-training L2 public speaking. Regarding length of unfilled pauses, most subjects are capable of making longer mean pauses to mark the beginning of a new point

for their final presentations. Overall, figures indicate Spanish informants can improve the key to signal new topics. Contrarily, there were fluctuations and irregularities when it comes to using filled pauses since there was no improvement in their final (APT) public speeches. To sum up, Spanish learners of English have not been able to improve the two prosodic forms consistently in their after-training oral presentations.

## 4. Conclusion

Results appear to indicate that the new teaching proposal does not seem to be enough for improving the non-native prosodic competence consistently. The Spanish presenters' intentions to mark new topics prosodically did not actually match the production of both key and unfilled pauses, even after treatment. All in all, these results seem to confirm previous studies such as Ramirez-Verdugo (2003) when she concludes that prosodic acquisition does not take place automatically.

## 5. References

Bang-Rouhet, P. (2020). Los españoles hablan alto en el bar, no en las empresas y la voz es el motor de todo. La Razón newspaper. Accessed 12/04/2025.

Boersma, P. and Weenik, D. (2010). Praat: Doing Phonetics by Computer [Computer Program]. Version 5.1.07, http://www.praat.org/ . Accessed 24/03/2025.

Brazil, D. (1997). The Communicative Value of Intonation. Cambridge: C.U.P.

DeVito, J. A. (2012). The Essential Elements of Public Speaking. Boston: Pearson Education.

Mitchell, J. M. (2003). Emerging Futures: Innovation in Teaching and Learning in VET. Australian National Training Authority, Australia.

Ramírez-Verdugo, M.D. (2003). A Contrastive Analysis of Non-native Interlanguage English Intonation Systems and their Implication in the Organisation of Information from a Functional Perspective. A Study Based on a Computerized Corpus of Spanish Learners of English (PhD Thesis). Universidad Autónoma de Madrid, Spain.

Wells, J. C. (2006). English Intonation. An Introduction. Cambridge: C.U.P.

Wichmann, A. (2000). Intonation in Text and Discourse. Oxford: Longman.

# Implementation of PRAAT in teaching of "Phonetics and Phonology of Spanish" of the Degree in Hispanic Philology: Learning variation from the first year of university

María Heredia Mantis
*Universidad de Granada*
herediamantis@ugr.es

## Abstract

In this paper we present a proposal for pedagogical improvement in the teaching of Spanish acoustic phonetics. The proposal consists of a two-hour practical seminar in which students analyze acoustic samples from their own speech using the PRAAT software. We present how it has been developed in the present academic year for the first time with great success. We also use this pedagogical methodology to introduce the teaching of dialectal and sociolectal phonetic variation even though the subject does not include it in its syllabus.

**Keywords**: teaching phonetics, PRAAT as a teaching tool, teaching innovation, didactics of phonetics, teaching variation

## 1. Introduction

This paper presents a didactic proposal for the course "Phonetics and Phonology of Spanish" (Degree in Hispanic Philology/Studies). Our teaching space is poorly adapted. Material resources are scarce and there is no phonetics laboratory to study acoustic phonetics. There is a lack of a manipulable and articulated anatomical reproduction of the human respiratory system and of the upper and lower articulatory organsto appreciate how sound production works. We are limited to showing digital recreations or recordings of the articulatory organs in movement Although we have printed support material (Hidalgo & Quilis 2012 as the handbook of the course, but also Martínez Celdrán 1991, Martínez Celdrán & Fernández Planas 2007, Fernández Planas 2022, Llisterri 2023-2025), it is not accessible to students: it does not include sound material that complements the spectrograms and the resolution quality of the spectrograms is so low that some acoustic features of the phonos are imperceptible or illegible. To fill these gaps, we have introduced a mandatory seminar in the course. The planning of this seminar took into account the initial study of the phonic variation between dialectal varieties of Spanish. The material worked on in this seminar didn't reflect the peninsular standard, but different speeches and accents.

## 2. Methodology

The pedagogical intervention is oriented to students of "Phonetics and Phonology of Spanish" in the first year of Hispanic Philology Grade. The subject "Acoustic Phonetics" can be evaluated through individual activities and the final exam of the course, it is an axial content. This is related to the topics of the syllabus: "Phonetics of Spanish", "The vowels of Spanish" in its subsection "Acoustic description of the vowels", "The consonants of Spanish" in its subsection "Acoustic description of the consonants", and "The suprasegmental units".

### 2.1. Objectives

a) To handle the PRAAT program at user level for acoustic analysis.
b) To interpret the acoustic features of the sounds and suprasegmental units of Spanish speech.
c) To understand the acoustic variation between dialectal and sociolectal varieties of Spanish.

### 2.2. Material resources

Due to the space limitations of the faculty, we must find an available time slot outside the usual class time and to split the group into two practical subgroups. The session took two hours for

each subgroup. This took place at the end of the semester, when the theoretically contents had already been studied. The classroom was equipped with one computer per student. It was specifically configured with PRAAT installed to work with sound through headphones.

## 2.3 Activities
The proposed activity was as follows:
Record and upload in the best possible quality in .mp3 or .wav format your pronunciation of the following words:
1. El coche; 2. Los coches; 3. Chaquetilla; 4. Piqueta; 5. Engordaba; 6. Llaneza; 7. Lotería; 8. Los osos; 9. Las osas; 10. El viento norte era muy frío
Students had a virtual assignment in Moodle to upload their sound files.

## 3. Results
Through the different segments proposed, the students were able to study the acoustic properties of practically all the sounds of Spanish, specially:
a) Stridency in [s], [θ], and its dialectal allophonic variants [s̪] or [s̄], [s] and [s̺], and in other less strident, [t͡ʃ] and its dialectal variant [ʃ], and [x] and its southern variant [h].
b) Occlusion in voiceless, with more marked VOT, and voiced, with somewhat less marked VOT, and absence of occlusion in contextual intervocalic approximant variants.
c) The distinction between treble (front) and bass (back) sounds, and between dense (low) and diffuse (high), and the acoustic difference in vowel sounds, clearly marked by the frequencies of the first and second formant, and consonant sounds, with less stable formants.
d) The presence of formants of sonority and nasality, as well as the analysis of articulatory stress as a function of the sonority and duration of the sound.
We took advantage of the dialectal diversity between the students to analyze dialectal variation. Here, we select the production of the segment "los osos" in three students.

In comparative, we see the difference between the first production (northern peninsular dialect), the second (southeastern peninsular dialect), and the third (Argentine dialect). In the first image and in the third, three [s] productions are perceived, the first more apical and the third more seseada (fronted and coronal), as opposed to the second in which only one clearly explosive
[s] is perceived, since the implosives are enormously leninized, instead, the preceding vowel [o] is much more lowered (as described in Herrero de Haro & Hajek 2020:146-147).

## 4. Discussion and conclusion
The academic results improved with respect to the previous academic year. In the first exam session, the reading of spectrograms in the exam was optional: those who answered it improved their score and answered correctly. In the second exam session, the spectrogram reading was mandatory, but the improvement in the overall scores with respect to the previous year was also achieved. The dropout rate for the course fell from 16% to 12.5%. The pass rate in the first call rose from 53% to 71%, and the overall pass rate rose from 62.5% to 78%. Likewise, the average score among those who passed the course also increased, as the percentage of students who obtained a "B" or "A" (more than 7 points out of 10) rose from 41% to 51.5%. The success of the pedagogical format of a reduced seminar on the use of the analytical tool PRAAT is confirmed by the quantitative data. The objectives are achieved by most of the students and that it represents a positive innovation in the development of the course. In future years, we propose to improve the planning, with more time dedicated to the seminar, and the implementation of a self-evaluation system through a questionnaire on the Moodle platform. As a further purpose, we foresee the elaboration of multimodal pedagogical material in digital format.

## 5. References

Fernández Planas, Ana Maria (2022). El dominio fónico del español: Estado de la cuestión. Estudios de Fonética Experimental, Special issue 1, 9-76.

Herrero de Haro, Alfredo & Hajek, John (2020). Eastern Andalusian Spanish. Journal of the International Phonetic Association, 52(1), 135-156.

Hidalgo Navarro, Antonio & Mercedes Quilis Merín (2012). La voz del lenguaje: fonética y fonología del español. Valencia: Tirant humanidades.

Llisterri, Joaquim (2023-2025). Fonética del español. <https://joaquimllisterri.cat/home.html>. Accessed on 13/06/2025.

Martínez Celdrán, Eugenio (1991). Fonética experimental: teoría y práctica. Madrid: Síntesis.

Martínez Celdrán, Eugenio & Ana María Fernández Planas (2007). Manual de fonética española: Articulaciones y sonidos del español. Barcelona: Ariel.

# The Role of Teacher Input Variability in L2 English Vowel Acquisition: A longitudinal study

Alejandra Pesantez, Karolina Broś
*Universität Zürich, University of Warsaw*
alejandracarolina.pesantezpesantez@uzh.ch

**Abstract**

L2 vowel contrast acquisition is challenging for learners when their L1 lacks equivalent phonemic categories. While several studies have compared L2 learner production with native English speakers, relatively little is known about how intra-group variability among teachers in the production of vowel contrasts over time affects how learners of English acquire it. This paper analysed productions by Ecuadorian learners of English and their teachers measured at three time points, using the Euclidean Distance (ED) in the F1/F2 vowel space for the vowel pair /i-ɪ/ as the variable of interest. The results show intra-group differences in teacher productions which, combined with changes of teachers across semesters, introduces substantial variability in the input that Ecuadorian learners are exposed to. This in turn makes it more difficult for the students to produce the vowel contrast /i-ɪ/, highlighting the need to study input variability in the context of L2 phonological category learning.

**Keywords:** L2 acquisition, vowel contrasts, Spanish learners of English, input variability

## 1. Introduction and background

L2 vowel contrast acquisition is challenging for learners when their L1 lacks equivalent phonemic categories (Pesantez & Dellwo, 2022; Escudero, 2005). Spanish-speaking learners of English, for example, often struggle to differentiate lax-tense vowel pairs such as /i-ɪ/ and /u-ʊ/ as these phonemic contrasts do not exist in their native vowel system (Flege & Wayland, 2019). Recent research has highlighted that inconsistent or non-target-like input can hinder the acquisition of L2 vowel contrasts (Escudero et al., 2014). According to the Second Language Linguistic Perception (L2LP) model (Escudero & Chládková, 2010), the phonetic quality of the input plays a critical role in shaping learners' perception and category formation. While several studies compare L2 learner productions with native English speakers, relatively little is known about how intra-group variability in the production of vowel contrasts among teachers affects learner outputs over time. To address this issue, we compared Ecuadorian student and teacher productions of /i-ɪ/ at three stages of university-level English vowel acquisition.

## 2. Methodology

The study analysed data from 21 monolingual Ecuadorian (EC) Spanish-speaking university students (mean age=21) enrolled in an English teacher-training program in Ecuador. Participants were in their third semester at the time of the first recording (T1), fourth semester at the second recording (T2), and fifth semester at the third recording (T3). They had no prior experience abroad and received a total of 1,104 hours of English teaching. All students followed the same curriculum. Instruction during the third and fifth semesters included English phonetics and phonology courses, respectively, focusing on pronunciation training, including the contrast /i-ɪ/. The instructor group consisted of eight L1 Ecuadorian Spanish lecturers of English (mean age=37). Four of them (A, B, C, D) were responsible for teaching English during T1, three (A, B, E) during T2, and four (C, F, G, H) during T3. Both EC learners and lecturers completed a controlled picture-naming task in which they produced monosyllabic CVC/CVCC words containing the eight target vowels (/i, ɪ/) following the methodology used by Flege and Wayland (2019). All participants were individually recorded in a sound-treated room at the university's radio station, using a Zoom H2n recorder (44.1 kHz sampling rate, 16-bit).

## 2.1. Acoustic Analysis

Acoustic analyses were conducted using Praat (Boersma & Weenink, 2022). The mean of the F1 and F2 of /i/ and /ɪ/ was the primary acoustic metric used, taken from the central portion (middle 50%) of each vowel, which helps control for coarticulation. Lobanov's z-scores were used to normalize vowels (Lobanov, 1971). We calculated the Euclidean Distance (ED) in the F1/F2 vowel space for each /i-ɪ/ pair across the three time points.

## 3. Results

Fig. 1 shows that not all teachers produce the same quality or clarity of vowel contrast although those teachers who happened to be lecturers at more than one timepoint were consistent in their productions. The intra-group differences, however, introduce variability in the input that EC learners are exposed to, with a dichotomy between teachers whose scores are similar to the students (low level of vowel contrast preservation) and those whose scores are higher at each learning stage. We tested these differences using linear mixed effects models, which showed that the differences between the teachers are significant at each timepoint ($F(7,43)=31.463$, $p<.001$) and overall lead to the lack of significant differences in mean Euclidean distance scores over time (confidence intervals are substantial, Fig. 2). A general model on all the data shows a significant difference between students and teachers at each timepoint ($F(1, 27.6)=13.36$, $p<.01$) and no significant interaction ($F(2, 278.4)=.95$, $p=.38$) nor timepoint effect ($F(2, 278.4)=2.93$, $p=.055$). Faced with such variability, students fail to advance in their pronunciations, although there is a trend in the desired direction. While there was an improvement in ED scores from T1 to T2 ($t=-3.871$, $p<.001$, based on emmeans), there was no further improvement from T2 to T3 ($t=-1.587$, $p=.25$).

Figure 1. Mean Euclidean Distance (z-score normalized) for the English vowel pair /i-ɪ/ produced by Ecuadorian learners at T1, T2, and T3, compared with their lecturers (letters A-H).
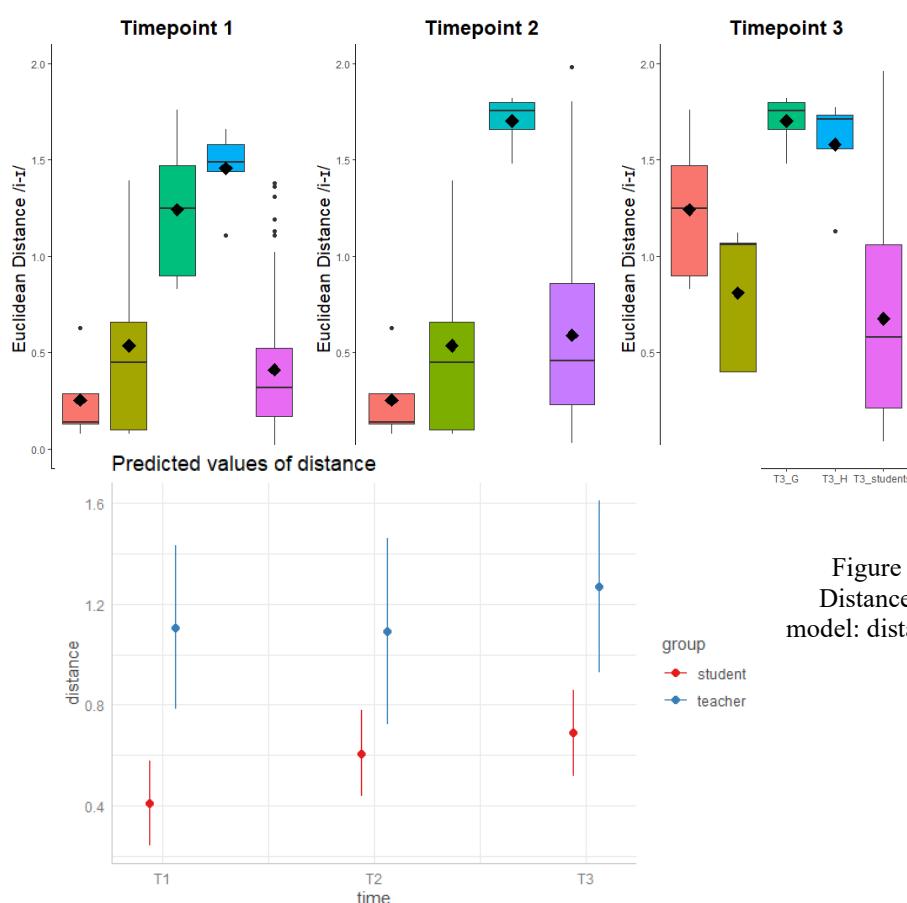


## 4. Discussion and conclusion

Overall, the results show that students do not reach the target although they do improve their pronunciations as a group. This may be because the uneven input productions they are exposed to hinder their



Figure 2. Values of Euclidean Distance predicted by the global model: distance~timepoint*group + (1 |speaker)).

learning process, and they may have to rely exclusively on the study materials (authentic productions used in class). More research is needed to disentangle the roles of model speakers/materials and non-native teacher productions in L2 vowel contrast acquisition. Since most L2 teachers are not native English speakers and their speech is heard more often than the speech of model speakers in authentic material, this issue plays a crucial role in understanding the factors that influence L2 pronunciation learning.

## 5. References

Boersma, P. & Weenink, D. (2022). Praat: Doing phonetics by computer (6.1.51) [Computer software]. http://www.praat.org/

Escudero, P. (2005). Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization. PhD thesis, LOT Dissertation Series 113, Utrecht University

Escudero, P. & Chládková, K. (2010). Spanish listeners' perception of American and southern British English vowels. The Journal of the Acoustical Society of America, 128(5), EL254-EL260. https://doi.org/10.1121/1.3488794

Escudero, P., Simon, E. & Mulak, K. E. (2014). Learning words in a new language: Orthography doesn't always help. Bilingualism: Language and Cognition, 17(2), 384–395. doi:10.1017/S1366728913000436

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 233–277). Timonium, MD: York Press.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. The Journal of the Acoustical Society of America, 49(2B), 606-608. https://doi.org/10.1121/1.1912396

Pesantez, A. & Dellwo, V. (2022). L2 phonetic accuracy development in a foreign language setting: a longitudinal study. In B. Blecua, J. Cicres M. Espejel & M. J. Machuca (Eds.), Propuestas en fonética experimental: enfoques metodológicos y nuevas tecnologías (pp. 225–228). https://dugi-doc.udg.edu/handle/10256/20770.
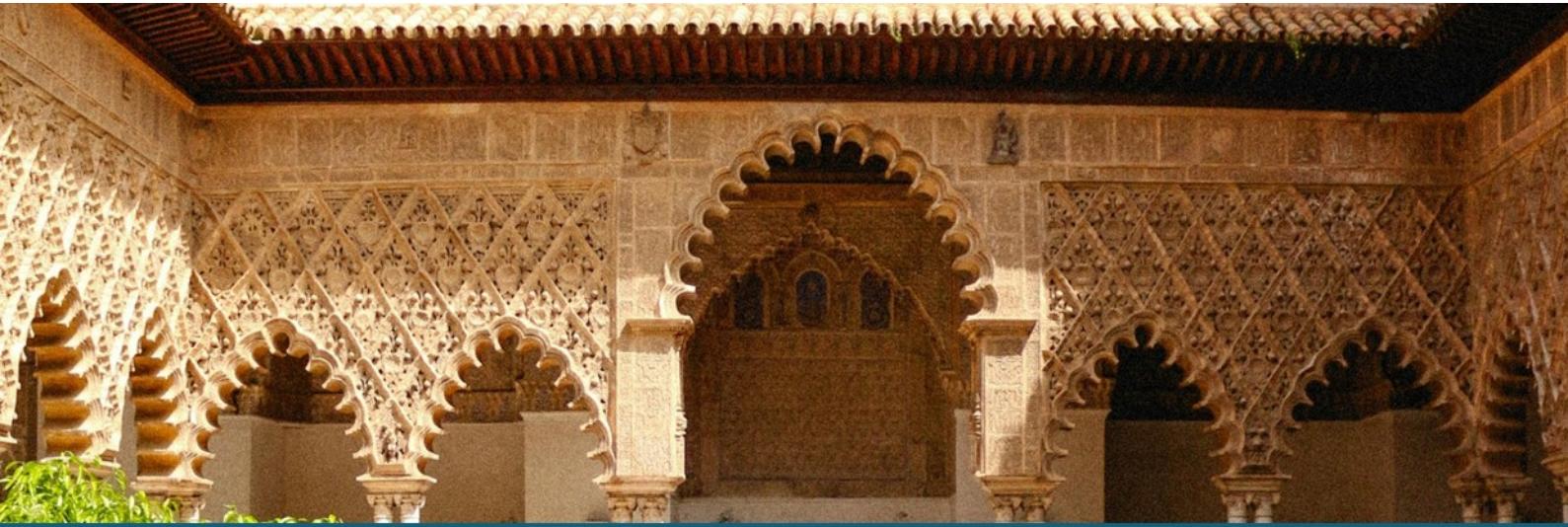
# HISTORICAL PROCESSES

# The Role of Phonetic Constraints on Historical Tapped /r/ in RP

Delia Belando
*University of Murcia*
delia.belandob@um.es

**Abstract**
This paper examines some phonetic factors influencing the now completed sound change in the phoneme /r/ in Received Pronunciation (RP). During the twentieth century, a shift from an alveolar tap [ɾ] to a postalveolar approximant [ɹ] was taken place in intervocalic position (e.g., *carry, period*). Previous empirical studies have documented the overall decline of tapped /r/ in both intervocalic and linking contexts, yet little research has adopted a micro-analytical perspective to investigate the phonetic factors conditioning tap occurrence during periods of heightened variability. Addressing this gap, the present study focuses on the 1940s–1960s, decades identified as crucial in the retreat of the alveolar tap. The analysis is based on a spoken corpus of early recordings spanning eight decades (1919-1999) and comprising 165 RP speakers and nearly 5,000 tokens of /r/. The present study concentrates on intervocalic /r/ in the mid-century period and examines both segmental and suprasegmental predictors. Segmental factors include the quality of the vowels preceding and following /r/, while suprasegmental factors consider lexical stress and syllable length. Preliminary results indicate that tap realisation is significantly favoured after close front vowels (e.g. /iː, ɪ, ɪə/) and before schwa /ə/, whereas close-mid and open back vowels (e.g. /ɒ, ɔː, ɔɪ/) and close back vowels (e.g. /ʊ, uː, ʊə/) strongly disfavor tapping. At the prosodic level, taps occur most frequently in unstressed syllables adjacent to stressed ones and are more common in shorter, particularly disyllabic, lexical items. These findings highlight the role of articulatory and prosodic conditioning in the final stages of the loss of the alveolar tap in RP and contribute to a more fine-grained understanding of completed sound change.

**Keywords:** sound change, tapped /r/, Received Pronunciation, phonetic factors

## 1. Introduction and background

The current paper explores the diachronic variation of a now complete sound change in Received Pronunciation (henceforth RP), that is, the shift in the phoneme /r/ from alveolar tap [ɾ] to postalveolar approximant [ɹ], during the twentieth century (Cruttenden, 2014; Wells, 1982). Previous empirical-based studies have shed light into the trajectory of this change, pointing to a prominent decline in the usage of tapped /r/ both in internal (e.g., *carry*, *very*, *period*) as well as linking /r/ (e.g., their own, for example, there are) positions (MacKenzie, 2017). From a sociolinguistic perspective, Fabricius (2017) and Belando (2023) suggested that the 1950s are a crucial decade for understanding this sound change given the fast social changes taking place. Thus, while the broader historical shift from tap to approximant /r/ is, to a certain extent, well-established, there is no research adopting a micro-analystical approach to delve into the linguistic predictors that may influence tap occurrence in intervocalic position. This paper aims to address this research gap by exploring phonetic factors.

## 2. Methodology

A spoken corpus of early recordings was compiled spanning almost a century (i.e., 1920s-1990s). The corpus includes a sample of 165 RP speakers (80 males and 85 females) and almost 5,000 analysed tokens of /r/. For the purpose of this study, data was analysed from the 1940s-1950s, which is the period when the use of taps exhibits higher variation. Phonetic factors have been divided into segmental and suprasegmental predictors. At the segmental level, the quality of the surrounding vowels flanking the /r/ context has been considered, hypothesizing that

vowel quality may affect the articulatory ease or likelihood of using a tap. Suprasegmental factors consider the prosodic features of the word, including lexical stress and syllable length. The number of syllables in the lexical item is analysed to explore any correlation between syllable length and the presence of a tap, as increased syllable count may lead to greater articulatory reduction (Bybee, 2015), such as schwa elision.

## 3. Results

Based on a preliminary analysis in the 1940s–1950s, the preceding vowel quality significantly affected the likelihood of tap realisation. Taps were most frequent after close front vowels (i.e., /iː, ɪ, ɪə/), in words such as *period* and *very*. In contrast, taps were notably less frequent after close-mid and open back vowels (i.e., /ɒ, ɔː, ɔɪ/) and close back vowels (i.e., /ʊ, uː, ʊə/), with tokens such as *worried, story, Tories, foreign, security*, and *during* rarely exhibiting tapping. Additionally, the quality of the following vowel showed significance: words realised with an approximant (no tap) were often followed by the weak vowel /i/, while taps tended to occur before a schwa /ə/, suggesting that the nature of the vowel after /r/ influences the tap occurrence. At the prosodic level, tapped /r/ tended to occur in unstressed syllables adjacent to stressed ones, particularly in medial positions. Shorter words, especially disyllabic items (e.g., *very*, *courage*, *series*), showed higher rates of taps than longer words.

## 4. Discussion and conclusion

The current study shows that the occurrence of tapped /r/ in mid-century RP may be strongly influenced by both segmental and prosodic factors. Taps are more frequent after close front vowels and before /ə/, while they are less common following close-mid, open back, and close back vowels, and before the happy i. Prosodically, taps tend to appear in unstressed syllables adjacent to stressed ones, as observed in previous work (Fabricius, 2017; Mackenzie, 2017). Further research may explore later decades (e.g., 1960s, 1970s), where the frequency of taps was comparatively lower.

## 5. References

Belando, Delia. 2023. Tapped /r/ in RP: a corpus-based sociophonetic study across the twentieth century. *Linguistics Vanguard*, *9*(1), 3-12. https://doi.org/10.1515/lingvan-2022-0092

Bybee, Joan. 2015. *Language change*. Cambridge: Cambridge Textbooks in Linguistics.

Cruttenden, Allan. (2014). *Gimson's Pronunciation of English*. Routledge. https://doi.org/10.4324/9780203784969

Fabricius, Anne. 2017. Twentieth-century received pronunciation: prevocalic /r/. In R. Hickey (Ed.), *Listening to the Past: Audio Records of Accents of English,* 39-65. Cambridge University Press. https://doi.org/10.1017/9781107279865.004

MacKenzie, Laurel. 2017. Frequency effects over the lifespan: A case study of Attenborough's r's. *Linguistics Vanguard*, *3*(1).

Wells, John. C. (1982). *Accents of English 2: The British Isles*. Cambridge University Press.

# Historical and dialectological variation of palatal phonemes in Extremaduran Spanish

Javier Escamilla, Ander Egurtzegi
*University of the Basque Country, Centre National de la Recherche Scientifique (CNRS-IKER)*
javierescamilla@usal.es, ander.egurtzegi@iker.cnrs.fr

## Abstract

The aim of this research is to use dialect atlases within the optics of current linguistics to obtain new results that may contribute to the analysis of diachronic and geographical phonetic variation. Specifically, our investigation attempts to explain the historical and dialectological distribution of the delateralization of /ʎ/ and the constriction of /ʝ/ in Extremaduran Spanish. Our project presents an innovative methodology that consists in the study of three atlases developed in different periods of the 20th century, which allows us to establish a chronological periodization of these specific changes and to thoroughly study the dialectal variation of the phenomena. We combine traditional data with modern methods: Bayesian models for statistical analysis, data visualization to communicate the results, and cartographical plots to draw conclusions upon them. We conclude that delateralization, common in the South of the region already in the 1930s, has increased during the last century, spreading to almost all Extremadura with minimal exceptions. Constriction of /ʝ/ does not show a regular distribution through time or space but seems consistent in Badajoz and the Central areas of Cáceres.

**Keywords:** historical linguistics, dialectology, phonetics, Bayesian statistics, cartography

## 1. Introduction and background

This research focuses on the historical and dialectological variation of the palatal phonemes in Extremaduran Spanish, which has commonly been referred as a transitional variety between Northern and Southern Spanish varieties. Specifically, we analyze diachronic and geographical data of two phenomena involving palatal consonants: delateralization of /ʎ/ to /ʝ/ (as in caballo 'horse' [kaβ̞aʝo] instead of [kaβ̞aʎo], see Navarro Tomás, 1918 & 1964; Clevería Nadal, 1993; García Mouton & Molina, 2012; Rost Bagudanch, 2014 & 2017) and the progressive constriction of /ʝ/ (as in ayer 'yesterday' [aʝer] > [ajˠer]; see Ariza, 1980; Montero Curiel, 1991). Extremaduran Spanish has been commonly ignored in several research areas, including linguistics (Montero Curiel, 1991). Thus, research on the region often resorts to traditional references, which mainly consist of local monographies and cultural research (see Fernández de Molina, 2014). These references point to a clear difference in the articulation of the palatal phonemes between Extremaduran regions (i.e., Cáceres and Badajoz) in the early 20th century. Cáceres usually maintains the distinction between the lateral /ʎ/ and the central palatal /ʝ/, while Badajoz presented delateralization of /ʎ/ since the first documented atlas (ALPI, 1962). This difference has changed through the years, as more recent atlases (i. e. CaLiEx, 2000) and papers (Montero Curiel, 2006) show how the merger between the palatal segments has spread in the last decades, leaving notwithstanding some conservative areas, such as Fuente del Maestre and Valle de Santa Ana in Badajoz or some areas of the Galician-Portuguese continuum in Cáceres. Furthermore, the articulation of /ʝ/ is not homogeneous in the region.

## 2. Methodology

Our investigation uses data extracted from three linguistic atlases developed during the 20th century: ALPI (developed in the 1930s), ALEx (formerly ALEP, late 1970s) and CaLiEx (late 1990s). However, because of the independent methodologies employed, localities and words examined in each atlas (e.g. yo, caballo) do not necessarily correspond with these in the others.

Data was coded manually from each atlas and then analyzed using the R statistical language (through the packages dplyer, ggplot2, brms, and sf, among others). Our analysis includes both external (historical data, dialectological information) and linguistic (word, place and manner of articulation, voicing, constriction level, and phonological context) variables of interest.

## 3. Results and discussion

We built two Bayesian regression models to analyze our data using the variables above as predictors: A Bernoulli model to predict the delateralization of /ʎ/ and a 3-level cumulative model to study the degree of constriction of /ʝ/. Figure 1 shows the posterior probability of delateralization through time and space, and Figure 2 that of the realization of /ʝ/. In both cases, colored municipalities are these in our dataset and each atlas represents a period (left-to-right). Our results on delateralization are in line with the observations in the previous literature. The merger of palatal phonemes was already observed in the 1930s, and it has later spread to most of Badajoz and parts of Cáceres. Fala-speaking areas bordering Portugal and localities from the Transierra Castellana still maintain /ʎ/, as well as some linguistic islands in Badajoz.

The distribution of the variation in the pronunciation of /ʝ/ is irregular. While more constricted productions are persistent in the Southwest of Badajoz, they do not spread uniformly. Nonetheless, /ʝ/-assibilation has become more common through the years, and it is now characteristic of most of Badajoz, while affrication is (almost) unattested.

## 4. Conclusion

This research combines traditional data from early atlases with modern analytical methods such as Bayesian statistics or cartographical plots to create an innovative methodology for the study of historical and dialectological phonetic variation in an under-researched variety.
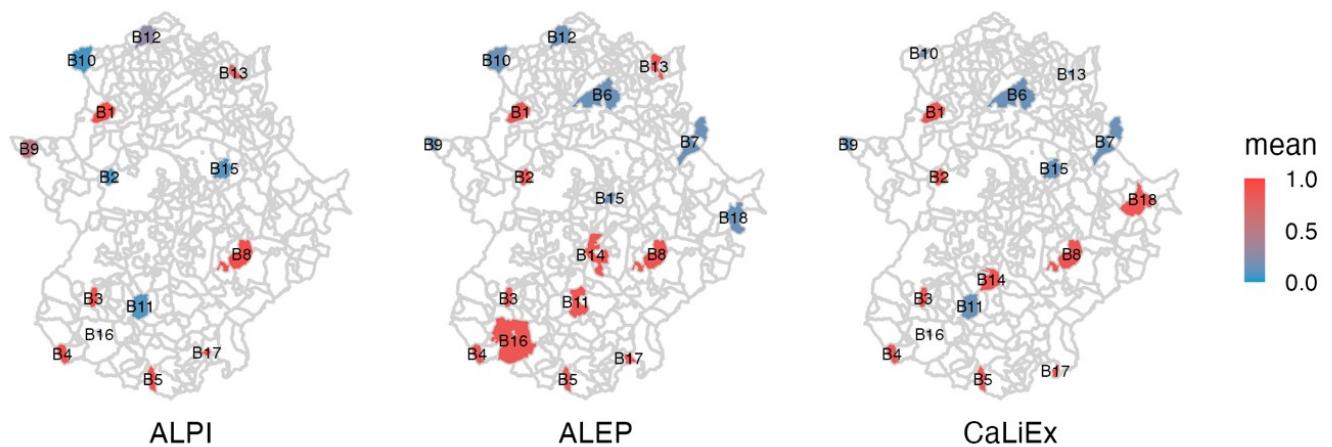


*Figure 7 Mean posterior probability of delateralization through time and space*
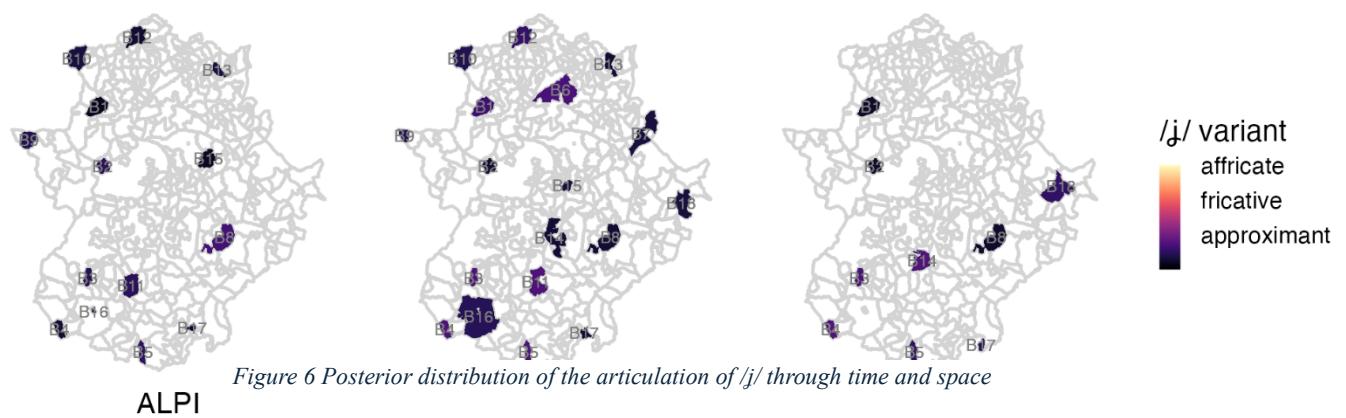


*Figure 6 Posterior distribution of the articulation of /ʝ/ through time and space*

## 5. References

Alonso Pascua, Borja (coord.) (2025-). Hacia un nuevo hito en la atlantografía dialectal ibérica: primeros pasos para la publicación del Atlas Lingüístico de Extremadura (ALEx).

Ariza, Manuel (1980). Apuntes de geografía lingüística extremeña. Anuario de estudios filológicos, 3, 21-29.

Clavería Nadal, Gloria (1993). Observaciones acerca de la historia del "yeísmo". In Ramón Lorenzo Vázquez (coord.), Actas do XIX Congreso Internacional de Lingüística e Filoloxía Románicas (229-242). Santiago de Compostela: Fundación Pedro Barrié de la Maza.

Fernández de Molina Ortés, Elena (2014). La investigación científica en el habla de Extremadura: monografías dialectales y estudios sobre fonética y fonología extremeña. Anuario de estudios filológicos, 37, 5-20.

García Mouton, Pilar & Molina Martos, Isabel (2012). The /ʎ/ - /ʝ/ merger (yeísmo) in Central Spain: Advances since the ALPI. Dialectologia, Special Issue, III: Linguistic Atlas of the Iberian Peninsula (ALPI): Progress and Perspectives, 23-42.

García Mouton, Pilar (coord.); Fernández-Ordóñez, Inés; Heap, David; Perea, Maria Pilar; Saramago, João; Sousa, Xulio (2016). ALPI-CSIC, digital edition of Navarro Tomás, Tomás (dir.), Atlas Lingüístico de la Península Ibérica. Madrid: CSIC. www.alpi.csic.es. Accessed on 10/06/2025.

González Salgado, José Antonio (2000). Cartografía Lingüística de Extremadura. https://geolectos.com/. Accessed on 10/06/2025.

Montero Curiel, Pilar (1991). Fonética extremeña en el Atlas Lingüístico de la Península Ibérica, Anuario de Estudios Filológicos, 14, 317-334.

Montero Curiel, Pilar (2006). El extremeño. Madrid: Arco/Libros.

Navarro Tomás, Tomás (1918). Manual de pronunciación española. Madrid: Centro de Estudios Históricos.

Navarro Tomás, Tomás (1964). Nuevos datos sobre el yeísmo en España. Thesaurus: Boletín del Instituto Caro y Cuervo, 19(1), 1–17.

Navarro Tomás, Tomás (1962-). Atlas Lingüístico de la Península Ibérica. Madrid: CSIC.

Rost Bagudanch, Assumpció (2014). Una panorámica del yeísmo: ¿un proceso acabado o en construcción? Revista Internacional de Lingüística Iberoamericana, 12(23), 141–163.

Rost Bagudanch, Assumpció (2017). Variation and phonological change: the case of 'yeísmo' in Spanish. Folia Lingüística, 51(1), 169-206.

# The reception of Spanish vowels in English and French through Hispanisms

Estrella Ramírez Quesada
*Universidad de Córdoba*
l62raque@uco.es

**Abstract**
This paper presents a comparative analysis of how vowels of Spanish loanwords are assimilated into English and French. To do so, it examines words of Hispanic origin that have entered the English and French lexicons in the last century. This paper therefore focuses on language contact at the phonetic level, which has been little explored compared to studies of languages in contact at the grammatical or lexical level, allowing us to understand tendencies in phonetic variation across languages.

**Keywords:** vowels, phonetics, variation, Spanish, language contact

## 1. Introduction and background

Since Trubetzkoy (1939) introduced the concept of phonological deafness, whose relevance has been highlighted, among others, by Gil (2007), we have had a phonetic and phonological explanation of how we perceive the sounds of a foreign language. By virtue of this phenomenon, the native ear gives preference to some features over others when "understanding" the sounds of a foreign language, based on the characteristics of the target language. In this sense, the application of this concept allows us to understand how, from a phonetic perspective, borrowings from a language function in the target linguistic system.

However, numerous factors intervene in the functioning of borrowings at the phonic level, not only phonological but also phonetic and orthographic, all of which are interrelated. Vendelin and Peperkamp (2006) have mentioned that this interrelationship has not been routinely studied and that a variety of factors are involved in the introduction of oral loanwords. Thus, Peperkamp and Dupoux (2003) argue that greater emphasis should be placed on phonetics, and Boersma and Hamann (2009) mention that the structure of the target language —including phonetics— plays a fundamental role in perception and, therefore, in adoption as well.

## 2. Methodology

With this in mind, this paper aims to show how Spanish vowels are received through the Spanish words present in English and French. By studying the phonetic and phonological systems of the source and target languages, we will understand the phonetic variation experienced by vowels and, furthermore, why some features are privileged over others. Despite the increasing importance of loanword phonology studies in the last two decades (Kang, 2010: 2295), there is still work to be done, as widely spoken languages, such as those presented in this paper, have barely been addressed from a contrastive perspective in the field of phonetic borrowing.

To this end, we take into account, thanks to dictionaries and pronunciation repositories (such as Forvo), whose recorded pronunciations can be compared by using Praat, the corpus formed by the Spanish words that entered English in the 20th century (see Cannon, 1996) and the newly appearing Hispanic words in the Diccionnaire de la Académie Française (9th edition). After methodological restrictions, the final corpus consists of 63 Spanish loans for English and 20 Spanish loans for French. We thus deal with the pronunciation of recent Hispanic words consolidated in the standard by speakers, mainly in European areas. The 20th century constitutes a suitable period to observe the accommodation of words because there has been time enough for consolidation, avoiding vacillation of forms, and, given that this period is also relatively recent, ways of adaptation are indicative of current phonetics.

## 3. Preliminary results

In the third part, we present a systematization of the phonetic variation experienced by Spanish vowels (considering their wide field of dispersion) in the target languages. The Spanish sounds considered are [i], [u], [e], [o], [a]. On its behalf, the number of English monophthongal sounds goes up to twelve units: [ʌ], [ʊ], [uː], [ɪ], [iː], [ɛ] (or [e]), [ɒ] (or [ɑ]), [ɔː], [æ], [ɑː] (or [ɑ]), [ɜː] (or [ɜ]), [ə], whereas French sounds considered are [a], [ɑ], [e], [ɛ], [ə], [i], [o], [ɔ], [ɞ], [œ], [u], [y], [ă], [ɛ̃], [õ]. Diphthongs are also contemplated in the three languages.

The adaptation of Spanish vowels in loans confirms that there is a prevalence of features associated with timbre in stressed vowels, with more varied solutions in unstressed positions, as preliminary results (some of them already in press) have shown. There is a tendency in English to have the long vowels [iː] and [uː] in stressed syllables, as well as the short vowels [ɪ] and [ʊ] in non-stressed positions for the Spanish [i] and [u], respectively. The Spanish sounds experiencing more variation in loanwords are mid-vowels [e], [o], and the open vowel [a], being adapted phonetically or orthographically in due course, especially in unstressed positions. As English and French have more vowels, it is easier to assign a specific sound to Spanish units, but the structure of the recipient language is favoured and, for instance, diphthongs are preferred in word-final positions instead of short vowels different from [ə] in English and nasals variants are to appear in French contexts of nasalisation despite not being so in origin. Further results are to be presented in relation to the different sounds.

## 4. Discussion and conclusion

Finally, we will discover each language's preference for specific sound features based on the phonetic structure of the target language. Preliminary results suggest the role of orthography is less important in the subsystem of vowels than in the case of consonants and, in both cases, being dependant of oral adaptation. This indicates that oral introduction of loans is more important in many cases than receiving the written form. The loanword adaptation relies mostly on the phonological features and structure of the recipient language but takes into consideration the similarity in phonetic terms of languages in contact. Lastly, with this knowledge, we will be able to delve deeper into the phonetic variation of vowels, from an interlinguistic point of view. All of this can lead to an improvement in understanding phonetic variation in the field of language contact and contrastive linguistics in all its applications, such as the study of interference for foreign language learners.

## 5. References

Académie Française (2024). Dictionnaire de l'Académie française <https://www.dictionnaire-academie.fr/>

Boersma, Paul & Hamann, Silke (2009). Loanword Adaptation as First-Language Phonological Perception. In Calabrese, Andrea & Wetzels, W. Leo (eds.), Loan Phonology (11-53). Amsterdam: John Benjamins Publishing Company.

Cannon, Garland (1996). Recent Borrowings from Spanish into English. In Rodríguez González, Félix (ed.), Spanish loanwords in the English language (41-60). Berlin: Mounton de Gruyter.

Forvo (2025). Forvo. The Pronunciation Dictionary. <https://forvo.com/>

Gil, Juana (2007). Fonética para profesores de español: de la teoría a la práctica. Madrid: Arco/Libros.

Kang, Yoonjung (2010). Tutorial Overview: Suprasegmental Adaptation in Loanwords. Lingua 120: 2295-2310. DOI: 10.1016/j.lingua.2010.02.015

Peperkamp, Sharon & Dupoux, Emmanuel (2003). Reinterpreting Loanword Adaptations: The Role of Perception. In Solé, Maria-Josep, Recasens, Daniel & Romero, Joaquín (eds.),

Proceedings of the 15th International Congress of Phonetic Sciences (367-370). Barcelona: Universitat Autònoma de Barcelona.

Trubetzkoy, Nikolái S. (1939 [1973]). Principios de fonología. Madrid: Cincel.

Vendelin, Inga & Peperkamp, Sharon (2006). The Influence of Orthography on Loanword Adaptations. Lingua, 116, 996-1007. DOI: 10.1016/j.lingua.2005.07.005

# A Contribution to the Historiography of Spanish Diachronic Phonetics: Henri Gavel's 1920 Essay on the Evolution of the Pronunciation of Castilian from the 14th Century According to the Theories of Grammarians and Other Sources

Daniel M. Sáez Rivera
*University of Granada*
dmsaezrivera@ugr.es

**Abstract**

Due to the high degree of specialization required in phonetic studies—which demands strong technical expertise—historiographical approaches, more humanistic, have not traditionally been a central focus of the discipline. Nevertheless, there is currently a growing trend in the historiography of phonetics, both internationally and within Spain. Although numerous studies in Linguistic Historiography (LH) have addressed the history of Spanish phonetic research, they have largely overlooked the historiography of historical phonetics itself. This paper aims to present, examine, and reclaim the figure of Henri Gavel and his Essai sur l'Évolution de la Prononciation du Castillan depuis le XIVme siècle d'après les théories des grammairiens et quelques autres sources (Paris: Édouard Champion, 1920). Gavel's work stands as a virtually unknown precedent to Amado Alonso's posthumously published masterpiece De la pronunciación medieval a la moderna (1955, 1967), which was groundbreaking in its direct use of Golden Age Spanish grammatical texts to trace the history of the Spanish language—a line of inquiry already initiated by Colombian scholar Rufino José Cuervo in the late 19th century.

**Keywords:** Historiography of Phonetics, History of Spanish Pronunciation, Old Spanish Grammars, Henri Gavel, Amado Alonso

## 1. Introduction and background

Due to the high degree of specialization required in phonetic studies—which demands strong technical expertise—historiographical approaches, more humanistic, have not traditionally been a central focus of the discipline (Koerner, 1995). Nonetheless, there is currently a strong trend in the historiography of phonetics, both internationally and within Spain. The Linguistic Historiography (HL) of Spanish phonetics is undergoing a clear resurgence, as noted in the prologue to the special issue edited by Quijada van den Berghe (2024). Although these studies have addressed and reassessed the work of phoneticians from the Spanish Golden Age (16th and 17th centuries), as well as from the 18th and 19th centuries—particularly pioneers from the late 19th century like Araujo and from the early 20th century like Navarro Tomás—the HL of diachronic phonetic studies remains scarce.

However, as early as Catalán (1974: 275–276), the work of Amado Alonso and his De la pronunciación medieval a la moderna ('From Medieval to Modern Pronunciation') is defended—a work that draws on early grammars and phonetic treatises as sources for the history of Spanish pronunciation. In a footnote, Catalán cites Cuervo (1895) and Gavel (1920) as precedents. While the contributions of Cuervo and Alonso are well known (their importance as pioneers in using grammars as historical linguistic sources is highlighted by Sáez Rivera, 2023), the figure of Gavel has gone almost completely unnoticed. Nevertheless, he was a respected Romanist and a Basque studies specialist in his native France (see the obituary by Boussagol, 1960), although his 1920 work was subject to criticism even upon its release (notably the critical review by Bourciez, 1922).

In this study, we will present Henri Gavel and his 1920 work, exploring why it has remained virtually unknown, positing hypotheses such as the absence of reprints due to certain errors—e.g., dating the Cantar de mio Cid to the 14th century based on the manuscript's date. At the same time, we aim to revalue Gavel's work from a historiographical perspective,

highlighting his role as an advocate and disseminator of the work of prominent phoneticians and philologists of the period such as Navarro Tomás, Araujo, and Menéndez Pidal, as well as his pioneering stance on several still-relevant diachronic debates.

## 2. Methodology
In studying the work of Henri Gavel (1880–1960), particularly his Essai sur l'Évolution de la Prononciation du Castillan depuis le XIVme siècle d'après les théories des grammairiens et quelques autres sources (Paris: Édouard Champion, 1920), we follow the analytical approaches of Linguistic Historiography as proposed by Koerner (1978). Rather than a cumulative historical narrative that emphasizes recent achievements while discarding earlier work as flawed—an attitude commonly found in phonetics (see Koerner 1995: 172)—we seek to contextualize Gavel's 1920 work within its own time, recognizing both its advances and its errors. In doing so, we aim to conduct a form of inspirational Linguistic Historiography (Sáez Rivera, 2020), so that Gavel's work can serve both as a positive model for studies on the history of Spanish phonetics and also as a cautionary tale.

## 3. Results
The Essai sur l'Évolution de la Prononciation du Castillan depuis le XIVme siècle (1920) was Gavel's first work, the result of his doctoral thesis defended in Toulouse before a committee that included renowned historians of the Spanish language such as Américo Castro and Ramón Menéndez Pidal (Boussagol, 1960: 245).

Each chapter of Gavel's work follows a similar structure: it begins with a synchronic description of Spanish phonetics, drawing primarily on the work of Tomás Navarro Tomás, thereby contributing to the dissemination of his ideas abroad. This is followed by a description of synchronic phonetic variation in the period, drawing from a variety of sources and distinguished by the inclusion of numerous personal observations and direct informant testimonies. He then analyzes historical phonetic and orthographic descriptions, showing direct familiarity with manuscript sources—particularly the Poema del Mío Cid, for which, following Menéndez Pidal's edition (1908–1911), he offers a detailed account. However, he erroneously dates the text to the 14th century based on the manuscript's date, hence the reference to the 14th century in the book's title.

To trace developments during the Spanish Golden Age, Gavel relies primarily on grammarians and phoneticians from the 16th and 17th centuries. However, he never consults them directly but rather through secondary sources—especially Cuervo (1895) and the foundational documentary compilation by La Viñaza (1893), a cornerstone in Spanish historiographical studies.

What stands out in Gavel's analysis is his Romanist and comparative approach, which leads him to contrast phonetic variation in Spanish with that of other European languages—especially Germanic and Romance ones—but also Basque (a language in which he was a recognized expert, even in its phonetic aspects) and Southern French dialects. He also proposes theories on the influence of Gregorian chant in Romance spelling, anticipating Wright's 1982 controversial theories on the impact of the Carolingian reform of Latin on the emergence of Romance-language writing.

## 4. Discussion and conclusion
Alongside many accurate phonetic and variationist observations, Gavel's work contains several errors (some of which were already noted by Bourciez in 1922), such as considering the Mío Cid a 14th-century work and relying on methodologies limited by his time—like using edited texts and documentary collections instead of original sources. In this regard, Amado Alonso

made significant advances, supported by access to the rich early holdings of Harvard University's library, where he taught from 1946 until his death in 1952.

Despite its flaws, Henri Gavel's contribution represents an interesting little-known chapter in the international history of Hispanism and in the history of phonetic studies of Spanish pronunciation. His work deserves to be explored further and connected to the rest of his research—particularly his work as a Basque studies specialist, with publications such as Éléments de phonétique basque (Champion, 1920), and also as a Hispanist, having authored, for example, a Grammaire historique de la langue espagnole (1951).

## 5. References

Alonso, A. (1955). De la pronunciación medieval a la moderna. Vol. 1 Madrid: Gredos.

Alonso, A. (1967). De la pronunciación medieval a la moderna. Vol. 2. Madrid: Gredos.

Boussagol G. (1960). Henri Gavel (1880-1960). Bulletin Hispanique, 62(2), 245-246.

Bourciez E. (1922). H. Gavel, Essai sur l' évolution de la prononciation du Castillan depuis le XIVe siècle, d'après les théories des Grammairiens et quelques autres sources. Bulletin Hispanique, 24(4), 381-384

Cuervo, Rufino José (1895). Disquisiciones sobre la antigua ortografía y pronunciación castellanas. Revue Hispanique, 2, 1-69.

Catalán, Diego (1974). Lingüística iberorrománica: crítica retrospectiva. Madrid: Gredos.

Gavel, Henri (1920). Essai sur l'Evolution de la Prononciation du Castillan depuis le XIVme siècle d'aprés les théories des grammariens et quelques autres sources. Paris: Édouard Champion.

Koerner, E. F. K. (1978). Four Types of History Writing in Linguistics. In Toward a historiography of linguistics: selected essays (pp. 55-62). Amsterdam: John Benjamins.

Koerner, E. F. K. (1995). History and Historiography of Phonetics. A state-of-the-art account. Professing linguistic historiography (pp. 171-202). Amsterdam: John Benjamins

Quijada Van den Berghe, Carmen (2024). Historiografía de la fonética y fonología en la tradición hispánica / Historiography of Phonetics and Phonology in the Hispanic Tradition. Loquens, 11(1-2), 1-2 https://loquens.revistas.csic.es/index.php/loquens/article/view/149/372

Sáez Rivera, Daniel M. (2020). La concepción del cambio lingüístico en Menéndez Pidal a la luz de la sociolingüística histórica. In Eugenio Bustos Gisbert y Rafael Cano Aguilar (Eds.), Noventa años de Orígenes del español (pp. 43-75). Valencia: Tirant Lo Blanch.

Sáez Rivera, Daniel M. (2023). Chapter 3: Las gramáticas del español como fuentes de datos lingüísticos. In Steven Dworkin, Gloria Clavería Nadal y Álvaro S. Octavio de Toledo y Huerta (Eds.), Lingüística histórica del español / The Routledge Handbook of Spanish Historical Linguistics (pp. 28-38). London: Routledge.

Viñaza, Conde de la. (1893). Biblioteca histórica de la Filología castellana. Madrid: Real Academia Española.

Wright, Roger (1982). Late Latin and Early Romance. In Spain and Carolingian France. Liverpool: Francis Cairns.

# Reconstructing Southern Spanish through the verses of Diego Hurtado de Mendoza

Laura Crespo Quesada
*Autonomous University of Madrid*
laura.crespoq@gmail.com

**Abstract**

This study investigates the phonetic imprint of southern Spanish variety in the poetry of Diego Hurtado de Mendoza (1503–1575). From a metrical-rhythmic framework, the research examines rhyme and metric patterns across a corpus comprising Mendoza's *Poesía erótica*. The findings reveal the coexistence of linguistic variants, including syllabic restructuring, bilabial and sibilant neutralization, and the gradual loss of Latin initial /f/ aspiration. These results not only represent the phonetic landscape of 16th-century Granada, but they also demonstrate the utility of poetic metrics and rhyme as methodological tools for reconstructing historical phonetic variation. The study contributes to filling the blanks in the Habsburg-era phonetic research and enhances our knowledge of language standardization in Early Modern Spanish.

**Keywords:** Historical Phonetics, Southern Spanish variety, metric and rhyme, Diego Hurtado de Mendoza, Golden Age Spanish poetry

## 1. Introduction and background

This study is situated within the theoretical framework of historical phonetics and phonology, with a particular focus on the phonetic component during the Habsburg period in Spain. The reconstruction of historical phonetics and phonology dates back to the late 19th century, and became more significant throughout the 20th century and the early 21st century. Numerous studies have concentrated on the historical reconstruction of Castilian Spanish, led by eminent scholars such as Menéndez Pidal, Alarcos, Amado and Dámaso Alonso, Echenique, Lapesa, and more recently, Ariza, Cano Aguilar, Penny and Pla Colomer.

However, in recent years, an evident gap has emerged concerning studies on historical phonetics and phonology specifically during the Habsburg era. This project arises from the necessity to revisit and expand the description of the phonetic component and peninsular diatopic variation in Golden Age Spanish, offering an innovative methodology that enables the extraction of logical conclusions grounded in a thorough and systematic analysis of the corpus. This paper aims to explore the state of the language through a detailed analysis of the metrics and rhyme in the poetry of Diego Hurtado de Mendoza (1503–1575, Granada). The selected corpus includes Mendoza's Poesía erótica (Díez Fernández 1995) and the edition by Díaz Larios & Gete Carpio (1990), based on the princeps edition of 1610, comprising 33 sonnets.

## 2. Methodology

Building on the philological tradition of Alarcos, D. Alonso, Cuervo, Lapesa, Montoliú, Penny, and more recently, Echenique o Pla Colomer, this study treats metrics and rhyme as indispensable tools for inferring the evolutionary traits of the phonetic component of Spanish, as well as for establishing relative and absolute chronologies derived from them. Fixed metrical patterns, combined with the analysis of rhyme (understood as the repetition of identical vowel and consonant segments from the last stressed syllable onward), enables the reconstruction of phonetic characteristics in Spanish across different historical stages.

While poetic language has often been viewed as artificial and distinct from popular speech, studies, such as Echenique & Pla Colomer (2013), suggest that poets worked within authentic linguistic foundations. Thus, verse can reflect genuine linguistic variation, despite occasional poetic licenses. This approach enables the identification of diachronic phonetic traits in the

corpus, with a particular focus on phenomena related to vocalism and consonantism: syllabic restructuring, bilabial and sibilant desphonologization, and the loss of aspiration of the Latin initial /f/.

## 3. Results

Diego Hurtado de Mendoza's poetic corpus exhibits both diphthongized and heterosyllabic variants in imperfect and conditional verb forms, reflecting the coexistence of competing linguistic norms in 16th-century spoken Spanish. This variation points to an ongoing process of phonological reconfiguration in which standardization had not yet fully dominated regional forms. As for bilabial phonemes, the data indicate that desphonologization was already underway in the southern Iberian Peninsula, although this phenomenon is not noticeably prominent in Mendoza's language, suggesting a transitional stage. Regarding sibilants, Mendoza's verse provides evidence of incipient neutralization affecting the predorsodental, apicoalveolar, and prepalatal pairs, a process that aligns with broader southern tendencies, though still incomplete. Finally, with respect to the evolution of Latin initial /f/, signs of aspiration remain detectable in the poetic texts, albeit with clear signs of phonetic weakening and loss.

## 4. Discussion and conclusion

This investigation confirms that the linguistic choices in Mendoza's poetry reflect the sociolinguistic realities of 16th-century southern Spanish. The cultivated zeal of poets and their proximity to centers of power contributed to the standardization of the language, aligning it with the most prestigious sociolects of the time. Moreover, it has shown an active linguistic variability, revealing a dynamic interaction between emerging standard norms and regional speech patterns.

The analysis of metrics and rhyme has proven to be a fruitful methodological tool for reconstructing the phonetic system of the language across different chronological stages, especially in contexts where direct phonetic data are unavailable. This approach not only enriches our understanding of early modern dialectal variation, but it also invites reconsideration of poetic language as a type of authentic linguistic expression rather than mere literary feature.

## 5. References

Alarcos, Emilio (1988). De nuevo sobre los cambios fonéticos del siglo XVI. In M. Ariza, A. Salvador & A. Viudas (eds.), Actas del I Congreso Internacional de Historia de la Lengua Española, vol. 1 (47–60). Madrid: Arco Libros.

Alarcos, Emilio (1991 [1965]). Fonología española. Madrid: Gredos.

Alonso, Amado (1969–1976 [1955]). De la pronunciación medieval a la moderna en español. R. Lapesa (ed.), Madrid: Gredos.

Alonso, Dámaso (1962). Temas y problemas de la fragmentación fonética peninsular. In D. Alonso, Obras completas, I: Estudios lingüísticos peninsulares (13–291). Madrid: Gredos.

Ariza, Manuel (2012). Fonología y fonética históricas del español. Madrid: Arco Libros.

Cano Aguilar, Rafael (1998). Historia del andaluz. In A. Narbona, R. Cano Aguilar, & R. Morillo (eds.), El español hablado en Andalucía (27–122). Barcelona: Ariel.

Cano Aguilar, Rafael (2002 [1988]). El español a través de los tiempos. Madrid: Arco Libros.

Cuervo, Rufino José (1895). Disquisiciones sobre antigua ortografía y pronunciación castellanas. Parte I. Revue hispanique, 2, 1–69.

Cuervo, Rufino José (1898). Disquisiciones sobre antigua ortografía y pronunciación castellanas. Parte II. Revue hispanique, 5, 273–313.

Echenique Elizondo, María Teresa (2013). Fuentes y vías metodológicas para el estudio de la pronunciación castellana a través de su historia. De Amado Alonso al siglo XXI. In M. T. Echenique Elizondo & F. J. Satorre Grau (eds.), Historia de la pronunciación de la lengua castellana (31–59). Valencia: Tirant Humanidades.

Echenique Elizondo, María Teresa & Pla Colomer, Francisco (2013). Reconstrucción fonética y periodización a la luz de la métrica y la rima. In M. T. Echenique Elizondo & F. J. Satorre Grau (eds.), Historia de la pronunciación de la lengua castellana (63–106). Valencia: Tirant Humanidades.

Lapesa, Rafael (1981). Historia de la lengua española. Madrid: Gredos.

Menéndez Pidal, Ramón (1933). El lenguaje del siglo XVI. Cruz y Raya, 6, 327–383.

Michelena, Luis (2011 [1963]). Lenguas y protolenguas. In J. A. Lakarra & Í. Ruiz Arzalluz (eds.), Obras Completas, I. San Sebastián/Vitoria: Seminario de Filología Vasca «Julio de Urquijo» / Diputación Foral de Gipuzkoa / Universidad del País Vasco.

Montoliú, Manuel de (1947). Notas sobre algunos cambios fonéticos de la lengua española en el siglo XVI. Boletín de la Real Academia Española, 26, 369–383.

Navarro Tomás, Tomás (1966). Estudios de fonología española. Nueva York: Las Américas Publishing Company.

Penny, Ralph (2006 [1993]). Gramática histórica del español. Barcelona: Ariel.

Pla Colomer, Francisco Pedro (2020). Historia de la lengua y métrica medieval. Boletín de la Real Academia Española, 100(322), 635–662.

Pla Colomer, Francisco Pedro (2024). Evolución de las sibilantes castellanas en su contexto iberorrománico: una historia latente en los testimonios poéticos. Romance Philology, 76(1-2), 45-70.
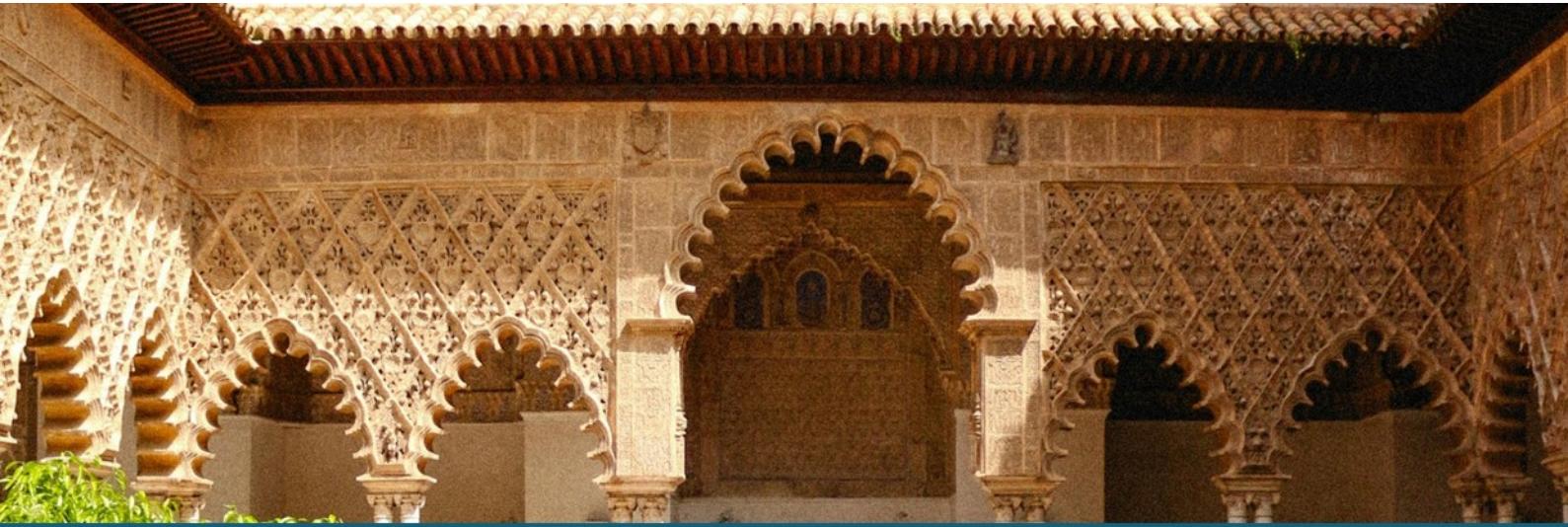
# SOCIOPHONETICS

# Phonetic Correlates of Voices in Gendered Product Advertising

Inke Henneberger, Melanie Weirich
*Friedrich-Schiller-Universität Jena*
inke.henneberger@uni-jena.de

## Abstract

This study examines phonetic characteristics of voices in German radio advertising, focusing on gender-related product contexts. Product categories were classified by gender connotation through a rating experiment. Linear mixed models show that only female-categorized voices vary significantly in fundamental frequency (f0) and its variation: higher values occur with female-connoted products. Male-categorized voices show no comparable context-dependent differences. The findings suggest that female-categorized voices in radio ads are more strongly adapted to product contexts than male-categorized voices.

**Keywords:** advertising voices, gender-related connotation, f0

## 1. Introduction and background

Research shows male-categorized voice-overs dominate advertising globally (Matthes et al. 2016). According to Leonardo et al. (2020), female-categorized voices in US advertisements appear mainly in household contexts and are underrepresented in other areas. Rodero et al. (2012) found voices match gendered products but do not significantly affect attention or brand recall, indicating tradition drives voice choice.

Regarding phonetic parameters an analysis of German TV ads found that female-categorized voices with higher f0 appear more in female-connoted products, while male-categorized voices with higher f0 rarely occur in male-connoted contexts (Henneberger 2021). The present study analyses phonetic variation in German radio advertisements, focusing on gender-related product connotations. Prior work suggests deeper voices are used for more expensive products, with female-categorized voices tending to show a stronger pattern (Henneberger & Weirich to appear).

## 2. Methodology

Corpus: 222 advertisements featuring 296 distinct voices (171 categorized as male (m), 125 as female (f)) were collected between November 2023 and January 2024. The spots were recorded during morning show programs on four German radio stations in two different broadcast regions, targeting both older and younger listener demographics in each region.

Product Context and Gender-Related Connotation: All spots in the corpus were assigned to 14 product categories. In a rating experiment (N = 126; 74 female, 45 male, 7 diverse; age: 18–80, mean 34 years), participants were asked to rate the target audience of the product categories on a scale from 1 [female] to 7 [male]. Based on the average ratings, the advertising contexts were classified as predominantly female, predominantly male, or ambivalent (Table 1).

| Predominantly female contexts Nf= 38, Nm=56 | Predominantly male contexts Nf=22, Nm=44 | Ambivalent contexts Nf=30, Nm=42 |
|---|---|---|
| **Shopping & Fashion (2.29)** **Furniture & Interior Design (2.69)** **Musicals & Theatre (2.80)** **Food & Household goods (2.87)** | Electrical appliances & Technology (5.08) Hardware store (5.48) Cars (5.58) | Vacation & Travel (3.62) Concerts (no classical music) (3.79), Gastronomy (3.89) Entertainment Media (TV, press) (3.98) |

| Health & Medicine (2.93) | Sports (4.56) |
| --- | --- |
| | Finance & Insurance (4.92) |

Table 1: Classification of product categories by gender-related connotation based on the rating experiment, mean ratings on a scale from 1 [female] to 7 [male] in parentheses (Note that spots containing job offers, self-promotion of the radio stations, or not fitting any of the categories created for the rating experiment (Nf=35, Nm=28), are not included in the subset of this study)

## 3. Analysis

To reduce background noise like music or effects, spots were processed using an audio source separation tool (LALAL.AI 2024). The quality of this separation was checked prior to analysis. For acoustic analysis, recordings were annotated using WebMAUS (Kisler et al. 2017) and checked manually. Average fundamental frequency (f0) and its variation (SD) were extracted using Praat (Boersma & Weenink 2022). Statistical analyses were done in R (R Core Team 2022), and linear mixed effects models (LMMs) using lme4 (Bates et al. 2015) were run to test the effects or possible interaction of gender-related connotation (predominantly male, predominantly female, ambivalent) and voice gender on f0 parameters. The speaker was added as random effect. Models were finalized using likelihood ratio tests and post hoc Tukey tests were run with emmeans (Lenth 2016).

## 4. Results

The final model with mean f0 as the dependent variable includes a significant interaction between the categorized gender of voices and the gender-related connotation of product categories ($\chi^2(2) = 8.48$, $p < .05$). Post hoc Tukey tests show significant differences only for female-categorized voices: a higher f0 is found for predominantly female-rated products (mean f0 = 240 Hz) compared to predominantly male-rated products (mean f0 = 201 Hz, $p < .001$) and ambivalent product contexts (mean f0 = 213 Hz, $p < .01$). Male-categorized voices do not show significant f0 differences across product contexts (see Figure 1, left plot).
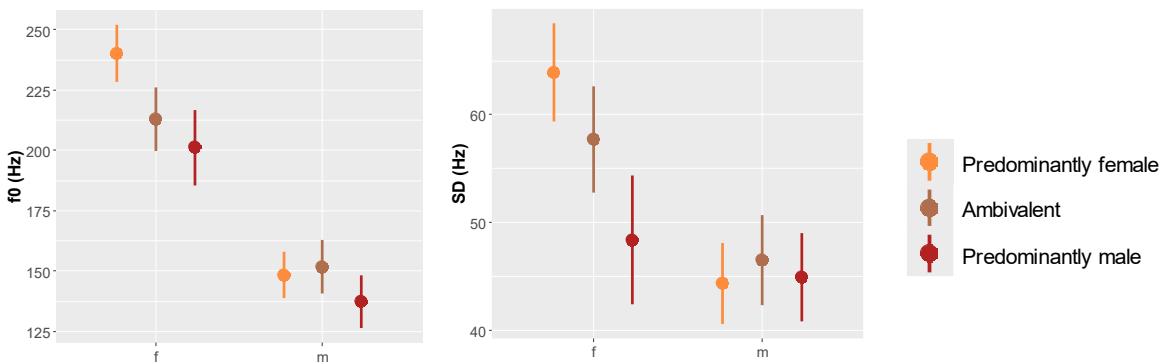


Figure 1: Model plot of predicted values for f0 (left) and f0-Variation (SD, right) showing the interaction between voice gender and gender-related connotation.

Regarding f0 variation (SD, see Figure 1, right plot), the final model also includes a significant interaction between voice gender and product connotation ($\chi^2(4) = 17.03$, $p < .01$). Post-hoc tests reveal that female-categorized voices exhibit significantly lower f0 variation in predominantly male-rated contexts (48 Hz) than in predominantly female-rated contexts (64 Hz, $p < .001$) and ambivalent product contexts (58 Hz, $p < .05$). No significant variation differences are found for male-categorized voices.

## 5. Discussion and conclusion

This study examines whether product contexts in radio advertisements and their gender-related connotations affect the f0 parameters of advertising voices, and if this effect differs between

female- and male-categorized voices. Results show that only female-categorized voices display significant context-dependent differences, with higher mean f0 and f0 variation in female-connoted product contexts. Outside gender-congruent contexts, both parameters are lower, appearing less stereotypically feminine. This aligns with observed stronger f0 differences between products with different price categories (Henneberger & Weirich to appear). Overall, female-categorized voice selection in German radio ads appears to be more influenced by product fit than male-categorized voices. More fine-grained analyses regarding the link between f0 parameters and the individual product categories show interesting patterns revealing a non-linear relationship between gender-connotation and acoustic phonetic parameter. For example, the gender-related target group of the entertainment media product category is assessed as rather balanced, but female-categorized voices in this category have the lowest f0 in the corpus on average. Future research will explore this in more detail also including other phonetic parameters like formants and voice quality.

## 6. References

Bates, Douglas, Mächeler, Martin, Bolker, Ben & Walker, Steve (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48.

Boersma, Paul & Weenink, David (2022). Praat: Doing phonetics by computer Version 6.0.40. http://www.praat.org. Accessed on 22/05/2025.

Henneberger, Inke (2021). Stereotype Werbestimmen – Untersuchung von geschlechterspezifischen Stimm- und Sprechtypen in der deutschen Primetime-Werbung anhand phonetischer Parameter. Master thesis, Halle (Saale): Martin-Luther-Universität Halle-Wittenberg.

Henneberger, Inke & Melanie Weirich (In press). Werbestimmen: Zusammenhang zwischen inhaltlichen Aspekten von Radiowerbung und Parametern der Grundfrequenz. In Proceedings Phonetik und Phonologie 20. Halle (Saale): Martin-Luther-Universität Halle-Wittenberg.

Kisler, Thomas, Reichel, Uwe D. & Schiel, Florian (2017). Multilingual processing of speech via web services. Computer Speech & Language, 45, 326–347.

LALAL.AI. (2024). Vocal Remover & Instrumental AI Splitter. https://www.lalal.ai/. Accessed on 22/05/2024.

Lenth, Russell V. (2016). emmeans: Estimated Marginal Means, aka Least-Squares Means. https://CRAN.R-project.org/package=emmeans. Accessed on 22/05/2025.

Leonardo, Angelas, Pollard, Alyssa & Clark, Roger (2020). The gender of product representatives and voice-overs in television commercials: An update. Sociology Between the Gaps: Forgotten and Neglected Topics, 5(1), 1–6.

Matthes, Jörg, Prieler, Michael & Adam, Karoline (2016). Gender role portrayals in television advertising across the globe. Sex Roles, 75(7), 314–327.

R Core Team (2022). R: A language and environment for statistical computing. https://www.R-project.org. Accessed on 22/05/2025.

Rodero, Emma, Larrea, Olatz & Vázquez, Marina (2012). Male and female voices in commercials: Analysis of effectiveness, adequacy for the product, attention and recall. Sex Roles, 68, 349–362.

# Muy de Granada [´mu ð̞e ɣra´na / ´muj ð̞e ɣra´nað̞a]. A case of multifaceted phonetic variation

Rocío Cruz Ortiz, Antonio Manjón-Cabeza Cruz
*Universidad de Granada*
rociocruz@ugr.es, amanjoncabeza@ugr.es

## Abstract

This paper analyses the monophthonguisation of the Spanish adverb muy /muj/ or /mwi/ 'very' into mu [mu] in terms of phonetic, dialectal, sociolinguistic, pragmatic and stylistic variation. Monophthongisation is common in Andalusian Spanish and sporadic in other varieties of Spanish. In Granada, it is the most common variant, especially in younger speakers, speakers with low levels of education, and women. Additionally, lexical diffusion phenomena appear to be at play, as certain high-frequently words trigger monophthonguisation, such as *bien* 'well', *bueno* 'good', and *bonito* 'pretty'. The rate of occurrence of the [mu] variant in our data (58.5%) leads us to question the stigma attached to monophthonguisation in Granada. This is demonstrated by the weak sociolinguistic stratification observed, since all the social groups studied use both variants to a greater or lesser extent.

**Keywords:** sociolinguistics, variationism, Andalusian Spanish, adverb *muy* 'very', prestige

## 1. Introduction and background

Despite its widespread occurrence in Andalusia and elsewhere in Spain, the phonetic variation of the adverb *muy* 'very', which can appear in its full or monophthong realisation within the same speaker, has received little attention. There are no historical or geolinguistic studies analysing this. We have observed the sporadic appearance of the written form mu since the 14th century, but only in isolated examples. It was not until the 19th century that numerous examples of this form began to appear, particularly among realist and naturalist writers, as well as Andalusian folklorists. This is also indirectly confirmed in linguistic atlases of peninsular European Spanish (ALEA, ALECMAN). Notably, this phonetic variation is not mentioned in the works on phonetics of the Royal Spanish Academy. Some 20th century bibliography refers to the form without the diphthong as vulgar or typical of gente tosca 'coarse people'. However, data from our Granada corpus refute this as [mu] is the most common variant.

## 2. Methodology

We present a variationist sociolinguistic study which uses inferential statistical techniques, such as logistic regression, to reveal that there are many factors that influence variation. The data has been extracted from the PRESEEA corpus (Project for the Sociolinguistic Study of Spanish in Spain and America), especially from the Granada subcorpus, This is an oral corpus consisting of 54 semi-structured interviews, each averaging 40 minutes in length. The sample is uniformly stratified according to age (three generational groups of 18 informants each), gender (27 women and 27 men) and educational level (three groups of 18 informants each depending on their level of education (primary, secondary or university education)). As this is an preliminary study, we have decided to analyse the first ten minutes of each interview, totalling 701 tokens. It is worth mentioning that other exploratory studies have also been carried out for the subcorpora of other cities in the Hispanic world.

## 3. Results

Monophthongisation is much more common in Andalusia than in other varieties of Spanish; this realisation is only found on the Iberian Peninsula and in some Caribbean locations. Old written references and the very negative judgement of Spanish linguists suggest that the form

[mu] is strongly stigmatised in prestigious varieties of Spanish: the Spanish spoken in central and northern Spain. However, this stigmatisation does not seem to apply in Andalusia. While social factors such as age and level of education appear to align with the pattern of stigmatisation, other factors suggest that the monophthong form [mu] carries significant local prestige. The fact that it is the most common variant (58.5%) seems to confirm this. It is also important to note that social groups with a high level of education sponsor it. Women prefer the monophthong form and this appears to contradict the usual assumptions about women's preference for prestigious variants. It is also striking that, in interviews, women use far more adverbs muy (regardless of their phonetic realisation) than men. This fact is related to studies which highlight the tendency of women in certain Spanish-speaking communities to use more intensifiers in their discourse. This also helps us introduce the novel and interesting pragmatic factor that the form [mu] is becoming specialised with the value of intensification, as opposed to the form [muj], which maintains a more neutral character.

## 4. References

Alarcos Llorach, Emilio (1971). Fonología española. Madrid: Gredos.

Albelda, Marta, Cestero, Ana María, Guerrero, Silvana & Samper, Marta (2020). Variación sociopragmática y geolectal en el uso de la atenuación. Lengua y Habla, 24, 1-53.

Moliner, María (1986). Diccionario de uso del español. Madrid: Gredos.

Quilis, Antonio (1997). Principios de fonología y fonética españolas. Madrid: Arco/Libros.

Real Academia Española y Asociación de Academias de la Lengua Española (2011). Fonética y fonología. Madrid: Espasa-Calpe.

# Retroflexion and Resistance: A Sociophonetic Study of Identity, Register, and Accent Ideology in Malayalam-English Bilingual Speech

Anupama Thekkadavan, Ashima Saseendran
*The English and Foreign Languages University, Hyderabad*
anupamaphdlandp21@efluniversity.ac.in, ashima431998@gmail.com

## Abstract

This sociophonetics study investigates the role of retroflex consonants (/ʈ/, /ɖ/, /ɭ/) in the spoken English of Malayalam-English bilinguals in Kerala, India. Far from being dismissed as mere phonological interference from the first language, retroflexion is examined here as a socially meaningful and ideologically charged feature of bilingual speech. The research asks not only how these sounds are used, but also what they reveal about speakers' linguistic identity, gender dynamics, and shifting perceptions of "correct" English. Adopting a mixed-methods approach, the study analyses data from 40 participants drawn from both urban and semi-urban settings. These participants took part in a range of communicative tasks, including classroom discussions, semi-structured interviews, role-play performances, and matched-guise perception tests. Through acoustic phonetic analysis, the study finds that retroflexion occurs most frequently in informal registers and tends to be downplayed or replaced in more formal contexts. This variation suggests that speakers are aware of sociolinguistic norms and often adjust their pronunciation based on setting and audience. Interestingly, patterns of retroflexion also reflect gendered strategies. Many male speakers retain retroflex sounds during casual speech, whereas several female participants exhibit more pronounced shifts in register, pointing to a stronger orientation toward social expectations or linguistic stigma. These findings are supported by qualitative responses, where speakers express an implicit awareness of retroflexion as a nonstandard trait. Some even describe actively avoiding these sounds in formal or academic settings, a behaviour described in this study as phonetic self-erasure, the deliberate suppression of native phonetic features to align with perceived norms of prestige English. This research argues that retroflexion is not simply a residue of the first language but a symbolically potent marker of regional affiliation, cultural positioning, and speaker agency. Within India's multilingual and postcolonial context, these findings underscore the need to rethink rigid "standard" pronunciation models. Rather than policing variation, pronunciation pedagogy should aim to embrace linguistic diversity and acknowledge the social realities of bilingual speakers. Ultimately, this study demonstrates how small phonetic features carry big social meanings and how language becomes a site where ideology, identity, and expression intersect.

**Keywords:** sociophonetics, retroflexion, bilingual, phonetic self-erasure, identity

## 1. Introduction and background

This study investigates the sociophonetic dynamics of retroflexion in English speech among Malayalam-English bilingual speakers in Kerala, India. Retroflex consonants such as /ʈ/, /ɖ/, and /ɭ/, common in Dravidian languages, are often perceived in English loanwords as phonological interference. However, this research reframes retroflexion as a socially meaningful phenomenon that interacts with speakers' identities, gender, social registers, and ideologies about accent and prestige. By moving beyond purely phonetic descriptions, the study aims to understand how bilingual speakers modulate or resist retroflex articulations as part of their sociolinguistic strategies in diverse communicative contexts. This contributes to broader discussions on multilingual speech practices and accent ideology in postcolonial India.

## 2. Methodology

Data were gathered from 40 Malayalam-English bilingual speakers, sampled from urban and semi-urban settings in Kerala. The research employed a mixed-method approach combining quantitative acoustic analysis and qualitative sociolinguistic techniques. Participants participated in sociolinguistic interviews, role-play tasks designed to elicit varying speech registers, classroom recordings, and matched-guise experiments to simulate formal versus informal interactions. Acoustic measurements were conducted using Praat software to identify and quantify retroflex articulations in key lexical items. Qualitative data included speaker reflections on language attitudes, identity, and perceived stigma surrounding their English accents.

## 3. Results

Analysis revealed consistent use of retroflex consonants in casual, peer-oriented speech, with examples such as "school" pronounced as /skuːɭ/ and "data" as /ɖeːʈʈaː/. Conversely, formal speech contexts showed marked reduction or suppression of retroflexion, with speakers favouring alveolar pronunciations, reflecting accommodation to perceived standard English norms. Gender differences were evident: male speakers tended to maintain retroflex sounds more persistently in informal settings, whereas female speakers exhibited greater flexibility, often suppressing retroflexion to align with standard Indian English expectations. Qualitative findings highlighted linguistic anxiety and awareness of accent-related stigma, especially in academic and professional domains.

## 4. Discussion and conclusion

The study introduces the concept of "phonetic self-erasure" to describe deliberate suppression of native phonetic traits due to sociolinguistic pressures. Overall, these findings challenge simplistic views of retroflexion as mere interference, instead positioning it as a socially loaded resource for expressing local identity, resistance, and aspirations toward prestige. The research calls for a critical reassessment of pronunciation norms in Indian English education, advocating for recognition and validation of regionally rooted phonetic repertoires.

## 5. References

Annamalai, E. (2001). Managing multilingualism in India: Political and linguistic manifestations. Sage Publications.

Bamgbose, A. (1998). Tensions in Nigerian English. World Englishes, 17(3), 273–287. https://doi.org/10.1111/1467-971X.00099

Bansal, R. K. & Harrison, P. J. (1973). A textbook of phonetics for Indian students. Orient Longman.

Chakrabarti, P. (2015). Sociolinguistic perspectives on Indian English accents: Negotiating identity and ideology. Journal of Sociolinguistics, 19(2), 177–199. https://doi.org/10.1111/josl.12117

Foulkes, P. & Docherty, G. (2006). The social life of phonetics and phonology. Journal of Phonetics, 34(4), 409–438. https://doi.org/10.1016/j.wocn.2005.08.002

Gumperz, J. J. (1982). Discourse strategies. Cambridge University Press.

Kachru, B. B. (1992). The other tongue: English across cultures (2nd ed.). University of Illinois Press.

Kang, O. (2010). Sociophonetics: A brief introduction. In N. Hornberger & S. McKay (Eds.), Sociolinguistics and language education (pp. 111–135). John Wiley & Sons.

Krishnamurti, B. (2003). The Dravidian languages. Cambridge University Press.

Ladefoged, P. & Maddieson, I. (1996). The sounds of the world's languages. Blackwell.

Mehrotra, C. M. (1998). Indian English: Its evolution and current usage. Macmillan India.

Mohanan, K. P. (1994). The theory of lexical phonology. Springer. https://doi.org/10.1007/978-94-011-1142-4

Nair, S. (2009). Gender and language use in Kerala: An ethnographic perspective. South Asian Linguistics, 13(1), 45–68.

Sankaran, S. (2014). Language ideology and linguistic identity in South India: The case of Malayalam English bilinguals. International Journal of Bilingualism, 18(3), 278–295. https://doi.org/10.1177/1367006912441416

Sharma, D. (2005). Indian English and its dialects: Phonetic and phonological perspectives. In R. B. Singh (Ed.), English in India: New directions (pp. 135–155). Macmillan.

Wells, J. C. (1982). Accents of English (Vol. 1). Cambridge University Press.

Yadav, R. R. (2013). Accent variation and language attitudes among Indian English speakers: A sociophonetic approach. World Englishes, 32(4), 565–581. https://doi.org/10.1111/weng.12056

# Creaky voice across apparent time in German: Insights from the Plapper corpus

Joshua Penney[1], Melanie Weirich[2], Stefanie Jannedy[3]

[1]*Macquarie University,* [2]*Friedrich Schiller University,* [3]*Leibniz-Centre General Linguistics*

joshua.penney@mq.edu.au, melanie.weirich@uni-jena.de, jannedy@leibniz-zas.de

**Abstract**

In this study, we conducted an apparent time study on the prevalence of creaky voice in read speech produced by Germans between the ages of 18-75. Using an automatic creak identification method, we found that female speakers produced more creak than males, with the effect mitigated by age: more creak was found in younger compared to older females. This suggests that increased use of creaky voice by young female speakers is an ongoing change in German.

**Keywords:** creaky voice, voice quality, sociophonetics, sound change, apparent time

## 1. Introduction and background

Creaky voice is a non-modal voice quality characterised by a combination of low, irregular fundamental frequency (F0), glottal constriction, and damping of individual glottal pulses. There are several different acoustic manifestations of creaky voice (Keating et al. 2015), but a common feature for most types of creaky voice is low F0. In recent years, there have been suggestions that the use of (non-contrastive) creaky voice is on the rise, especially among young, female speakers in English speaking countries, and creak has been linked to a number of social meanings (Pearce 2023; Podesva 2013; Yuasa 2010). Differences in creak between speaker groups in languages other than English have been less extensively studied, with most studies focusing on creaky voice's prosodic boundary marking function (e.g. Carlson et al. 2005; Ogden 2001; cf. van Hugte & Heeren 2024). Very little research has examined creaky voice or voice quality more generally in German (Köser 2014; Peters 2003). However, recent work identified differences in voice quality between multiethnolectal Kiezdeutsch and Standard German speakers from Berlin, suggesting that voice quality may play a role in indexing social identity in German (Penney et al. 2024). In this study, we present an apparent time analysis examining the prevalence of creaky voice in German speakers across a wide range of ages, to investigate whether there is evidence for creaky voice becoming more common in younger speakers, and whether there are gender-based differences in the occurrence of creak.

## 2. Methodology

### 2.1. Data

The data for this study are taken from the Plapper corpus. Plapper is a smartphone application developed to crowdsource speech data from German speakers from across Germany (Jannedy 2022; Jannedy & Weirich 2023; Weirich et al. 2024). Participants took part in spontaneous and read speech tasks as their voices were recorded through their phone, including a consistent set of read sentences, which are examined here.

### 2.2. Speakers

Data was initially included for 250 speakers. This represented all available German-born speakers aged 18-75 who had produced at least 3 of the 7 read sentences, and for whom segment boundaries had been hand checked/corrected. Only data from speakers identifying as male or female were included due to low numbers of speakers of other genders. 5 speakers were excluded as outliers (see below) resulting in data from 245 speakers (F: 162; M: 83).

**2.3 Analysis**

Instances of creaky voice were automatically identified using the Antimode method (Doreen 2017; Dallaston & Docherty 2019). F0 was estimated for all sonorant segments using REAPER (Talkin 2015), which provides reliable pitch tracking at low frequencies such as those in which creak commonly occurs. F0 distributions are bimodal when speakers produce creak; the method identifies as creaky all speech that falls below an antimode between a speaker's two F0 modes (see Dallaston & Docherty 2019 and White et al. 2022 for details).

For each speaker, we first calculated an overall creak prevalence score, which is the percentage of creak occurring across all of a speaker's available speech. Eight outliers with prevalence scores above 20% were manually checked and 5 of these were excluded. We then fitted a generalised linear mixed effects model to analyse effects of gender and age and their interaction on the binary dependent variable of whether creak was present or not. Random intercepts were included for speaker and for sentence.

**3. Results**

Figure 1 illustrates estimated creak prevalence according to gender and age as lines, and raw individual speaker creak prevalence scores as data points. There was a significant interaction between gender and age (p = 0.02). From Figure 1 it can be seen that, considered across all ages, females showed higher creak prevalence scores than males (raw mean scores: F: 6.68%; M: 5.51%). From the figure, it can also be seen that the effect of higher creak prevalence in the females is not consistent across age: creak was more likely in younger female speakers, and less likely in older female speakers; for males, there was little change across age.
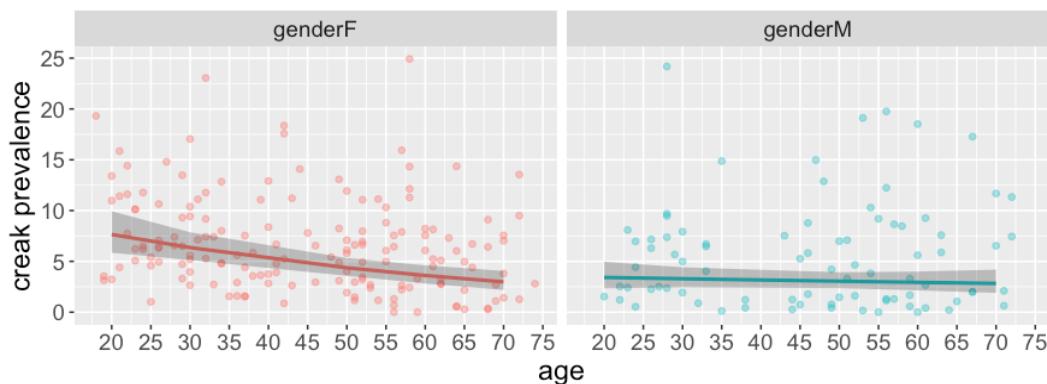


Figure 1. Estimated creaky voice prevalence according to age and gender. Individual data points represent raw speaker prevalence scores.

**4. Discussion and conclusion**

Although there are some speakers of both genders who produce high rates of creak, on the whole, the results suggest that female speakers produce creaky voice more than males in German. Moreover, the interaction between gender and age demonstrates that for females the degree of creak increases as age decreases, with younger females producing more creaky voice than older females. This result across apparent time suggests that increased use of creaky voice may be an ongoing change that is increasingly being taken up by young, female speakers, similar to what has been claimed to be the case in English (Podesva 2013; Yuhasa 2010). We note that this analysis is preliminary and utilises an automatic identification method; in future work we plan to integrate complementary analysis methods and extend this analysis to a larger set of data. Nonetheless, the results present a compelling case for a gendered increase in creaky voice, which, as far as we are aware, is the first report of creak functioning in this way as a sociolinguistic marker in German.

**5. References**

Carlson, R., Hirschberg, J. & Swerts, M. (2005). Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates. Speech Communication, 46, 326–333.

Dallaston, K. & Docherty, G. (2019). Estimating the prevalence of creaky voice: A fundamental frequency-based approach. In S. Calhoun, P. Escudero & P. Warren (eds.), Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia (532–536).

Dorreen, K. (2017). Fundamental frequency distributions of bilingual speakers in forensic speaker comparison. University of Canterbury.

Jannedy, S. (2022). Plappern für die Wissenschaft – Eine App gegen die Datenlücke. Bericht über das Forschungsjahr 2020/2021: ZAS; Impressionen, 115–126.

Jannedy, S. & Weirich, M. (2023). Plapper - a smartphone app for selfrecording: Exploration of vowel spaces. 5th Phonetics & Phonology in Europe (PaPE 2023) (157–158).

Keating, P., Garellek, M. & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.

Köser, S. (2014). Phrasen-finale Phonationsänderungen und ihre Rolle beim turn taking. In D. Barth-Weingarten & B. Szczepek Reed (eds.), Prosodie und Phonetik in der Interaktion (20–45). Verlag für Gesprächsforschung.

Ogden, R. (2001). Turn transition, creak and glottal stop in Finnish talk in interaction. Journal of the International Phonetic Association, 31, 139–152.

Pearce, J. (2023). Creaky voice in three Scots varieties: Using F0 based identification to consider social and linguistic factors. In R. Skarnitzl & J. Volín (eds.), Proceedings of the 20th International Congress of Phonetic Sciences, Prague, Czechia (1816–1820).

Penney, J., Weirich, M. & Jannedy, S. (2024). Increased Breathiness in Adolescent Kiezdeutsch Speakers: A Marker of Multiethnolectal Group Affiliation? Language and Speech. https://doi.org/10.1177/00238309241269059.

Peters, B. (2003). Multiple cues for phonetic phrase boundaries in German spontaneous speech. In M. J. Solé, D. Recasens & J. Romero (eds.), Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain (1795–1798).

Podesva, R. (2013). Gender and the social meaning of non-modal phonation types. Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society, 427–448.

Talkin, D. (2015). REAPER: Robust epoch and pitch EstimatoR. https://github.com/google/REAPER. Accessed on 21/05/2025.

van Hugte, T. B. R & Heeren, W. F. L. (2024). Exploring Interspeaker Variation in Creaky Voice in Dutch. Journal of Voice. https://doi.org/10.1016/j.jvoice.2024.05.011.

Weirich, M., Duran, D. & Jannedy, S. (2024). Gender and age based f0-variation in the German Plapper Corpus. In I. Lapidot & S. Gannot (eds.), Proceedings of Interspeech, Kos, Greece (1565–1569).

White, H., Penney, J., Gibson, A., Szakay, A. & Cox, F. (2022). Evaluating automatic creaky voice detection methods. Journal of the Acoustical Society of America, 152, 1476–1486.

Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women. American Speech, 85, 315–337.

# Tracing Sociolinguistic Change among Women in Málaga: Indexical Meaning and the Andalusian Phonetic System

Matilde Vida Castro, María Cruces Rodríguez
*Universidad de Málaga*
mvida@uma.es, mariacruces@uma.es

## Abstract

This study examines how linguistic change among women in Málaga (Spain) is shaped by broader social transformations affecting their roles in contemporary society. It explores the idea that women use certain linguistic features as indexical tools to align themselves with particular social identities. The analysis centres on three key phonological variables of Andalusian Spanish: the realisation of the coronal fricatives /s/ and /θ/ in syllable onset, the velar /x/ in medial position, and the alveolar /s/ in final word position. Data were drawn from 48 women representing five generational cohorts, with speakers distributed across four oral corpora collected at distinct synchronic points: Arqueoling (1960–1980), PRESEEA-Málaga I (1995), PRESEEA-Málaga II (2015), and MAYHO (2024). These corpora provide a rich temporal perspective on variation, with the latest corpus specifically designed to capture emerging patterns associated with new models of womanhood in Andalusian society. The main hypothesis posits that younger, educated women will exhibit a decline in the use of traditional Andalusian variants, favouring features more closely aligned with the standard variety of Peninsular Spanish. This linguistic shift is interpreted as a reflection of evolving social identities characterised by increased empowerment and independence. Preliminary findings support this hypothesis, revealing a near-complete abandonment of several traditional features among younger women—not only those traditionally stigmatised, such as the merger of /s/ and /θ/ or the weakening of /x/, but also others more widely accepted, including the elision of postnuclear /s/. Furthermore, linguistic and social factors, such as interactional register, syllable stress, and neighbourhood of residence, appear to influence usage patterns. These results suggest that phonological variation in Málaga is not merely a matter of regional identity but also a dynamic expression of shifting gender roles and social aspirations.

**Keywords:** Andalusian Spanish, phonetics, gender, indexicality, sociolinguistic variation

## 1. Introduction and background

This paper explores the relationship between linguistic change and social transformation among women in the city of Málaga, Spain. It argues that the phonological shifts observed in recent decades are closely tied to evolving gender roles and broader social dynamics. Drawing on the concept of indexicality, the study considers how specific linguistic features are used agentively by speakers to signal alignment with particular social identities (Eckert 1989:265, 2012; Labov 2001: 385-411, Villena & Vida 2020: 167).

The analysis focuses on three salient phonological variables in Andalusian Spanish: the realisation of the coronal fricatives /s/ and /θ/ in syllable onset, the velar /x/ in medial position, and the alveolar /s/ in final word position (Narbona et al. 2003). These features serve as key indicators of linguistic variation across generations of female speakers in Málaga and provide insight into the socio-indexical meanings embedded in phonetic choices.

## 2. Hypothesis and research questions

The central hypothesis of this study is that certain groups of women—particularly younger, educated speakers represented in the most recent corpus—will exhibit a lower frequency of traditional Andalusian phonological variants. Instead, they are expected to favour features more closely associated with the standard variety of Peninsular Spanish. This linguistic shift is

interpreted as a strategy for projecting a redefined social identity marked by greater autonomy, empowerment, and alignment with contemporary models of womanhood.

In addition, the study seeks to determine whether the linguistic changes observed in Andalusian Spanish began to emerge during the final decades of the twentieth century among the youngest generations or whether educated women born in the early twentieth century were already using more prestigious variants, at least in formal contexts. To explore this question, the analysis will quantify the occurrence of each phonological realisation among women in the Arqueoling corpus, which includes recordings made before the 1980s.

## 3. Methodology

This study draws on data from 48 female speakers spanning five generational cohorts, distributed across four oral speech corpora compiled at different synchronic moments. Each corpus includes women from more than one generation, allowing for both cross-sectional and diachronic insights into linguistic variation: Arqueoling (1960–1980), PRESEEA-Málaga I (1995), PRESEEA-Málaga II (2015), and MAYHO (2024). The Arqueoling corpus includes historical audio recordings from public institutions and private contributors (Ávila et al. 2024). The other three corpora were compiled according to the methodological guidelines of the PRESEEA project (Moreno-Fernández 1993, 1996). The most recent corpus, MAYHO, was designed to focus on capturing the sociolinguistic dimensions of the so-called "new woman", with attention to emerging identity patterns and linguistic change.

More than 1,000 instances of the three phonological variables under study were analysed following a perceptual judgement on their phonetic realisation. The data were coded in Microsoft Excel and subjected to a descriptive analysis based on the calculation of both absolute and relative frequencies (percentages). Chi-square tests of independence were conducted to explore possible associations between sociolinguistic variables and the different realisations of each phoneme. Cramér's V was also used to measure effect size, following the interpretation criteria proposed by Levshina (2015: 209).

All statistical analyses and data visualisations were carried out using the R programming language (R Core Team 2025) within the RStudio development environment.

## 4. Results

Preliminary findings indicate that younger women have largely abandoned traditional Andalusian phonological features. This shift is evident not only in the rejection of strongly stigmatised variants—such as the merger of /s/ and /θ/ and the weakening of the velar /x/—but also in the decline of features previously considered more socially acceptable and widespread, like the elision of postnuclear /s/. These results suggest a significant realignment of linguistic practices among younger generations. Additionally, the use of different variants appears to be conditioned by a range of linguistic and social factors, including the formality of the interactional register, the speaker's residential neighbourhood, and prosodic elements such as syllable stress. These factors contribute to a complex sociolinguistic landscape in which language variation reflects individual identity and broader social dynamics.

## 5. Discussion and conclusion

The preliminary results indicate that the reduced traditional Andalusian features among younger women may reflect evolving social identities. Nonetheless, further detailed analysis is required to substantiate these trends and to elucidate the underlying social factors driving this linguistic variation.

## 6. References

Ávila-Muñoz, Antonio, Molina-García, Álvaro, Villena-Ponsoda, Juan Andrés & Sionakidis, Georgios (2024). Las voces de Málaga. Reconstrucción histórica y archivo sonoro de la Ciudad Lingüística. Granada: Comares.

Eckert, Penelope (1989). The Whole Woman: Sex and Gender Differences in Variation. Language Variation and Change 1: 245–267.

Eckert, Penelope (2012). Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. Annual Review of Anthropology 41: 87–100.

Labov, William (2001). Principles of Linguistic Change. Volume 2: Social Factors. Oxford: Blackwell.

Moreno-Fernández, Francisco (1993). Proyecto para el estudio sociolingüístico del español de España y América (PRESEEA). Lingüística 5: 268–271.

Moreno-Fernández, Francisco (1996). Metodología del 'Proyecto para el estudio Sociolingüístico del español de España y América' (PRESEEA). Lingüística 8: 257–287.

Narbona-Jiménez, Antonio, Cano-Aguilar, Rafael & Morillo-Velarde, Ramón (2003). El español hablado en Andalucía. Sevilla: Fundación José Manuel Lara.

R Core Team. (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/. Accessed on 09/06/2025

Villena-Ponsoda, Juan Andrés & Vida-Castro, Matilde (2020). Variation, identity and indexicality in southern Spanish. On the emergence of a new variety in urban Andalusia. En M. Cerruti & S. Tsiplakou (eds.), Intermediate language varieties. Koinai and regional standards in Europe. 150-183. Amsterdam: John Benjamins.

# Sociophonetics beyond the word: building connected speech processes into accounts of variation and change

Gerard Docherty, Paul Foulkes
*Griffith University, University of York*
gerry.docherty@griffith.edu.au, paul.foulkes@york.ac.uk

## Abstract

This paper addresses the paradox that while connected speech processes are ubiquitous in unscripted speech, they figure only marginally in sociophonetic studies of phonological variation and change. We present findings from an exploratory analysis of variation across six lexical collocations found frequently within a corpus of spoken conversations produced by 40 speakers of Australian English. Implications of the findings are discussed for methodological and theoretical approaches to investigating phonological variation and change.

**Keywords:** connected speech, sociophonetics, Australian English, phonological variation

The phonetic realisation of words in connected speech is routinely different from that of words produced in isolation (Ernestus (2012); Warner (2021)). The label "connected speech processes" (CSPs) captures such differences, referring to reduction or elision of segments or even entire syllables, assimilations of place and manner of articulation and of voicing, and simplifications of segment sequences. CSPs appear commonly at word margins, where the immediate context is strongest, but can also impact on segmental properties throughout a word. CSPs vary both across and within languages (e.g. Ingram (1989), Nolan and Kerswill (1990)), are more or less likely to occur across different speech styles (e.g. Warner and Tucker (2011)), and can be an expression of different types of pragmatic meaning over and above the denotation of the lexical item concerned (e.g. Hawkins' (2004)). CSPs therefore constitute phonetic resources available to speakers in determining how they sound in particular communicative contexts.

In this light, it is instructive to consider the body of work on social dimensions of phonological variation and change (PVC); see Kendall and Fridland (2021) and Strelluf (2024) for recent overviews. Based to a very large extent on corpora of natural unscripted speech, in which CSPs are found in abundance, work on PVC is largely focused around the properties of individual lexical items. The central object of study, the phonological variable, is typically defined in relation to the canonical form of lexical items, and this approach is reinforced by the fact that the automatic segmentation, forced alignment and interrogation of large corpora are underpinned by large lexical repositories such as CELEX or CMUdict. This lexical focus is clear in vowel studies in which investigators search a corpus for all lexical items containing an accented token of a target vowel category and treat the identified tokens as an equivalence set for the purposes of acoustic analysis of that variable. The immediate segmental or prosodic context beyond target items may be introduced as a predictor in modelling variation in the realisation of a particular phonological category, but there are relatively few PVC studies focusing on variable phenomena that arise directly from the juxtaposition of target words within connected speech; examples include studies of V-V hiatus (Penney et al. (2024)), and of the realisation of the nucleus of "the" (Cox et al. (2023).

This paper presents an exploratory investigation of the sociophonetic properties of a subset of CSPs in a corpus of unscripted English. We focus on six lexical collocations that in connected speech styles give rise to variability in phonetic realisation that would not typically be factored in to analyses of the phonological variables contained within the relevant lexical items. Our research questions are: (a) what is the nature and range of phonetic variability observed in the realisation of these collocations as evidenced by a corpus of Australian English

conversations, and what factors can be adduced to account for variation that is observed; (b) to what extent do we find cross-speaker variation in the way in which the focused CSPs are realised; (c) what are the implications of our findings for analysis of the vowel categories associated with the lexical items investigated?

Our study focuses on the performance of 40 young speakers from Perth WA engaged in unscripted conversational dyads lasting around 30 minutes. Speakers are balanced equally across male/female and neighbourhood of residence (as a proxy for 'class').

Our dataset comprises contexts where variation in phonetic realisation arises from the collocation of two lexical items: going to (n=205), kind of (122), sort of (119), want to (108), don't know (277), and used to (61). The grammatical item in each pair would typically be excluded from a variationist analysis by virtue of 'stop list' filtering, but the other item would be included in searches for words containing the relevant phonological variable. In all of the collocations investigated, the content word is amongst the most frequent items matching the relevant vowel category, and the collocation in focus is amongst the most frequently encountered frames for the corresponding content word (e.g. across the entire corpus used is one of the most frequent exponents of the goose lexical set, and all of the tokens of used in this corpus are found in used to sequences).

The exploratory nature of this study lends itself to a phonetic profiling approach (Docherty and Foulkes (2005)) which does not pre-define the parameters for analysis, but rather allows these to emerge from a combined auditory-spectrographic analysis which characterises the full range of CSPs observed in the data. Our quantitative analysis explores the diversity encountered within the phonetic realisations observed (see examples in Fig 1), and the extent of any relationship between CSPs and local speech rate, as well as the extent of variation across individual speakers or groups of speakers defined by sex and neighbourhood.

We present quantitative data relating to all of the CSP contexts listed above, and discuss the findings in relation to previous accounts of cross-speaker variation in CSPs in Australian and other varieties of English. We also consider the implications of our findings for conventional methodologies for sociophonetic research and for theoretical approaches to accounting for variation and change in the realisation of phonological categories.
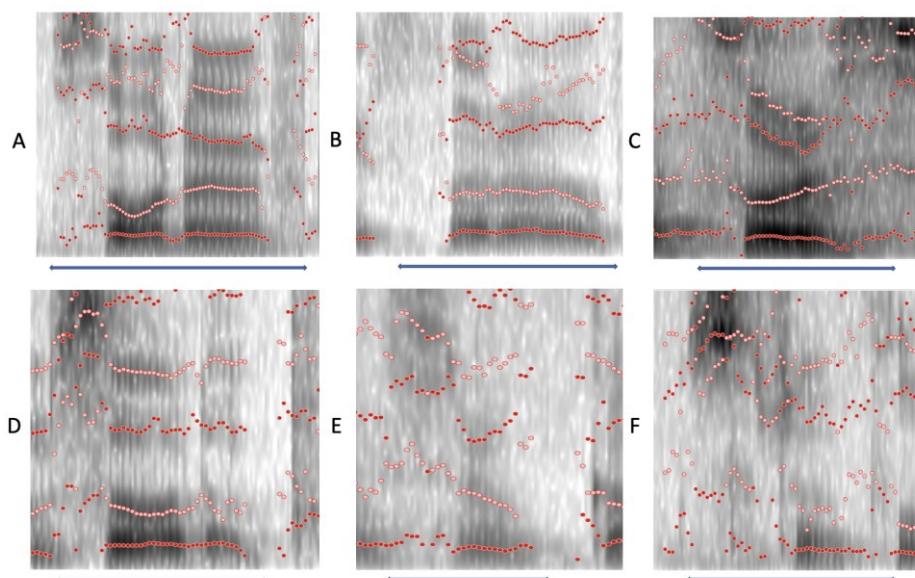


Figure 1: Spectrograms (x = t, y = 0-5 Khz) exemplifying diverse phonetic realisations of sort of within the dataset (horizontal lines map approximately to sort of - formant estimates are shown as an aid to readers). A: vocalic nuclei clearly demarcated (of longer than sort) with a clear tap realisation for /t/; B: as with A, but with a weaker /t/ realisation; C: /t/ realised as an approximant [ɹ]; D: clearly audible realisation of /t/ as [ʔ]; E: single syllable realisation with continuous vowel glide; F: vowel of sort is devoiced between the obstruent Cs.

**References**

Cox, F., Penney, J. & Palethorpe, S. (2023). Fifty years of change to prevocalic definite article allomorphy in Australian English. International Phonetic Association. Journal of the International Phonetic Association, 53(3), 804-834. https://doi.org/https://doi.org/10.1017/S002510032200007X

Docherty, G. & Foulkes, P. (2005). Glottal variants of /t/ in the Tyneside variety of English. In W. J. Hardcastle & J. M. Beck (Eds.), A figure of speech A festschrift for John Laver (pp. 173-197). Lawrence Erlbaum Associates.

Ernestus, M. (2012). Message-related variation: segmental within-speaker variation. In A. C. Cohn, C. Fougeron & M. Huffman (Eds.), The Oxford Handbook of Laboratory Phonology (pp. 99-102). Oxford University Press.

Hawkins, S. (2004). Roles and representations of systematic fine phonetic detail in speech understanding (vol 31, pg 373, 2003). Journal of Phonetics, 32(2), 289-289. https://doi.org/10.1016/j.wocn.2004.01.001

Ingram, J. C. L. (1989). Connected speech processes in Australian English. Australian Journal of Linguistics, 9(1), 21-49. https://doi.org/10.1080/07268608908599410

Kendall, T. & Fridland, V. (2021). Sociophonetics. Cambridge University Press.

Nolan, F. & Kerswill, P. E. (1990). The description of connected speech processes. In Studies in the Pronunciation of English: A Commemorative Volume in Honour of A.C. Gimson (pp. 295-316). Taylor and Francis. https://doi.org/10.4324/9781315688367

Penney, J. Cox, F. & Gibson, A. (2024). Hiatus resolution and linguistic diversity in Australian English. Phonetica, 81(2), 119-152. https://doi.org/10.1515/phon-2023-0029

Strelluf, C. (Ed.). (2024). The Routledge Handbook of Soiophonetics. Routledge.

Warner, N. (2021). Processes in Connected Speech. In R.-A. Knight & J. Setter (Eds.), The Cambridge Handbook of Phonetics (pp. 133-156). Cambridge University Press. https://doi.org/DOI: 10.1017/9781108644198.006

Warner, N. & Tucker, B. V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. The Journal of the Acoustical Society of America, 130(3), 1606-1617. https://doi.org/10.1121/1.3621306

# The Social Significance is in a State of Constant Change: Language Attitudes and the New Interpretation of the [ts] Allophone in Malaga Andalusian Spanish

Pablo Pérez Molina
*Universidad de Málaga & Università di Torino*
pabloperezm@uma.es

**Abstract**

The primary aim of this proposal is to report findings derived from a linguistic attitudes experiment regarding the social significance that the speech community of Málaga associates with two specific variants: the voiceless aspirated [tʰ] and the alveolar affricate [ts], both derived from the /-st-/ sequence (e.g., fiesta). Previous research comparing two oral speech corpora collected 20 years apart (PRESEEA-Málaga I and II) indicated a significant shift in the distribution of the affricate: while in 1995 it was predominantly found among young women with higher education, by 2015 its presence had increased significantly among young women without formal education. This suggests a change in its indexical properties over time. To examine the current social meaning of this feature, a matched-guise survey was distributed to 120 participants. Judges were exposed to acoustically modified stimuli representing both variants. The results suggest that the affricate is not currently a highly salient marker for the community as a whole. However, age-graded differences indicate that older speakers associate the affricate with higher status but less local authenticity, whereas younger speakers show no significant differentiation. Current evidence suggests a mechanical shift is taking place, primarily influenced by the standardisation of the Andalusian koine.

**Keywords:** Linguistic attitudes, Andalusian Spanish, Indexicality, Social meaning, Phonetic variation

## 1. Introduction and background

In the last decades, the speech community of Málaga has undergone significant sociolinguistic changes, partly due to the standardisation of the regional "intermediate variety" or Andalusian koine. Within this context, the realization of the cluster /-st-/ as an affricate [ts] has shown a complex evolution. While the weakening of /s/ in coda is a staple of southern Peninsular Spanish, the affricated realization is a relatively recent phenomenon. Diachronic comparisons between the PRESEEA-Málaga I (1995) and PRESEEA-Málaga II (2015) corpora reveal a striking pattern: the feature has shifted from being associated with young, educated women to being more prevalent among young women with lower educational levels. This trajectory challenges the initial hypothesis of a prestige-driven "change from above" and suggests a process of enregisterment or functional reinterpretation. It is essential to examine the indexical properties of these sounds to understand if the community perceives this variant as a marker of local identity, prestige, or a mechanical articulatory consequence.

## 2. Methodology

To measure the social meaning of the affricate, a matched-guise experiment was designed following the tradition of Lambert et al. (1960). The stimuli were created from recordings of four native speakers from Málaga: a male and a female university student (22 years old), and a male (63) and female (53) with vocational training.

### 2.1. Stimuli preparation

Speakers were recorded in a semi-directed interview and a picture description task to elicit spontaneous productions of words containing /-st-/ (e.g., *estatua, fiesta*). A carrier phrase of 2-5 seconds containing at least two tokens was selected for each speaker. Using Praat software,

the original /-st-/ segments were manipulated to create two guises for each speaker: one with the post-aspirated realization [tʰ] and another with the affricated realization [ts]. The acoustic parameters (duration, intensity, center of gravity) were adjusted to match the reference values for Málaga Spanish. This resulted in eight target stimuli. Additionally, four distractor phrases were selected from a reading task, ensuring they contained no marked Andalusian features or controversial content.

## 2.2. Data collection and analysis

Two versions of the questionnaire were created to ensure no participant heard the same speaker producing both variants. The survey was distributed digitally to 120 participants (89 women, 31 men), primarily under 35 years old, divided into locals (63) and non-locals (57). Participants rated the speakers on a 6-point semantic differential scale across 13 socio-affective traits (e.g., intelligent, formal, local accent). A Principal Component Analysis (PCA) and Factor Analysis were performed to group these traits, followed by mixed-effects linear regression models (using R) to assess the influence of linguistic and social factors, treating subject and speaker as random effects

## 3. Results

The PCA revealed three main dimensions of evaluation: Status (grouping education, culture, and socioeconomic level), Accent (grouping perceived Andalusian/Málaga accent), and Sympathy (grouping confidence and pleasantness). Regarding Status, the overall effect of the variant was not statistically significant. However, a significant interaction with age was found: older participants rated the affricated [ts] guise as having higher status than the post-aspirated guise, whereas younger participants did not differentiate between them. The speaker's voice identity had a stronger effect on status perception than the phonetic variant itself. Regarding Accent, the affricated variant was generally associated with a stronger local accent, though the result was not statistically significant overall. However, older listeners perceived the [ts] variant as less traditionally Andalusian than the post-aspirated one. Non-local participants tended to rate all stimuli as having a stronger regional accent compared to local judges. Finally, regarding Sympathy, no significant differences were found between the two allophones, suggesting that the affricated variant does not trigger distinct affective reactions compared to the post-aspirated one.

## 4. Discussion and conclusion

The results suggest that the affricated realization of /-st-/ is not a highly salient stereotypic marker for the general population of Málaga, as indicated by the lack of drastic differences in evaluation between the guises. This supports the hypothesis that the spread of [ts] is a "mechanical change" driven by articulatory economy (preference for open syllables) rather than a prestige-driven change. However, the age-graded differences are crucial. Older speakers associate [ts] with higher status and less local accent, likely interpreting this novel, clearer articulation as a feature of the standard or "correct" speech, contrasting it with the traditional aspiration. Younger speakers, who are the primary users of the intermediate variety, show no such differentiation, indicating that the feature has become normalized or "unmarked" in their speech. This lack of strong indexicality among the youth confirms that the variant is fully integrated into the leveling process of the Andalusian koine.

## 5. References
Eckert, P. (2008). Variation and the indexical field. Journal of Sociolinguistics, 12(4), 453-476.

Moya Corral, J. A. (2007). Noticia de un sonido emergente: la Africada Dental procedente del grupo -ST- en Andalucía. *Revista de filología de la Universidad de La Laguna*, 25, 457-466.

Pérez Molina, P. (2023). Indexicalidad y significado social de un alófono emergente: la variante africada procedente de /-st/ en la variedad andaluza malagueña. Málaga: Repositorio Institucional de la Universidad de Málaga.

Pérez Molina, P. & Vida Castro, M. (2025). Cambio fonético y significado social en la ciudad de Málaga. Estudio actitudinal sobre el proceso de aspiración de la oclusiva dental sorda procedente del segmento /-st-/. In A. M. Ávila Muñoz (coord.), *Regreso a la ciudad lingüística* (127-164). Valencia: Tirant Humanidades.

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

Vida Castro, M. (2022). On competing indexicalities in southern peninsular Spanish. A Sociophonetic and Perceptual analysis of Affricate [ts] through time. *Language Variation and Change*, 34(2), 137-163.

Villena Ponsoda, J. A. (2008). Sociolinguistic patterns of Andalusian Spanish. *International Journal of the Sociology of Language*, 193-194, 139-160.

Villena Ponsoda, J. A., Vida Castro, M. & Molina García, Á. (2022). Coherence in a levelled variety. The case of Andalusian. In K. Beaman & G. Guy (eds.), *The Coherence of Linguistic Communities* (239-257). New York: Routledge. https://doi.org/10.4324/9781003134558-18

# Pronunciation of syllable- and word-final /s/ in Venezuelan Spanish: regional and sociodemographic patterns of variation

Verena Weiland
*University of Bonn*
vweiland@uni-bonn.de

**Abstract**

This study investigates variation in syllable- and word-final /s/ pronunciation in four Venezuelan locations: two urban sites (Caracas in the lowlands, Mérida in the highlands) and two rural sites (San Antonio de los Altos in the lowlands, La Azulita in the highlands). Using the TiAlBa corpus, recordings from speakers balanced by age, gender, and education were analysed. Mixed-effects models and conditional inference trees evaluated the influence of linguistic and extralinguistic factors on /s/ realization. Results show that phonological context and elicitation method significantly affect /s/ pronunciation, with alveolar [s] mainly occurring before pauses in word lists and texts. Contrary to expectations, neither highland/lowland nor urban/rural factors showed significant overall effects. Notably, elision is significantly more frequent in Mérida during interviews and texts (as opposed to the wordlist) in specific phonological contexts, indicating local variation.

**Keywords:** /s/-pronunciation, variation, corpus phonology, Venezuela, highlands/lowlands

## 1. Introduction and background

Aspiration and elision of syllable- and word-final /s/ are well-known features of consonant pronunciation in lowland American Spanish (Henríquez Ureña 1921, Hualde 2014). /s/-weakening exhibits considerable variation, influenced by both internal factors – such as phonological context, word length, lexical frequency, and prosodic stress – and external factors, including speech register and sociodemographic variables such as age, gender, and education (e.g., Brown 2008, Medina Rivera 2011, Carrasquero 2010).

This study examines /s/ pronunciation in four Venezuelan locations – urban (Caracas, Mérida) and rural (San Antonio de los Altos, La Azulita) – across two regions (lowlands and highlands). It aims to advance our understanding of /s/ variation by integrating geographic, settlement, sociodemographic, stylistic, and linguistic factors. Venezuela presents a notable case with both highland and lowland regions, yet its Spanish varieties do not consistently follow typical geographic patterns. In Caracas, the prestige variety primarily shows aspiration (74%), with less frequent retention (16%) and deletion (10%) of /s/ (PRESEEA corpus 2004-2010; Carrasquero 2010), consistent with previous findings (Villamizar 1998). In Mérida (Venezuelan Andes), high rates of aspiration and deletion occur preconsonantally and prepausally (Longmire 1976, Chela-Flores & D'Aquino 2003).

## 2. Methodology

### 2.1 Corpus data

This study is situated within the framework of corpus phonology. The data come from the Tierras Altas y Bajas de Hispanoamérica (TiAlBa) corpus. Currently, this corpus comprises speech recordings from over 20 data collection points across 11 Hispanic American countries. The recordings were collected in 2023 and 2024, following the Fonología del Español Contemporáneo (FEC) protocol, which ensures a standardized methodology for data collection for linguistic purposes (Pustka et al. 2016, Pustka et al. 2018). The aim of the FEC protocol is to comprehensively capture the phonetic-phonological inventory of the Spanish language. The protocol uses a three-part methodology: a list of 125 mostly randomized words (including potential minimal pairs), a text containing some of these words to observe continuous speech,

and a semi-structured interview. The wordlist and the text provide the protocol with a distinctive feature, as corpora such as PRESEEA or Val.Es.Co., for instance, primarily consist of spontaneous speech. The semi-structured interview balances structure and spontaneity by guiding speakers toward specific topics while encouraging them to take the lead as "experts." This design fosters a conversational dynamic that closely approximates spontaneous interaction (Labov 1972: 36-37; Brown 2008: 38-39). At each site, twelve lifelong residents were recorded, evenly balanced by gender, age group, and education level.

## 2.2 Encoding

For the coding of /s/, nine pronunciation categories are distinguished based on the place of articulation, as shown in Table 1, by visual inspection of the sonagram in Praat (Boersma & Weenink 2021). The Kappa value (Fleiss et al. 2003) for this coding system is 0.91, indicating "almost perfect agreement" (Landis & Koch 1977, R package DescTools; cf. Andri et al. 2025).

| 1 | Alveolar fricative [s] or [z] |
|---|---|
| 2 | Glottal fricative [h] (aspiration) |
| 3 | Dental fricative [θ] or [ð] |
| 4 | Retroflex fricative [ʂ] or [ʐ] |
| 5 | Palatal fricative [ç] o [j] |
| 6 | Velar fricative [x] o [ɣ] |
| 8 | Elision [ø] |

Table 2: Numeral coding system for the realization of the Spanish consonant /s/

## 3. Results

The statistical analysis was conducted using R, employing mixed-effects linear models as well as a conditional inference tree (cTree), shown in Figure 1. A cTree is a predictive model that recursively partitions data by selecting predictors most strongly associated with the response variable. It is suitable for exploratory analyses, as it can reveal higher-order interactions, handle sparse data, and model categorical responses without distributional assumptions (see Tagliamonte & Baayen 2012, Levshina 2015). The dependent variable is pronunciation of /s/. The independent variables include the right and left phonological context, part-of-speech, phonological stress (stressed/unstressed), protocol type (wordlist, text, interview), landform (highland/lowland), settlement type (urban/rural), age, gender and education.

The results highlight that phonological context and protocol type are key factors influencing /s/ pronunciation variation. The alveolar fricative [s] occurs predominantly before pauses in the word list and text protocols, and in other phonological contexts mainly in the word list. Neither the highland/lowland nor the urban/rural factors show significant effects. The factor city of origin only matters in a very specific context: when in the interview and text protocols the right phonological context is neither a pause nor a plosive, pronunciation in Mérida differs from other cities (see nodes 6 and 7). In Mérida, elision is significantly more frequent in these contexts, with very few occurrences of alveolar /s/.
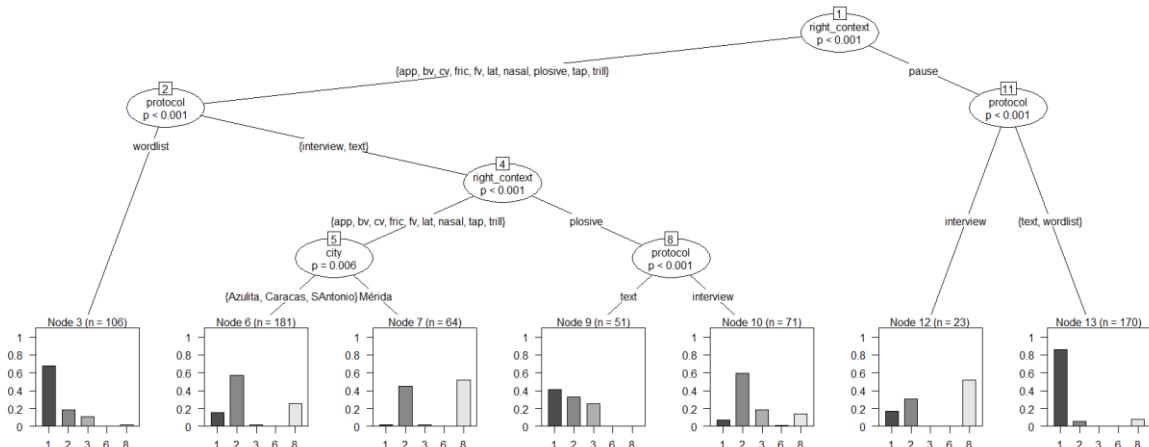
Fig. 1: Conditional inference tree predicting syllable- and word-final /s/ realization by the variables Left and right phonological context, protocol and city.

## 4. References

Andri, Signorell et al. (2025). DescTools: Tools for descriptive statistics. https://cran.r-project.org/web/packages/DescTools/DescTools.pdf. Accessed on 31/05/2025.

Boersma, Paul & Weenink, David (2021). Praat: doing phonetics by computer 6.4.34. http://www.praat.org/. Accessed on 31/05/2025.

Brown, Earl. K. (2008). A usage-based account of syllable- and word-final /s/ reduction in four dialects of Spanish. Albuquerque, New Mexico: Dissertation.

Carrasquero, Viktor (2010). Un caso de variación sociofonética: /-s/ posnuclear en el español actual de Caracas. Letras, 52(81), 63-86.

Chela-Flores, Godsuno & D'Aquino, Giovanna (2003). Reorientación de la división dialectal hispánica: Datos de variedades venezolanas e ibéricas. Opción, 41, 95-111.

Henríquez Ureña, Pedro (1921). Observaciones sobre el español de América. Revista de Filología Española, 8, 357-390.

Hualde, José Ignacio (2014). Los sonidos del español. Cambridge: Cambridge University Press.

Labov, William (1972). Sociolinguistic patterns. Oxford: Blackwell.

Landis, J. Richard & Koch, Gary G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1), 159-174.

Levshina, Natalia (2015). How to do linguistics with R: Data exploration and statistical analysis. Amsterdam: John Benjamins.

Longmire, Beverly Jean (1976). The relationship of variables in Venezuelan Spanish to historical sound changes in Latin and the Romance languages. Georgetown: University Press.

Medina-Rivera, Antonio (2011). Variationist Approaches: External Factors Conditioning Variation in Spanish Phonology. In Díaz-Campos, Manuel (ed.), The handbook of hispanic sociolinguistics. Malden: Wiley-Blackwell, 36-53.

Pustka, Elissa, Gabriel, Christoph & Meisenburg, Trudel (2016). Romance Corpus Phonology: from (Inter-)Phonologie du Français Contemporain (I)PFC to (Inter-) Fonología del Español Contemporáneo (I)FEC. In Draxler, Christoph & Kleber, Felicitas (eds.), Tagungsband der 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum. P und P12. 12.-14. Oktober 2016. Munich: Ludwig-Maximilians-Universität. https://epub.ub.uni-muenchen.de/29405/. Accessed on 31/05/2025.

Pustka, Elissa, Gabriel, Christoph, Meisenburg, Trudel, Burkard, Monja & Dziallas, Kristina (2018). (Inter-)Fonología del Español Contemporáneo/(I)FEC: metodología de un programa de investigación para la fonología de corpus. Loquens, 5(1).

Tagliamonte, Sali A. & Baayen, Harald (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. Language Variation and Change, 24(02), 135-178.

Villamizar, Thania (1998). Fonetismo. In Obediente Sosa, Enrique (ed.), El habla rural de la Cordillera de Mérida. Mérida: Universidad de Los Andes, Consejo de Desarrollo Científico, Humanístico y Tecnológico, Centro de Investigaciones Etnológicas, 27-94.

# Sociophonetic Variation and Adaptation in Malayalam: Language Contact between Malayalam, Hindi, and Bengali in Kerala

Anupama Thekkadavan
*The English and Foreign Languages University, Hyderabad*
anupamaphdlandp21@efluniversity.ac.in

## Abstract

Kerala has recently experienced a significant influx of migrant labourers from North India and Bangladesh, particularly from North India's Hindi-speaking belt and Bangladesh's Bengali-speaking regions. This demographic shift has transformed the state's linguistic landscape, creating dynamic multilingual contact zones in urban and semi-urban areas. Malayalam speakers increasingly engage in everyday communicative interactions with Hindi and Bengali speakers in these contexts. These sustained interactions not only foster lexical borrowing but also lead to phonological accommodation, where the phonological systems of Hindi and Bengali subtly but systematically influence the phonetic features of Malayalam. This study investigates sociophonetic variation in Malayalam speech within this multilingual environment, focusing on how language contact shapes phonological adaptation. This research identifies and examines specific change features using sociolinguistic interviews, participant observation, and acoustic phonetic analysis. For example, younger Malayalam speakers are observed replacing the aspirated plosive /ph/ with the fricative /f/, reflecting a feature commonly found in Hindi and Bengali, or substituting the labiodental /ʋ/ with the bilabial /b/ sound, as /ʋ/ is absent in the phonemic inventory of Bengali. Consequently, all /ʋ/ sounds are naturally replaced with /b/ in Bengali. Additionally, modifications in intonation patterns, retroflexion, and vowel length are examined as indicators of interactional alignment and identity negotiation. The study adopts an interdisciplinary framework integrating sociophonetics, language contact theory, and identity theory. It argues that phonological accommodation in Malayalam is not merely a linguistic consequence of contact but a socially embedded practice reflecting and mediating evolving relationships of power, solidarity, and cultural convergence. These adaptations often serve as strategies for inclusion, politeness, or social mobility in a rapidly transforming sociolinguistic environment. The findings illustrate that phonetic variation is closely tied to speakers' perceptions of social hierarchy, group membership, and linguistic prestige. This research contributes to a broader understanding of how multilingualism shapes language use at the phonetic level, offering new insights into the fluid nature of linguistic identity in contact zones. The study has implications for language planning, education, and social integration in linguistically diverse societies. It calls for further attention to the phonetic consequences of migration and multilingual interaction, especially in contexts undergoing rapid social and demographic change.

**Keywords:** Sociophonetics, language contact, Malayalam, phonological accommodation, identity

## 1. Introduction

Kerala, the southernmost state of India, is considered the most literate state. Thus, the people of Kerala have migrated to various other states in India and abroad to pursue better education and job prospects. Hence, no Keralites are available to do manual jobs in Kerala. In this scenario, there has been an increased labour migration from northern India and Bangladesh.

The north Indian Hindi-speaking community and the Bangladeshi Bengal-speaking community have to communicate mainly with the older and middle generations of the Malayalam-speaking community in Kerala. Though these non-Malayalam speakers have started to learn and speak Malayalam, the Malayalam-speaking local Keralites have also started to learn some Hindi and Bengali to facilitate communication. Kerala has thus evolved into a multilingual landscape, reshaping local communicative practices.

This study addresses this gap by focusing on sociophonetic variation in Malayalam shaped by cross-linguistic interaction. Sociophonetics bridges phonetic analysis and sociolinguistic inquiry and offers a robust framework to study how sound patterns encode social meaning, identity, and group dynamics. By examining not only what changes occur in speech but also why and in what social contexts, this research contributes to a nuanced understanding of linguistic adaptation in contact zones.

## 2. Discussion

The analysis reveals those phonetic shifts in Malayalam, such as:

1. Adopting the fricative /f/ for the aspirated plosive /pʰ/. E.g.: /phalam/ ☐ /falam/ (Fruit).

2. Substituting the labiodental approximant /ʋ/ with the bilabial plosive /b/. E.g.: /ʋellam/ ☐ /bellam/ (Water).

3. Loss of retroflexion: /na:ɳayam/ ☐ /na:nayam/ (Coin), /tuɭasi/ ☐ /tulsi/ (Tulasi, a medicinal plant).

4. Rhythm and Intonation Change: Malayalam is syllable timed with relatively flat intonation contours. Hindi and Bengali are more stress-timed, and pitch rises and falls are more pronounced, especially in yes-no questions. Malayalam speakers have started using Hindi-style rising intonation at sentence ends.

5. /s/ vs /ʃ/ distinctions collapsing or shifting: In Malayalam, there is a clear distinction between /s/ and /ʃ/. Hindi/Bengali may influence Malayalam speakers to confuse or merge these sounds. E.g.: /ʃabdam/ ☐ /sabdam/ (Sound)

These changes are emerging through naturalistic, recurrent interactions between speakers of different languages. These are not random changes but socially motivated adaptations reflecting perceived prestige, linguistic accommodation, and evolving norms of politeness and belonging.

## 3. Conclusion

Phonological accommodation in Malayalam emerges as a linguistic and social process shaped by migration, multilingualism, and interaction dynamics. Far from being passive outcomes, these phonetic variations represent active, adaptive strategies in a complex sociolinguistic ecology. The study underscores the value of sociophonetics in revealing how language contact reshapes speech patterns, social identities, and cultural affiliations, with broader implications for policy, education, and integration in multilingual societies.

## 4. References

Auer, P. (2007). Style and social identities: Alternative approaches to linguistic heterogeneity. Mouton de Gruyter.

Foulkes, P. & Docherty, G. J. (2006). The social life of phonetics and phonology. Journal of Phonetics, 34(4), 409–438. https://doi.org/10.1016/j.wocn.2005.08.002

Hock, H. H. & Joseph, B. D. (2009). Language history, language change, and language relationship: An introduction to historical and comparative linguistics (2nd ed.). Mouton de Gruyter.

Labov, W. (2001). Principles of linguistic change: Social factors (Vol. 2). Blackwell.

Sharma, D. (2005). Dialect stabilisation and speaker awareness in non-native varieties of English. Journal of Sociolinguistics, 9(2), 194–224. https://doi.org/10.1111/j.1360-6441.2005.00292.x

Thomason, S. G. & Kaufman, T. (1988). Language contact, creolisation, and genetic linguistics. University of California Press.

Trudgill, P. (1986). Dialects in contact. Blackwell.

# Real-time variation in creaky voice prevalence in Australian English

Hannah White, Joshua Penney, Felicity Cox
*Macquarie University*
felicity.cox@mq.edu.au

**Abstract**
There is a general perception that the prevalence of creaky voice has been increasing in recent years; however, few studies are yet to explore whether this is empirically the case. In the present study, we compare creak prevalence in the speech of teenagers from Sydney recorded in 1989 to that of teenagers recorded in the early 2020s. Results suggest that, at least for female speakers, there has been a significant increase in creaky voice prevalence in Australian English.

**Keywords:** creaky voice, Australian English, real-time variation

## 1. Introduction and background

Creaky voice is a voice quality generally characterised by low and/or irregular fundamental frequency (f0) (Keating et al. 2015). There is a general association between creak and young women's speech, as well as a perception that these young women are driving an increase in the prevalence of creak (e.g., Swannell, 2023).

Studies of British English (BrE) over the years have found more creak in the speech of men compared to women (e.g., Henton & Bladon 1988; Gittelson et al. 2021). In the US, however, studies have either found no difference in creak prevalence by gender (Abdelli-Beruh 2016) or a higher prevalence in women's speech (e.g., Yuasa, 2010).

Although limited, a few studies have investigated a change in creak prevalence over time, most using apparent-time methodologies. In BrE, Gittelson et al. (2021) found an increase in acoustic correlates of creak with increasing age for men but not for women. Oliveira et al. (2016) looked at just American women's speech finding no significant difference in creak prevalence between younger speakers and older speakers. In Australian English (AusE), Penney et al. (2020) found more evidence of creaky phonation among younger speakers compared to older speakers. A recent real-time study of Finnish provides support for a generational increase in creak prevalence among women but not men (Uusitalo et al. 2024).

Despite an increase in research on creak in AusE in recent years (e.g., Loakes & Gregory 2022; Penney et al. 2020; White et al. 2023), to our knowledge a real-time analysis of creak prevalence is yet to be conducted for this variety. This is the focus of the present study.

## 2. Methodology
### 2.1. Data

Data in the present study are from two corpora, both collected from the Northern Beaches area of Sydney. Contemporary data come from the Multicultural Australian English – Voices of Sydney (MAE-VoiS) corpus collected between 2020 and 2023 (Cox & Penney 2024). Conversations from 22 speakers (14F; 8M) were included in the contemporary dataset. Historical data were extracted from Cox (1996), collect in 1989. Sociolinguistic interviews from 38 speakers (16F; 22M) were included in this dataset. All speakers from both datasets were aged between 14 and 16 at the time of recording.

### 2.2. Data preparation and analysis

Creak was automatically detected using the Union method (White et al. 2022). This method combines an algorithm that relies on low f0 to identify creak (Dallaston and Docherty 2019; Dorreen 2017) and one that relies on other acoustic cues to creak such as low H1-H2 (Drugman et al. 2014). White et al. (2022) showed that combining these algorithms can improve creak

detection compared to using either method alone. After manually checking the output of the Union method, 10 speakers were excluded from the final analysis due to poor performance, resulting in 49 speakers (historical: 15F & 13M; contemporary: 14F & 7M).

A generalised linear mixed-effects regression model was run to investigate change in creak prevalence over time in AusE. The dependent variable was whether creak was identified or not in any given 10 ms interval. Independent variables included timepoint (historical vs contemporary) and gender (female vs male). Creak has been shown to mark phase-finality (e.g., Henton & Bladon 1988; White et al. 2023) so whether creak was in a pre-pausal (PP) word or not was also included, as was a 3-way interaction between these variables. Speaker and the phoneme during which the creak occurred were included as random intercepts, both with random slopes for PP position.

## 3. Results

Results showed all main effects and 2-way interactions were significant. Significant 2-way interactions are presented in Figure 1.
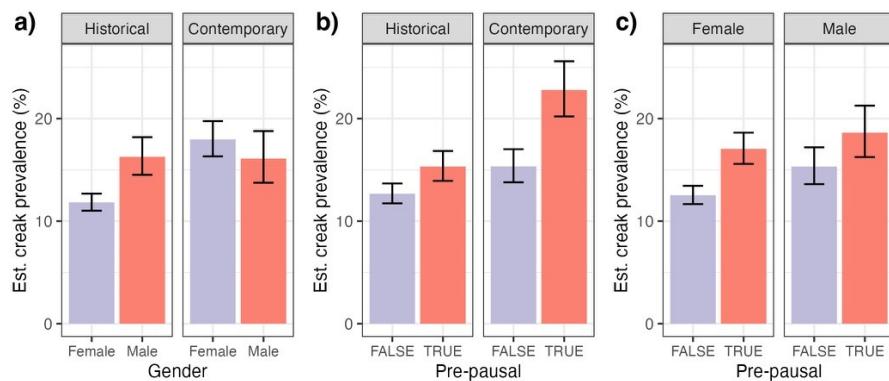


Figure 1. Estimated creak prevalence by a) timepoint and gender; b) timepoint and pre-pausal position; c) gender and pre-pausal position.

Post-hoc pairwise comparisons showed a significant increase in creak over time for females but not males (Figure 1(a)). It appears that while historically males were creakier than females, there is no gender difference in the contemporary data. Turning to Figure 1(b), we found creak is more likely in PP words at both timepoints (although this difference is greater for contemporary speakers). Over time creak has increased in both positions. In Figure 1(c), we see that regardless of gender, creak is more likely in PP position. No gender difference was found in PP position, but in non-PP position males are overall creakier than females.

## 4. Discussion and conclusion

The findings of the present study provide evidence for an increase in creak prevalence in AusE for females but not for males (as in Uusitalo et al. 2024). While historically, males were creakier than females, the contemporary data does not indicate any support for creak currently being a predominantly female feature in AusE. We acknowledge, however, that the number of contemporary males in our dataset is small, so further investigation is required.

The finding of more creak in PP condition in both datasets and for both genders supports the idea that creak is a marker of phrase-finality (e.g., Henton & Bladon 1988; White et al. 2023). It is unknown why, across gender, this effect is greater for contemporary speakers than historical speakers, but we speculate that it may involve changing intonation patterns. Finally, across timepoints males were creakier than females when creak was not in PP position, whereas in PP position there was no gender difference. This suggests that, while being phrase-final may

not preclude creak indexing social meaning, at least for these speakers, creak is not indexing gender in this position.

## 6. References

Abdelli-Beruh, N. B., Drugman, T. & Red Owl, R. H. (2016). Occurrence frequencies of acoustic patterns of vocal fry in American English speakers. Journal of Voice, 30(6), 759.e711–759.e720.

Cox, F. (1996). Vowel variation in Australian English [Unpublished doctoral dissertation]. Macquarie University.

Cox, F. & Penney, J. (2024). Multicultural Australian English – The New Voice of Sydney. Australian Journal of Linguistics, 44(2–3), 200–219.

Dallaston, K. & Docherty, G. (2019). Estimating the prevalence of creaky voice: A fundamental frequency-based approach. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds.), Proc. 19th ICPhS (581.1–5). Melbourne, Australia.

Dorreen, K. (2017). Fundamental frequency distributions of bilingual speakers in forensic speaker comparison [Unpublished master's thesis]. University of Canterbury.

Drugman, T., Kane, J. & Gobl, C. (2014). Data-driven detection and analysis of the patterns of creaky voice. Computer Speech and Language, 28(5): 1233–1253.

Gittelson, B., Leemann, A. & Tomaschek, F. (2021). Using Crowd-Sourced Speech Data to Study Socially Constrained Variation in Nonmodal Phonation. Frontiers in Artificial Intelligence, 3, 565682.

Henton, C. & Bladon, A. (1988). Creak as a sociophonetic marker. In L.M. Hyman & C. N. Li (eds.), Language, speech, and mind: Studies in honour of Victoria A. Fromkin (3–29). London; New York: Routledge.

Keating, P., Garellek, M. & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. In The Scottish Consortium for ICPhS 2015 (ed.), Proc. 18th ICPhS (821.1–5). Glasgow: University of Glasgow.

Loakes, D. & Gregory, A. (2022). Voice quality in Australian English. JASA Express Letters, 2(8), 085201.

Oliveira, G., Davidson, A., Holczer, R., Kaplan, S. & Paretzky, A. (2016). A comparison of the use of glottal fry in the spontaneous speech of young and middle-aged American women. Journal of Voice, 30(6), 684–687.

Penney, J., Cox, F. & Szakay, A. (2020). Glottalisation, coda voicing, and phrase position in Australian English. Journal of the Acoustical Society of America, 148(5), 3232–3245.

Swannell, C. (2023). Please stop frying your vocals, ladies. The Medical Republic. https://www.medicalrepublic.com.au/please-stop-frying-your-vocals-ladies/18024. Accessed on 29/05/2025.

Uusitalo, T., Nyberg, L., Laukkanen, A., Waaramaa, T. & Rantala, L. (2024). Has the prevalence of creaky voice increased among Finnish university students from the 1990's to the 2010's? Journal of Voice, 38(3):697–702.

White, H., Penney, J., Gibson, A., Szakay, A. & Cox, F. (2022). Evaluating automatic creaky voice detection methods. Journal of the Acoustical Society of America, 152(3), 1476–1486.

White, H., Penney, J., Gibson, A., Szakay, A. & Cox, F. (2023). Creak prevalence and prosodic context in Australian English. In Proc. Interspeech, 2023 (112–116). Dublin, Ireland.

Yuasa, I. P. (2010). Creaky Voice: A New Feminine Voice Quality for Young Urban Oriented Upwardly Mobile American Women? American Speech, 85(3), 315–337.
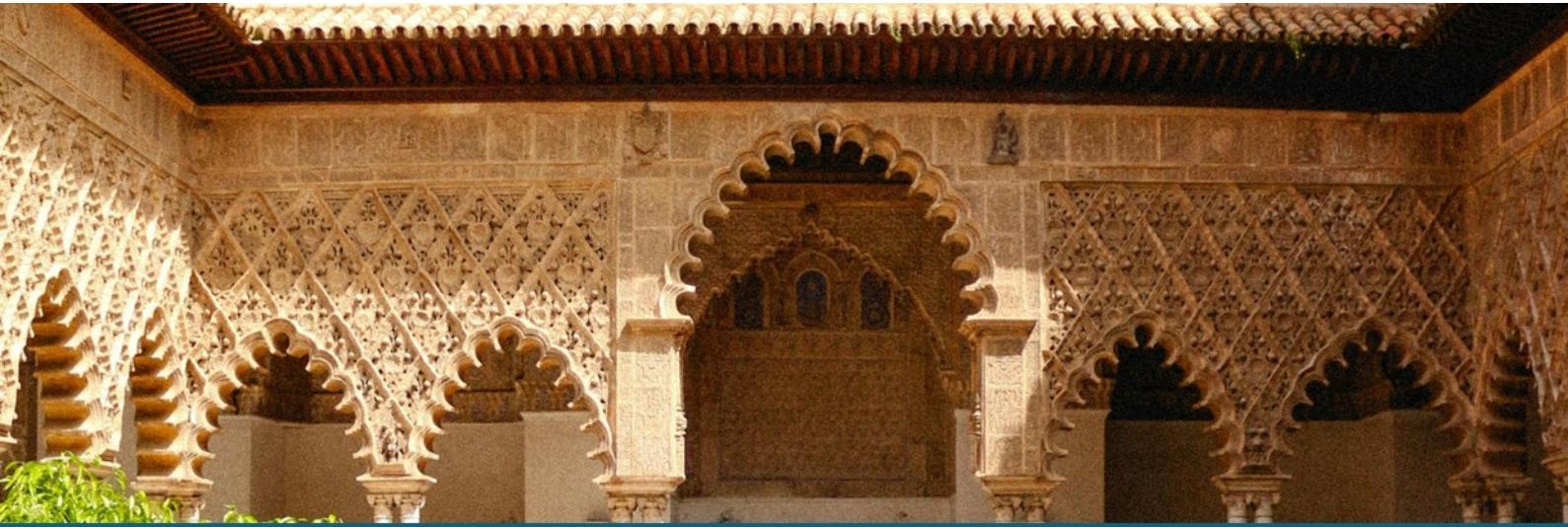
# PROSODIC VARIATION

# Prosodic variation as a function of planning scope: Evidence from individual differences

Jason Bishop, Mei-Ying Ki
*City University of New York*
jbishop@gc.cuny.edu, mki@gradcenter.cuny.edu

## Abstract

Recent research in phonetics and psycholinguistics has explored the extent to which speech production planning might be flexible—i.e., the idea that the scope of speakers' (pre)planning of an upcoming utterance is sensitive to both external (e.g. speaking conditions) and internal (e.g. cognitive limitations) factors. In this paper, we will discuss findings from a large-scale production study with American English speakers that bears on the role of a speaker-internal factor, namely working memory capacity (WMC). We show that prosodic variation related to silent pause durations is predicted by speakers' WMC in ways that suggest a planning mechanism. In particular, speakers with lower WMC tend to produce shorter pre-phrasal silent pause durations when the prosody allows for it. We argue these patterns reflect divergent planning strategies—more incremental planning for low-WMC speakers and less incremental planning for high-WMC speakers—at least at early stages of phonological planning.

**Keywords:** prosody, speech production planning, individual differences, silent pause

## 1. Introduction and background

Findings related to a wide range of phonetic and phonological patterns suggest that speech production planning, at least the earliest stages of it, unfolds in relatively large chunks—chunks that are better defined in terms of phrase-level prosodic units than in terms of one-or-two-word sequences (cf. Levelt et al., 1999). At the same time, there is considerable evidence that the scope of speakers' planning, rather than relying on a single "unit of planning," is instead flexibly adaptive to the limitations of the speaker. In the present study, we explore the idea, considered extensively in Krivokapic (2012), that partial planning of upcoming prosodic structure (which we assume to reflect early phonological planning; Keating and Shattuck-Hufnagel, 2002) can be inferred from pre-phrasal silent pause durations. It has been known for some time that silent pause durations are longer before longer utterances than before shorter utterances (Whalen & Kinsella-Shaw, 1997), which has been understood to reflect the extra time required to begin planning longer stretches of speech material. However, Krivokapic (2007) demonstrated that these pause durations are not only affected by the length of the upcoming utterance, but also its prosodic structure. In particular, Krivokapic found that pause durations were longer before utterances with a simple prosodic structure relative to those before complex prosodic structures, where complexity is defined as a branching structure in the prosodic hierarchy. Krivokapic argues her findings reflect the different options speakers have for planning the upcoming utterance; complex phrases allow for a more incremental planning strategy, whereby speakers can plan the first lower-level phrase during the silent pause, and then the second phrase while producing the first; simple prosodic structures allow for only one strategy: initiating planning of the whole higher-level phrase during the silent pause. The contrast is show, in Fig. 1, where the phrase levels referred to are based on those in Beckman and Pierrehumbert (1986).

## 2. Present study: Individual differences in planning strategies?

Based on the above, we assume that, other things being equal, silent pauses should be longer when speakers have a larger chunk of upcoming speech material to plan. However, if prosodic phrase structure can be used to break up this larger chunk into smaller, phrasal chunks with

less material to plan, less planning time will be required during the pause, resulting in shorter pause duration.
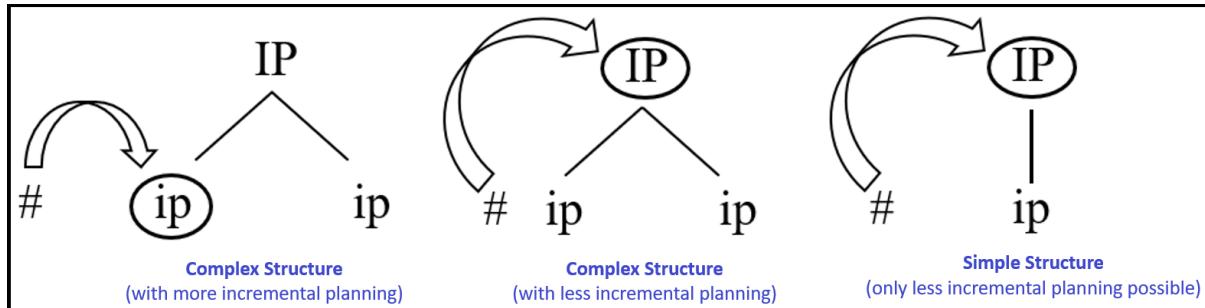


Fig 1. Different prosodic structures offer different planning strategies: IP=intonational phrases, ip=intermediate phrases; '#'=silent pause durations; and arrows show the target of planning during the silent pause. Crucially, complex phrases allow for more incremental planning, predicted to be preferred by speakers with lower WMC.

The question we ask here is whether speakers with lower WMC, who have been argued to favor smaller planning units (Wagner et al., 2010), show a preference for the more incremental approach. If they do, we predict that before complex prosodic phrases, speakers with lower WMC will produce shorter pause durations, since their strategy involves planning lower-level/shorter phrases; speakers, with higher WMC, since their strategy involves planning higher-level/longer phrases, will produce longer pause durations. Finally, before simple prosodic phrases, we do not predict any systematic difference in pause duration in relation to WMC, since simple prosodic phrases do not allow for different strategies.

## 3. Experiment
### 3.1. Methods
The present study's basic goal was to relate pause durations to the length and especially structure of an upcoming intonational phrase (IP). To that end, samples of connected speech from 100 speakers of American English who read a 165-word passage, originally reported in Bishop & Intlekofer (2020), were reanalyzed for the present study, along with 44 new speakers. Speech from all samples was annotated for phrase structure using a modified version of the ToBI guidelines (Beckman & Ayers, 1997) that prioritized break indices in decisions about phrasing. In addition to the speech production task, speakers also completed an automated version (Oswald et al., 2015) of the Reading Span task (Daneman & Carpenter, 1980), producing an estimate of each speaker's verbal WMC. Mixed-effects linear regression was then used to predict pause durations based on upcoming phrase length (in syllables), phrase structure (simple vs. complex), and speakers' WMC (reading spans). For the present purposes, the crucial test was for an interaction between phrase structure and speakers' WMC.

### 3.2. Results and discussion
A significant interaction, apparent in Fig 2., indicated that speakers with higher WMC produced longer pauses before complex prosodic structures but not before simple ones (B=7.1,SE=3.5,t=2.03, p>.05), supporting the idea that speakers engage in more or less incremental planning strategies depending on their own WMC profiles. To the extent that these kinds of patterns do reflect such strategies, this suggests that individual differences in planning scope may reflect a significant source
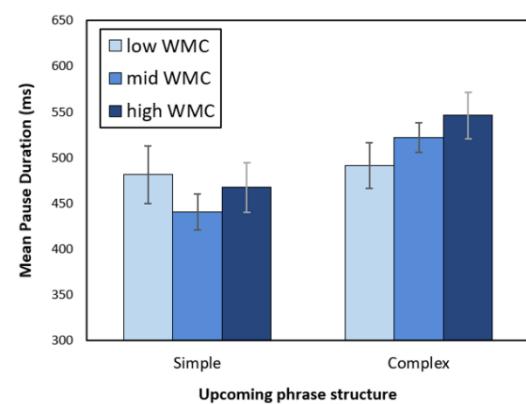


**Fig 2.** Silent pause duration as a function of upcoming phrase structure and speakers' WMC: Higher WMC means longer pause durations before complex, but not simple phrases.

of phonetic variation not typically considered. We explore some additional prosodic patterns related to phrase length and fundamental frequency that may further support this interpretation.

## 5. References

Beckmann, M. & Ayers, G. (1997). Guidelines for ToBI labeling (Version 3). Unpublished ms. Ohio State University.

Beckman, M. & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. Phonology Yearbook, 3, 255–309.

Bishop, J. & Intlekofer, D. (2020). Lower working memory capacity is associated with shorter prosodic phrases: Implications for speech production planning. Proceedings of Speech Prosody 10, 191-195.

Daneman, M. & Carpenter, P. (1980). Individual differences in working memory and reading. Journal of Verbal Learning & Verbal Behavior, 19, 450–466.

Keating, P. & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. UCLA Working Papers in Phonetics, 101, 112-156.

Krivokapic, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. Journal of Phonetics, 35, 162-179.

Krivokapic, J. (2012). Prosodic planning in speech production. In Susanne Fuchs, Melanie Weihrich, Daniel Pape, Pascal Perrier (eds.) Speech planning and dynamics. Peter Lang, pp. 157-190.

Levelt, W.J., Roelofs, A. & Meyer, A. (1999). A theory of lexical access in speech production. Brain and Behavioral Sciences, 22, 1-38, 1999.

Oswald, F. O., McAbee, S. T., Redick, T. S. & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. Behav Res, 47, 1343–1355.

Wagner, V., Jescheniak, J. & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36, 423−440.

Whalen, D. H., and Kinsella-Shaw, J. M. (1997). Exploring the relationship of breath intake to utterance duration. Phonetica, 54, 138-152.

# Exploring Phonetic Variation in Italian Poetic Speech: Intonation, Prosodic Features and Voice Quality

Federico Lo Iacono[1,3] & Valentina Colonna[2,3]

[1]*Univ. of Turin and Univ. of Genoa;* [2]*Univ. of Granada;* [3]*Laboratorio di Fonetica Sperimentale "Arturo Genre"*

federico.loiacono@unito.it, valentinacolonna@ugr.es

## Abstract

This study examines both recurring patterns and phonetic variation in Italian poetic speech, focusing on prosody, intonation, and voice quality. Based on 12 performances from the Voices of Italian Poets archive, the analysis aims to investigates the phonetic and prosodic variation of poetic speech in structure and delivery. Annotation was performed in Praat using an extended VIP protocol, adding tiers for phonetic segments, syllables, and intonational functions. Intonation contours were studied alongside pause distribution and selected acoustic parameters—such as H1–H2, spectral emphasis, HNR, duration, intensity, and pitch dynamics—extracted from key expressive and structural passages. Findings reveal both consistent patterns—like the poetic declarative, a gently descending contour with explanatory or emphatic function—and significant variation in intonation, voice quality, pausing, and rhythm. A common tendency toward softened voice quality also emerges, linked to poetic expressivity and specific acoustic traits. Overall, poetry reading is shown to be a highly expressive and variable phonetic register, where stable prosodic forms coexist with notable individual and stylistic diversity.

**Keywords:** poetic speech, intonation, voice quality, phonetic variation, Voices of Italian Poets

## 1. Introduction and background

Recently, experimental phonetics has increasingly focused on poetry reading [1, 2, 3], inspiring new analytical models and large-scale projects such as the Voices of Italian Poets (VIP) [4, 5, 6] and the Voices of Spanish Poets (VSP) [7, 8]. Despite significant progress in data collection and analysis, key aspects of poetic speech remain understudied—particularly intonation, which plays a central role in structuring poetic enunciation and pause management. Its relevance has long been acknowledged by linguists and literary critics for its impact on composition, performance, and versification rules [9, 10]. However, experimental research on poetic intonation remains scarce. A foundational account for English was proposed in [11] and later developed in [12]. In the Italian context, [4] identified distinctive features such as the poetic declarative, marked by a medium-low, gently descending contour. More recent studies [13, 14] have extended this analysis to other contour types—interrogative, continuative, enumerative— and have argued for the inclusion of voice quality in models of poetic intonation [15].

## 2. Hypothesis and selected corpus

This study aims to explore both the recurring patterns and the variation in prosodic and acoustic features of Italian poetic speech, starting from the hypothesis that poetry reading exhibits distinctive acoustic and linguistic traits—especially in its management of pauses, intonation, and voice quality. We analysed 12 readings by 12 (3 F, 9 M) contemporary Italian poets from the VIP archive [4], comprising 4 versions each of Alba by Giorgio Caproni, Alle fronde dei salici by Salvatore Quasimodo, and La vita by Sandro Penna, including performances by the poets themselves.

## 3. Methodology

The annotation was performed by the two authors using Praat [16] and an updated version of the VIP protocol [5]. Alongside the four standard VIP tiers—rhythmical words, inter-pausal units, utterances, and verses—three additional tiers were introduced (based on a first-pass annotation from WebMAUS [17]): phonetic segments, syllables, and intonation labels indicating the pragmatic function of contours, following the system proposed by [18] and adapted for poetic speech by [4]. Pauses were classified by both duration and their position in the verse (e.g., punctuation, caesuras, verse or stanza boundaries). Intonation patterns were extracted and compared using AMPER segmentation and alignment procedures [19]. In addition, acoustic parameters were measured at structurally or expressively relevant points in the poems, following [15]. These included spectral measures (H1–H2, spectral emphasis, Hammarberg index, H/L ratio), periodicity indicators (HNR, autocorrelation), and prosodic features such as syllable duration, intensity, and pitch dynamics.

## 4. Results

This study confirms both the distinctiveness and the complexity of prosodic structures in poetic speech, also highlighting marked variation in stylistic and phonetic features across the selected readings. The prosodic text emerges as a sophisticated linguistic construction, where a sense of continuation and suspension predominates. Except for the final verse—typically delivered with a clearly assertive, declarative contour—most intonational units, even those aligned with strong punctuation, are marked by less descending poetic declaratives [4], often conveying explanatory or emphatic intentions. This structure also influences interrogatives, which frequently lack pragmatic necessity and function as rhetorical questions, displaying flat or non-rising contours that diverge from standard Italian read speech (Fig. 1) [20]. Beyond intonation, the analysis reveals distinctive prosodic and pausal strategies, including longer pauses, shorter inter-pause units, and lengthened vocalic segments compared to other speech registers [21]. These features reflect reduced syllabic compression and increased adherence to metrical structure. Pause placement is often guided more by stylistic and metrical considerations than by syntactic or semantic boundaries, emphasizing a notable asymmetry between prosody, text, and metre. This highlights the need to consider the unique pragmatics of the poetic text when analysing intonation in oral readings. Finally, with respect to voice quality, the study identifies consistent strategies across readings that favour what [15] define as "softened" voices, characterized by specific acoustic features and often perceived as more poetic in nature. Overall, our study demonstrates that poetry reading aloud is marked by substantial variation across multiple levels of the text, prosodic structure, and voice quality. Yet within this variability, consistent patterns also emerge, aligning with findings from [4, 7, 8].
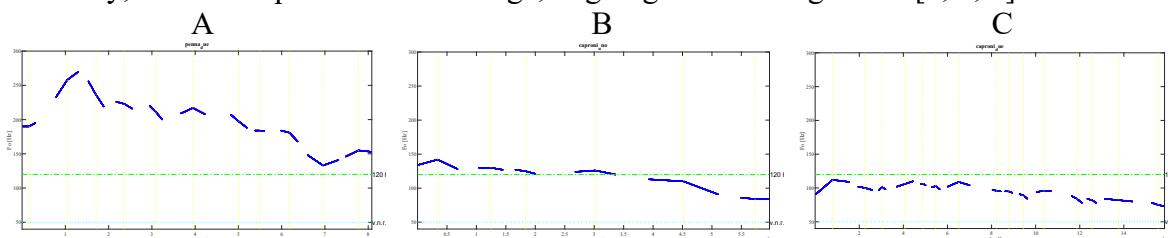


Figure 8 – A: Intonational pattern of a poetic declarative: "Un mare tutto fresco di colore" [A sea all fresh with color]; B: First intonational unit of an interrogative poetic utterance: "Io quale tram odo" [Which tram do I hear]; C: Second intonational unit of an interrogative poetic utterance: "Che apre e richiude in eterno le deserte sue porte?" [That opens and closes forever its deserted gates?]

## 5. References

Barbosa, P. A. (2023). The dance of pauses in poetry declamation. Languages, 8(1), 76.

Wagner, P. (2013). Meter specific timing and prominence in German poetry and prose. In O. Niebuhr (Ed.), Understanding Prosody: The Role of Context, Function and Communication (pp. 219-236). De Gruyter.

Meyer-Sickendiek, B., Hussein, H. & Baumann, T. (2017). Rhythmicalizer. In INFORMATIK 2017: Deep Learning in heterogenen Datenbeständen (2189-2200). Gesellschaft für Informatik.

Colonna, V. (2022). Voices of Italian Poets. Storia e analisi fonetica della lettura della poesia italiana del Novecento (Vol. 5, pp. 1-272). Edizioni dell'Orso.

Colonna, V. & Romano, A. (2023). VIP-RADAR: A model for the phonetic study of poetry reading. In Proceedings of the 20th ICPhS (pp. 594-598). Guarant International.

Lo Iacono, F., Colonna, V. & Romano, A. (2024). Preservation, conservation and phonetic study of the voices of Italian poets: A study on the seven years of the VIP archive. In Proceedings of INTERSPEECH 2024 – Speech and Beyond (pp. 3664-3668).

Colonna, V. Pamies Bertrán, A. and Damato, S. (2024) "Towards a phonetic history of the voices of Spanish poets: A first experimental study on the Generation of '27", Journal of Experimental Phonetics, 33, pp. 7–34. doi: 10.1344/efe-2024-33-7-34.

Colonna, V. (in press). "Poetry audio recordings and interdisciplinary phonetic study: The Voices of Spanish Poets Project: Challenges and Two Case Studies". In R. Dürr, K. Henkel & V. Keylin (eds.). Poetry, Music and Sound Art: Recent Medial Correlations. De Gruyter.

Marchenko, V. (2020). Speech intonation and music: a look at their dynamics within the song format. Journal of Language and Linguistic Studies, 16(2), 822-834.

Fabb, N. & Halle, M. (2008). Meter in poetry: A new theory. Cambridge University Press.

Byers, P. P. (1979). A formula for poetic intonation. Poetics, 8(4), 367-380.

Barney, T. (1999). Readers as text processors and performers: A new formula for poetic intonation. Discourse Processes, 28(2), 155–167.

Lo Iacono, F. (2024). A whirlwind of neo-avant-gardist voices: Unveiling sound experimentalism in the poetic performances of six members of the Gruppo 63. In Libro de resúmenes I Congreso Internacional de Teoría de la Lírica y Poéticas Comparadas (Vol. 1, pp. 89-90). Universidad de Salamanca.

Lo Iacono, F. (2025). Rhythm, Intonation and Voice Quality in Italian Experimental Poetry (1960s–1970s): Between Gruppo 63 and Mulino di Bazzano. CAIS Annual Conference. Panel: Italian Performance Poetry of the 20th and 21st Centuries.

Gafni, C. & R. Tsur. ""Softened" Voice Quality in Poetry Reading: An Acoustic Study." Style, vol. 51 no. 4, 2017, p. 456-481. Project MUSE.

Boersma, P. & Weenink, D. (2025). Praat: doing phonetics by computer [Computer program]. Version 6.4. 31, retrieved 3 May 2025 from http://www.praat.org/

Kisler, T. & Reichel U. D. and Schiel, F. (2017): Multilingual processing of speech via web services, Computer Speech & Language, Volume 45, September 2017.

Romano, A. (2014–2018). Etichette per l'analisi prosodica di file di parlato.

Contini, M., Lai, J.-P., Romano, A., Roullet, S., de Castro Moutinho, L., Coimbra, R. L., Pereira Bendiha, U. & Secca Ruivo, S. (2002). Un projet d'atlas multimédia prosodique de l'espace roman. In Proceedings of the International Conference Speech Prosody 2002 (227-230). Aix-en-Provence.

Colonna, V. & Romano, A. (2023). Prose or poetry? A perceptive phonetic study. In Proceedings of the 20th ICPhS(pp. 347-351). Guarant International.

# Whistled Gascon Occitan: Linguistic Encoding and Cognitive Insights from a Formant-Based Acoustic Modality

Michela Russo[1,2], Alexandre Di Caro[3]

*[1]SFL CNRS, [2]University of Paris 8 [3]UJML 3*

mrusso@univ-paris8.fr

## 1. Introduction

The Occitan language family, and more precisely its Gascon variety, harbors a lesser-known yet striking form of whistled speech, functionally akin to systems such as Silbo Gomero (Canary Islands) or Sfyria (Greek). In the remote mountainous zones of Béarn and the central Pyrenees, Gascon herders developed a whistled code to communicate over long distances—across valleys and grazing plateaus where oral speech would dissipate or become unintelligible. Though severely endangered and understudied, this whistled tradition persists in oral memory and echoes a long-standing agro-pastoral culture.

## 2. Corpus and Data Source

The present analysis is based on a newly available corpus of whistled Gascon Occitan, recorded and published in 2024 for the first time in open access. This dataset, hosted on the Lingua Libre platform (a Wikimedia Foundation project: https://meta.wikimedia.org/wiki/Lingua_Libre/Occitan_whistle), provides a structured and searchable collection of whistled utterances drawn from the Béarnais dialect of Occitan, specifically the traditional whistled speech of Aas, a village in the Ossau valley. These recordings represent a major breakthrough in the documentation of this endangered linguistic form. The whistled forms were produced by neo-speakers—individuals who acquired the whistled modality through revitalization efforts rather than through traditional intergenerational transmission. Despite this, the recordings offer a high degree of phonetic fidelity to the original system and provide the first reproducible basis for acoustic and phonological analysis of whistled Gascon.

Gascon Occitan presents a phonological system particularly suited for whistled transposition. It contains seven tonic vowels—/i, y, e, ɛ, a, o, u/—each of which can be mapped onto a distinct harmonic space in the whistled signal. This mapping follows the principle of formant transposition, particularly that of the second formant (F2), into the first harmonic (H1) of the whistled form. Much like in Sfyria and in Silbo Gomero (see Rialland 2005; Russo & Katsiouris 2025), vowel identity in whistled Gascon is determined not through the full formant spectrum but by the frequency position and modulation of H1, which functions as a spectral surrogate for F2.

Whistled languages, whether based on tonal (e.g., Mazatec) or non-tonal (e.g., Gascon, Greek) systems, showcase a unique bioacoustic interface between cognition, perception/production, and the environment. In non-tonal languages like Gascon, the whistled signal simplifies spoken input by selecting robust phonetic cues—primarily F2 for vowels and transitional contours for consonants. The brain's tonotopic organization in the auditory cortex allows listeners to interpret these cues efficiently, even when the spectral richness of speech is severely reduced.

Gascon whistlers likely developed perceptual compensation mechanisms—using amplitude modulations, frequency shifts, and vowel-consonant transitions—to reconstruct phonological structure from minimal signals. These cognitive strategies parallel those used in speech-in-noise or fast speech processing, where the listener 'fills in' missing information based on context and expectation.

Recent neurocognitive research on whistled languages (Carreiras et al., 2005) shows bilateral hemispheric processing, with increased involvement of the right auditory cortex,

which is typically associated with pitch, music, and prosody. In Gascon whistled speech, the reliance on musical-like pitch patterns to encode linguistic information may invoke similar mechanisms, suggesting a hybrid processing stream: linguistic in content, musical in modality.
Acoustic Phonetics of Whistled Gascon: An Untapped Corpus

While other whistled languages have been subject to detailed spectrographic and segmental analysis, whistled Gascon awaits such systematic treatment. Front vowels like /i/ and /e/ correspond to higher H1 frequencies, while back vowels /o/ and /u/ are associated with lower harmonics, though never below ~1200 Hz due to the physical constraints of human whistling.

Gascon consonants in whistled speech were reportedly encoded through modulations in intensity and pitch, with stops producing brief interruptions or dips in the harmonic trajectory— especially /p, t, k/—and fricatives signaled through sharp transitions rather than turbulence.

## 3. Results

Across whistled languages, we observe a convergence toward universal cognitive strategies. Whether in Gascon, Greek, Turkish, or Spanish, the simplification of speech to a single harmonic channel forces languages to adapt along predictable lines: prioritizing formant transitions, exploiting amplitude cues, and relying on rhythmic regularity. This shows not just the plasticity of speech but the adaptability of phonological systems under ecological pressure.
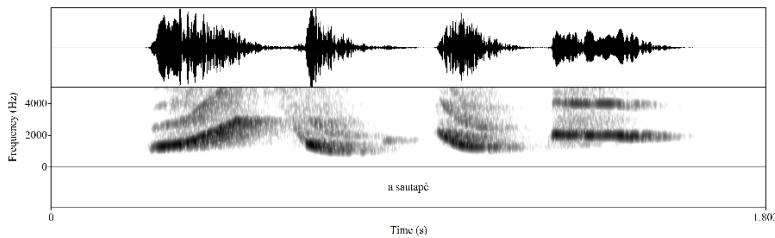


Figure 1. Spectrograms of a sautapè in whistled Gascon Occitan. Consonants are signaled through modulations in the harmonic trajectory (H1), which reflect articulatory loci: coronal segments (e.g., [t]) correspond to mid-to-high inflection zones (approx. 1800–2400 Hz).
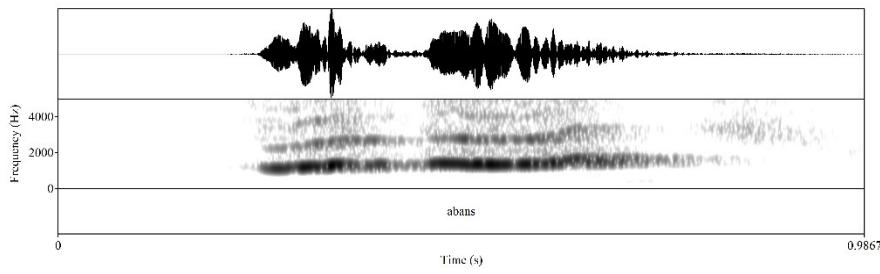


Figure 2. Spectrogram of abans in whistled Gascon Occitan: labials (e.g., [β], [p]) trend lower (below 1500 Hz). Transitions across places of articulation are visible as frequency inflections or interruptions in the whistled harmonic contour.

We have conducted a detailed analysis of the whistled corpus and mapped the segmental correspondences with spoken Gascon, identifying how phonological contrasts—particularly place of articulation and vowel quality—are acoustically encoded in the whistled signal. These findings are further contextualized through a cross-linguistic comparison with other whistled languages (e.g., Turkish, Silbo Gomero, Greek), allowing us to propose candidate universal strategies for phonological encoding under extreme spectral reduction.

# Segmental vs. Suprasegmental Glottalization in a Jin Chinese Dialect: Phonetic Pathways of Tone and Phonation

Xiaojing Du
*University of Cambridge*
xed20@cam.ac.uk

## Abstract

This paper examines how phonation supplements pitch in the tonal system of Huoji Chinese, a lesser-known Jin dialect. Focusing on historical checked tones and emergent neutral tones, we show that both are short and non-modal, yet differ in origin: checked tones retain segmental traces of lost stop codas, while neutral tones recruit prosodically induced glottalization or breathy voice. Based on data from 17 speakers, acoustic analyses of F0, H1–H2, HNR, and duration reveal that neutral tones vary between modal, creaky, and breathy realizations, with breathy voice surfacing when glottalization is absent. These findings demonstrate that tonal identity in Huoji is dispersed across pitch, phonation, and duration, challenging pitch-dominant models and highlighting phonation as an active resource in ongoing tonal restructuring.

**Keywords:** glottalisation, phonation, tonal variation, Jin Chinese, prosodic reduction

## 1. Introduction and background

This study examines tonal variation in Huoji Chinese, a Jin dialect of Henan, with a focus on two marginal categories: the historical checked tone (T5) and a prosodically reduced tone akin to Mandarin's neutral tone. Jin dialects occupy a transitional space between conservative Middle Chinese reflexes and Mandarin-driven innovations, making Huoji an ideal case for exploring how phonation and glottalization interact with pitch. Historically, checked tones were marked by stop codas (-p, -t, -k), which in Huoji have been reinterpreted through glottalization and changes in voice quality. Meanwhile, new reduction processes yield short, unstressed syllables resembling neutral tones. The overlap of these two categories raises key questions: how are tonal distinctions preserved when pitch contours converge, and can voice quality carry contrast where pitch fails?

Traditional models of tone privilege pitch as the primary cue. Yet under structural pressure—through merger, contact, or reduction—other phonetic dimensions emerge. In Huoji, phonation is not incidental but a core parameter of tonal dispersion, ensuring contrast even as pitch categories blur.

## 2. Methodology

Data come from 17 native speakers of Huoji, recorded during a 2024 pilot study and subsequent fieldwork (May, 2025) in Henan Province. Participants ranged in age from 11 to 83, balanced across gender and literacy. Recordings were made in quiet home settings using a 24-bit/48 kHz USB condenser microphone. After quality screening, 10 speakers (approx. 7000 tokens) were retained for detailed acoustic analysis.

## 3. Materials and elicitation

Three types of data were collected in order to capture both citation and contextualized tonal realizations. First, monosyllabic citation tones were elicited using minimal and near-minimal pairs covering all five tones of Huoji Chinese. Second, disyllabic constructions were recorded, including both compounds and affixed forms, to test the effects of reduction and sandhi. Finally, scripted dialogues were designed to elicit broad versus narrow focus, providing a controlled paradigm for examining prosodic prominence and tonal alternations.
Acoustic analysis

Segmentation was performed in Praat, with vowel nuclei marked from the onset of periodic energy to the offset of voicing. Four acoustic measures were extracted using VoiceSauce: fundamental frequency (F0), z-scored per speaker and checked for halving in creaky spans; H1–H2, used as a proxy for spectral tilt and voice quality; harmonics-to-noise ratio (HNR); and vowel duration, normalized per speaker. In addition to instrumental measurements, each token was auditorily and instrumentally coded for phonation type, including modal, creaky, breathy, devoiced, or the presence of a glottal stop. Importantly, breathy voice was observed primarily in neutral-tone contexts that lacked glottalization, suggesting that it may function as a supplementary cue when other laryngeal features are absent.

Statistical analysis

To examine how tonal categories are dispersed in acoustic space, we used multidimensional scaling (MDS). Each token was specified by normalized values of F0 (onset, offset, mean, delta), duration, H1–H2, and HNR. In baseline configurations with F0 and duration only, the checked tone (T5) overlapped substantially with T1 and T3, suggesting that pitch alone cannot sustain its distinctiveness. When HNR was added, T5 shifted outward, reflecting its systematically low harmonicity. With H1–H2 included, T5 became a fully separate cluster. Inter-category distance ($\mu x$) nearly doubled ($\approx 0.20 \rightarrow \approx 0.50$), quantitatively confirming that phonation cues enlarge the tonal space and preserve contrast. This approach highlights that voice quality is not peripheral but essential to the identity of Huoji's marginal tones.

## 4. Results

Checked tones consistently showed glottal stops, creaky phonation, short duration, negative H1–H2, and low HNR, confirming their segmental origin as relics of lost codas. Neutral tones showed more variable phonation: affixed forms often glottalized, while root-derived reductions alternated between modal, creaky, and breathy voice, with breathy realizations most frequent in weak contexts. Under focus, neutral tones re-expanded in duration, whereas checked tones resisted lengthening.

## 5. Discussion and conclusion

Our findings challenge the view that tone is solely a matter of pitch height and contour. In Huoji, tonal categories are dispersed not only along pitch but also across dimensions of phonation and duration. Glottalization emerges as both a historical residue and an adaptive feature, expanding the phonetic space for contrast maintenance. Including voice quality in tonal analysis is not just beneficial—it is essential for understanding systems where pitch categories are under structural pressure. As younger speakers increasingly conflate forms like 匹 [pʰiʔ] 'measure word' (historically checked tone) and 批 [pʰi] 'to approve' (historically level tone), the persistence of non-modal phonation becomes a key stabilizing mechanism. Phonation anchors contrast when pitch becomes ambiguous, ensuring intelligibility even as tone categories shift.

This multidimensional model aligns with proposals by Zhu and Kuang, who argue for integrated tone–phonation frameworks in Chinese. It also resonates with typological parallels in Zapotec and Xiapu Min, where glottalization and voice quality play systematic roles in tone systems. By treating phonation as a core parameter in tonal dispersion, we gain a more complete and resilient account of phonological variation and change.

## 6. References

Abramson, Arthur S. & Luangthongkum, Theraphan (2009). A fuzzy boundary between tone languages and voice-register languages. In Gunnar Fant, Hiroya Fujisaki & Jianfen Shen (eds.), Frontiers in Phonetics and Speech Science (149–155). Beijing: The Commercial Press.

Andruski, Jean E. (2006). Tone clarity in mixed pitch/phonation-type tones. Journal of Phonetics,34(3), 388–404.

Blankenship, Barbara (2002). The timing of nonmodal phonation in vowels. Journal of Phonetics,30(2), 163–191.

Boersma, Paul & Weenink, David (2020). Praat: doing phonetics by computer Version 6.1.15.http://www.praat.org/. Accessed on 31/05/2025.

Esposito, Christina M. (2010). Variation in contrastive phonation in Santa Ana Del Valle Zapotec.Journal of the International Phonetic Association, 40(2), 181–198.

Garellek, Marc & Keating, Patricia (2011). The acoustic consequences of phonation and toneinteractions in Jalapa Mazatec. Journal of the International Phonetic Association, 41(2), 185–205.

Gordon, Matthew & Ladefoged, Peter (2001). Phonation types: a cross-linguistic overview. Journal of Phonetics, 29(4), 383–406.

Keating, Patricia & Esposito, Christina M. (2007). Linguistic voice quality. UCLA Working Papers in Phonetics, 105, 85–91.

Kreiman, Jody & Gerratt, Bruce R. (2001). Toward a taxonomy of nonmodal phonation. Journal of Phonetics, 29(4), 365–381.

Kuang, Jie (2011). Production and perception of the phonation contrast in Yi. M.A. thesis, University of California Los Angeles.

Ladefoged, Peter & Maddieson, Ian (1996). The sounds of the world's languages. Oxford: Blackwell.

Laver, John (1980). The phonetic description of voice quality. Cambridge: Cambridge University Press.

Redi, Leah & Shattuck-Hufnagel, Stefanie (2001). Variation in the realization of glottalization in normal speakers. Journal of Phonetics, 29(4), 407–429.

Shue, Y.-L., Keating, Patricia, Vicenik, Christine & Yu, Kristine (2011). VoiceSauce: A program for voice analysis. Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), 1846–1849.
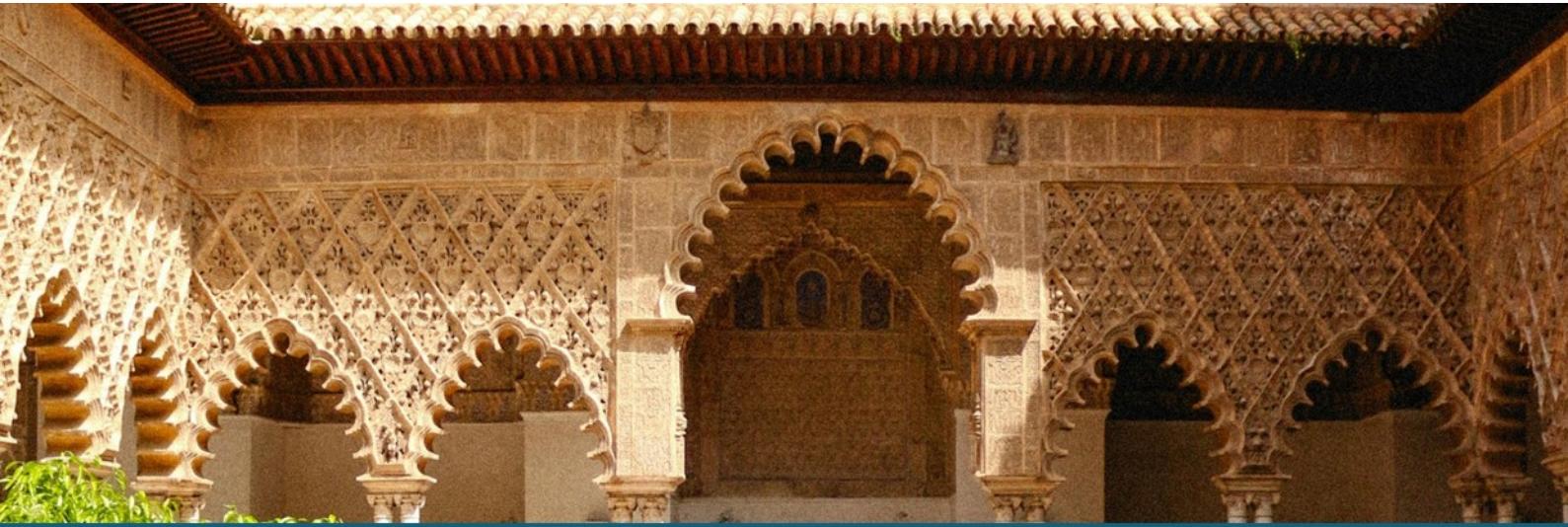
# LIQUID CONSONANTS

# Pennsylvania Dutch liquids: phonetic correlates of the sonority hierarchy

Rose Fisher, Erin Noelliste, David Bolter

*Michigan State University, University of Northern Colorado, Humboldt-Universität zu Berlin*

rosefisher96@gmail.com, erin.noelliste@unco.edu, dbolter.linguist@gmail.com

**Abstract**

This paper investigates the phonetic correlates of sonority of Pennsylvania Dutch (PD) liquid allophones. PD, a variety of German spoken in North America, is known to have a diverse set of liquid realizations, including clear [l], dark [ɫ], approximant [ɹ], tapped [ɾ], and vocalized [ɐ]. The realization of these allophones is partially phonologically conditioned and partially conditioned by language contact. In this paper, we investigate discrete phonetic properties of the aforementioned allophones, specifically intensity and duration, with the goal of understanding the phonological patterning with regards to the sonority hierarchy.

**Keywords:** liquids, sonority, Pennsylvania Dutch, intensity, duration

## 1. Introduction and background

PD is an under-studied Germanic variety, closely related to European varieties of German, specifically Palatinate German, which has been spoken in North America for well over 300 years (Keiser 2012). Due to its language history, PD has evolved such that it contains numerous allophones of /l/ and /r/, many of which mirror those found both in European varieties of German, as well as American English varieties (Keiser 2012). With regards to the current study of synchronic PD, we investigate the intensity and duration of the liquid allophones drawing on previous research on sonority from Parker (2002, 2008, 2011) and data from two modern PD dialects. Representative data for the PD liquid allophones in this study are provided in (1).

(1)      Liquid Allophones in Midwestern PD
a.       approximant [ɹ]      [fɹɔg]          Frog            'frog'
b.       vocalized /r/ [ɐ]     [dɛɐ]           Deer            'door'
c.       tap [ɾ]               [o.ɾə]          Ohre            'ears'
d.       clear [l]             [fɛn.sə.meɪs.lɪ]   Fensemeisli        'chipmunk'
e.       dark [ɫ]              [æɫ]            Owl             'owl'
Eastern PD (spoken in Pennsylvania) has all allophones in (1) except the tap.

In Parker's works on sonority (2002, 2008, 2011), he shows that rhotics must be divided into different levels of sonority and gives this universal ranking for the sounds investigated in the current study: rhotic approximants > flaps > laterals. Recent work on Bavarian German (cf. Noelliste 2019, Bolter & Noelliste forthcoming) adapts Parker's universal ranking to account for the fact that Bavarian varieties of German exhibit a flap that is less sonorous than the lateral, due to its shorter duration and lower intensity. The current study on PD takes a similar approach to the methodology developed in Bolter & Noelliste (forthcoming) by considering these liquids of a German variety in terms of phonetic correlates of the sonority hierarchy. This is particularly compelling for PD since there are multiple allophonic realizations of /l/ and /r/, including velarized [ɫ] and approximant [ɹ], both of which have been recognized as having more tongue body gestures that are more vocalic in nature (on [ɫ], see Proctor 2011, Sproat & Fujimura 1993; for [ɹ], Delattre & Freeman 1968, Zhou et al. 2008). Thus, the question we address is: Do the phonetic correlates of sonority (intensity and duration) illustrate for PD, like they do for Bavarian German, that different realizations of /r/ can be both more and less sonorous than /l/ (rhotic approximants > laterals > flaps/taps)?

## 2. Methodology

Liquid realizations for this study are taken from eight recorded tellings of the children's story *Frog, where are you?* (Mayer 1969). All eight speakers are native PD speakers; four speakers are from the Lancaster County, Pennsylvania Amish community, which has no tap, and four are from the Holmes County, Ohio Amish community, which has all the liquid allophones in (1). Each liquid realization was manually segmented in Praat (Boersma & Weenink 2021) and impressionistically labeled by allophone. Intensity and duration measurements were made between the boundaries of each realization using a Praat script. Segmentation was based in part on work by Skarnitzl (2009), who discusses challenges of segmenting /l/ manually for Czech. Both clear and dark realizations of /l/ were segmented apart from adjacent vowels by assessing the following properties: formant differences, lateral antiformants, lower intensity, simpler waveform, visible release phase of /l/. For [ɹ], formant differences and the general lower intensity were helpful, but the most informative characteristic was generally the considerably lower F3 as compared to most adjacent vowels.

## 3. Results

The results of this study show that, in general, the tap/flap in PD is, indeed, the least sonorous liquid having both the lowest duration and intensity of all the other liquids. The approximant has slightly lower duration than /l/ but higher intensity. Additionally, this study shows that clear [l] is consistently less sonorous (lower duration and intensity) than its allophone dark [ɫ].

| Mean Duration Readings | | | Mean Intensity Readings | |
|---|---|---|---|---|
| Sound | Mean Duration | | Sound | Mean Intensity |
| ɾ̞ | 0.08280781 | | ɾ̞ | 66.59220 |
| ɹ | 0.05857935 | | ɹ | 71.94749 |
| ɫ | 0.09682101 | | ɫ | 73.29079 |
| l | 0.06552746 | | l | 67.60225 |
| ɾ | 0.04126010 | | ɾ | 64.87863 |

## 4. Discussion and conclusion

Based on his results, Parker (2002, 2008, 2011) concludes that there is a universal sonority hierarchy where /l/ is always less sonorous than /r/. Our work, alongside Noelliste (2019) and Bolter & Noelliste (forthcoming), reveals that realizations of /l/ can be placed between realizations of /r/ (taps/flaps are less sonorous and approximants more sonorous than /l/) in Bavarian German and PD. To date, there is no extant corpus of recorded spontaneous PD speech; thus, the recordings used in this study contribute to the creation of such a corpus. The current ongoing study adds not only to the linguistic research on PD, but also contributes to the larger body of cross-linguistic studies on universal and relative sonority of liquids.

## 5. References

Boersma, Paul & Weenink, David (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.39. http://www.praat.org/. Accessed on 05/03/2025.

Bolter, David & Noelliste, Erin (forthcoming). A phonetic account of the sonority hierarchy with reference to Bavarian German <rl>. Journal of Germanic Linguistics.

Delattre, Pierre & Freeman, Donald C. (1968). A dialect study of American r's by x-ray motion picture. Linguistics, 6(44), 29–68.

Keiser, Steven Hartman (2012). Pennsylvania German in the American Midwest. Durham: Duke University Press.

Mayer, Mercer (1969). Frog, Where are You? New York: Dial Books.

Noelliste, Erin (2019). Bavarian German r-Flapping: Evidence for a dialect-specific sonority hierarchy. Glossa: a journal of general linguistics, 4(1).

Parker, Steve (2002). Quantifying the sonority hierarchy. Amherst, MA: University of Massachusetts dissertation.

Parker, Steve (2008). Sound level protrusions as physical correlates of sonority. Journal of phonetics, 36(1), 55–90. DOI: https://doi.org/10.1016/j.wocn.2007.09.00

Parker, Steve (2011). Sonority. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), The Blackwell companion to phonology (1160–1184). Oxford: Wiley-Blackwell. DOI: https://doi.org/10.1002/9781444335262.wbctp0049

Proctor, Michael (2011). Towards a gestural characterization of liquids: Evidence from Spanish and Russian. Laboratory Phonology, 2(2), 451–485.

Sproat, Richard & Fujimura, Osamu (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. Journal of Phonetics, 21(3), 291–311.

Skarnitzl, Radek (2009). Challenges in segmenting the Czech lateral liquid. In Anna Esposito & Robert Vích (eds.) Cross-modal analysis of speech, gestures, gaze and facial expressions (162–172). Berlin, Heidelberg: Springer.

Zhou, Xinhui, Y. Espy-Wilson, Carol, Boyce, Suzanne, Tiede, Mark, Holland, Christy & Choe, Ann (2008). A magnetic resonance imaging-based articulatory and acoustic study of "retroflex" and "bunched" American English /r/, Journal of the Acoustical Society of America, 123(6), 4466–4481.

# The liquids are flowing: intervocalic laterals in Cibaeño Dominican Spanish

Erik W. Willis
*Indiana University*
ewwillis@iu.edu

**Abstract**

Dominican Spanish is known for innovative liquid sound patterns, including pre-breathy rhotics, coda switching between laterals and rhotics, and coda vocalization. However, intervocalic laterals have received little attention in descriptions of the Spanish sound system. Typically described as a "light" or anterior approximant segment, these laterals contrast with the "dark" or velarized variants found in other languages. This study analyzes intervocalic laterals from 24 speakers of Cibaeño Dominican Spanish, recorded in Santiago and divided by social class and sex. Fifteen tokens per speaker were extracted from a picture description task, resulting in 360 tokens. Segment duration and formant values were measured, with normalization via the Bark scale and a vowel-based ratio to account for anatomical differences. F2 values in intervocalic position ranged widely, from 2,000 Hz to 900 Hz, with marked differences by sex. An assimilatory effect from the preceding vowel appears to influence F2 range. Many laterals showed characteristics typical of velarized realizations, with males producing more back articulations, suggesting a robust and variable lateral system in this variety.

**Keywords:** Caribbean Spanish, laterals, velarization, assimilation

## 1. Introduction and background

Dominican Spanish has long been recognized for its innovations in sound patterns that include extreme consonant lenition (Jiménez Sabater 1975), coda /s/ reduction (Alba 1990), nasalization of vowels (Haché 1981, Bongiovanni 2018), and rhotics with pre-breathy voice (Willis 2007), coda switching between laterals and rhotics (Willis and Ronquest 2021), and coda vocalization (Jiménez Sabater 1975, Pérez Guerra 1988, Alba 2004). The current project explores another potential innovation—variation in intervocalic /l/ realization—by examining the lingual positioning of the tongue through acoustic measures of F2 values and segment duration to describe intervocalic laterals in Cibaeño Dominican Spanish.

Most overviews of the Spanish sound system dedicate minimal discussion to intervocalic laterals. They are described as an approximant lateral segment that is "light" or anterior compared to "dark" laterals that are typically velarized as in American English (Olive et al. 1993).) or Catalan (Simonet 2010, Davidson 2020, Recasens and Espinosa 2005). Quilis reports an average duration of 53ms and an F2 above 1500 Hz. (1979). In coda position, Spanish laterals are claimed to assimilate in place of articulation resulting in a dental allophone [l̪]or alveo-palatal or palatalized allophone [lʲ]. These allophonic variations are still considered light or anterior. The current investigation developed from an observation of intervocalic laterals during an investigation of coda liquids in this Dominican variety.

## 2. Methodology

The corpus for the current study comes from 24 speakers of Dominican Spanish recorded in Santiago, Dominican Republic, and is divided by social class and sex. The first 15 tokens of intervocalic /l/ were extracted from a modified picture description interview task from each speaker for a total of 360 tokens. The tokens were not balanced for vowel quality. The audio was recorded with a USBpre external soundcard and a head-mounted Shure Wh10a microphone. Measures of segment duration and formant measures were extracted from the midpoint of the segment. The F2 measures were examined with and without normalization.

Normalization processes include the Bark scale and the Vowel Centroid Ratio (Solon 2016) to account for differences in anatomy.

## 3. Results

The data showed lateral F2 realizations that ranged from 2,000 Hz down to 900 Hz in intervocalic position. The raw F2 values and Bark scale analyses show large differences that exhibit different distributions based on speaker sex. Vowel centroid-normalized values also show differences but require additional analysis. The darkest or most velarized tokens approached normalized values for English dark laterals reported by Solon (2016). Our preliminary analysis also indicates that there is an assimilatory process from the preceding vowel that conditions the range in F2 values These low F2 values indicate a robust lateral variation conditioned by vowels that extends well into the F2 formant characterizations of a dark or velarized lateral (Recasens & Espinosa 2005).

## 4. Discussion and conclusion

These findings and the acoustic descriptions of the lateral segment suggest that there is much more potential variation of laterals than previously documented regarding duration and velarization. The laterals also exhibited a wide range of durations. Many intervocalic laterals had a duration below 40 ms, which Melero and Cisneros (2020) found to be perceptually associated with a Spanish phonological tap. Males and females showed different distributions, with the males producing more velarized productions. The Cibaeño Dominican Spanish intervocalic lateral shows F2 values that vary far beyond 1500-1600 range typically referenced for light /l/ velarized F2 values suggest that the lateral overflows the boundaries assumed for a light /l/ in Spanish (Hualde 2014).

## 5. References

Alba, Orlando (1990). Variación fonética y diversidad social en el español dominicano de Santiago. Santiago: Pontificia Universidad Católica Madre y Maestra.

Alba, Orlando (2004). Cómo hablamos los dominicanos. Santo Domingo: Grupo León Jimenes.

Bongiovanni, Silvina (2018). Production of Anticipatory Vowel Nasalization and Word-Final Nasal Consonants in Two Dialects of Spanish. PhD dissertation, Indiana University.

Davidson, Justin (2020). Spanish phonology in contact with Catalan: on implementations of gradience and discreteness in the study of sociolinguistic variation of laterals. In Rajiv Rao (ed.), Spanish Phonetics and Phonology in Contact, 337–359. Amsterdam: John Benjamins.

Haché de Yunén, Ana (1982). La /n/ final de sílaba en el español de los caballeros. In Orlando Alba (ed.), El Español del Caribe: Ponencias del VI Simposio de Dialectología, 261–271. Santiago: Universidad Católica Madre y Maestra.

Jiménez Sabater, Maximiliano A. (1975). Más datos sobre el español en la República Dominicana. Santo Domingo: Ediciones Intec.

Melero-García, Fernando & Alejandro Cisneros (2020). No es tan simple como parece: the effect of duration of one-closure rhotics on the perception of Spanish /ɾ/ and /r/. In Diego Pascual y Cabo & Idoia Elola (eds.), Current Theoretical and Applied Perspectives on Hispanic and Lusophone Linguistics, 233–255. Amsterdam: John Benjamins.

Olive, Joseph, Alice Greenwood & John Coleman (1993). Acoustics of American English Speech: A Dynamic Approach. New York: Springer-Verlag.

Pérez Guerra, Irene (1991). Un caso de prestigio encubierto en el español dominicano: la vocalización cibaeña. In Actas del III Congreso Internacional del Español de América, 3: 1185–1192.

Quilis, Antonio, Manuel Esgueva, María Luz Gutiérrez Araus & Margarita Cantarero (1979). Características acústicas de las consonantes laterales españolas. Lingüística Española Actual, 1: 233–344.

Recasens, Daniel & Aina Espinosa (2005). Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: evidence from two Catalan dialects. Journal of the International Phonetic Association, 35: 1–25.

Simonet, Miquel (2010). Dark and clear laterals in Catalan and Spanish: interaction of phonetic categories in early bilinguals. Journal of Phonetics, 38: 663–678.

Solon, Megan (2017). Do learners lighten up?: phonetic and allophonic acquisition of Spanish /l/ by English-speaking learners. Studies in Second Language Acquisition, 39: 801–832.

Willis, Erik W. (2007). An acoustic study of the pre-aspirated trill in narrative Cibaeño Dominican Spanish. Journal of the International Phonetic Association, 37: 33–49.

Willis, Erik W. & Rebecca E. Ronquest (2022). A socio-phonetic exploration of coda liquids and vocalization in Cibao Dominican Spanish. In Manuel Díaz-Campos (ed.), The Routledge Handbook of Variationist Approaches to Spanish, 563–585. London/New York: Routledge.

# Andalusian liquids and VOX: An analysis of lateral and rhotic variation in the political speech of southern Spain

Matthew Pollock
*Louisiana State University Shreveport*
mpollock@lsus.edu

**Abstract**

This study examines variable liquid phenomena in a corpus of political speech across four cities in Spain: Madrid, Cordoba, Malaga, and Seville. Following acoustic and auditory analysis of 50 politicians in three political parties, results show that liquid variation is shaped by linguistic context, speaker identity, and stylistic factors. Lateral and rhotic modifications are most common among male speakers, in coda position, and in unscripted contexts. Members of the VOX party show high rates of liquid opacity, suggesting a larger stylistic strategy on the part of alt-right politicians. These findings highlight the social and phonetic complexity of liquid behavior in political speech in southern Spain and draw attention to the variability of productions associated with Andalusian identity across the region.

**Keywords:** Andalusian Spanish, sociophonetic variation, liquids, political speech, regional variation

## 1. Introduction and background

Spanish politics experienced a turn to the right in the 2019 elections, marked by the surging popularity of the alt-right VOX party, a spin-off from the moderate-right Partido Popular (PP; Rama et al., 2021). This shift was particularly notable in Andalusia, when the socialist PSOE's 40-year control of the regional government ended. To contribute to our understanding of this socio-political shift, the current study compares how southern politicians differ from those in the capital, and how individuals across the political spectrum differ from alt-right speakers by examining the production of the lateral /l/, trilled rhotic /r/, and tapped rhotic /ɾ/ in the political speech of Andalusia. This builds on previous research into political speech in the region (e.g., Cruz-Ortiz, 2022; Pollock & Wheeler, 2022; Pollock, 2023).

### 1.1. Sociolinguistic style and liquids in Spain

We all style-shift as a part of constructing a coherent social identity (e.g., Coupland, 2001; Eckert, 2008). However, politicians in particular use language in a high-stakes way not only to reflect social position, but to appeal to voters. Researchers have explored this variation across peninsular contexts at both the phonetic and phonological level (e.g., Fernández de Molina Ortés, 2021; Hernández-Campoy & Cutillas-Espinosa, 2013; Pollock, 2023).

Sociolinguistic research in southern Spain in particular suggests that liquid processes correlate with speaker socioeconomic status, age, gender, and regional identity. Ruiz-Peña (2013) describes processes approaching rhotacism (i.e., where /l/ is produced as [ɾ]) and lateralization (i.e., the reverse), as being influenced by social difference; in particular, working-class Seville speakers affiliated with socialist political parties and proud of their Andalusian identity were most likely to demonstrate variable production. Henriksen and Willis (2010) and Henriksen (2014) both find that older speakers and men tend to produce more innovative reduced variants in Jerez, while in Málaga, Zahler and Daidone (2014) find that speakers under 35 years of age favor the innovative reduced variant.

Various linguistic factors have been used to describe variable production: these include word position (Henriksen & Willis, 2010; Willis, 2006), lexical stress (Henriksen & Willis, 2010; Zahler & Daidone, 2014), and phonetic context (Bradley, 2006; Henriksen, 2014). Henriksen (2014) found phonetic context, in particular, to be predictive of reduction in northern

central peninsular Spanish: preceding high back vowels (i.e., /u/) favored innovative reduced variants in León and Ciudad Real. In Malaga, Zahler and Daidone (2014) determined that innovative reduced variants were disfavored in stressed syllables, and that frequency and phonological neighbors play a role in production of the reduced variant. Perceptually, Herrero de Haro (2017) finds that difference is maintained between elided word-final rhotics and elided final /s/ (e.g., [amo] in amos and amor can be distinguished). This process of reduction reflects similar ones underway in other Romance languages like Portuguese, where the alveolar trill has become a voiceless velar fricative (Parkinson, 1988).

## 2. Methodology

Contributing to research on sociophonetic variation in Spanish political speech, this study analyzes sociophonetic variation in the production and elision of the lateral /l/, tap /ɾ/, and trill /r/ among 50 political figures from four cities: Seville, Málaga, Madrid, and Córdoba. The speakers were balanced by political party (VOX, Partido Popular, and PSOE), gender, and regional origin. All audio was drawn from scripted public speeches delivered between 2013 and 2022. For each liquid, tokens were auditorily and acoustically analyzed using Praat and analyzed as being normative, lateralized, elided, reduced, or rhotacized productions. Acoustic measures, including F2, F3 and intensity range, were used to support liquid coding. Independent variables included the social factors political party, region, gender, audience type, and a stylistic measure based on the speech timestamp, along with linguistic factors including position in word, phonetic environment, and segment duration. Mixed-effects logistic regression models were developed to identify significant predictors of variation.

## 3. Results

Mixed-effects logistic regression models reveal that regional variants are conditioned by both linguistic and social factors. Lateral variation is favored in coda position, and is more common among male speakers, especially those from Málaga and Córdoba. Rhotic variation is most frequent in word-final position, especially preceding pauses or coronals, and is favored by conservatives and speakers from Seville and Málaga. Trill variation, though less common, is conditioned by gender and orthographic context. In all three cases, unscripted speech contexts promote regional realizations. Acoustic measurements, including F3 and intensity range for laterals and F2 range for taps, further help differentiate allophonic realizations and correlate with articulatory effort.

Political party also plays an important role in describing liquid variation, with members of the far-right VOX showing the highest rates of lateral opacity, and employing these variants earlier in their speeches, suggesting part of a larger stylistic norm belonging to the populist alt-right. Male VOX speakers from the NCPS region are particularly likely to favor these forms. Tap opacity is also common among VOX and PP speakers, and trill variation, while limited, was favored by VOX politicians and before specific audiences, in shorter segments, and by male speakers.

## 4. Discussion and conclusion

These findings suggest that the Spanish liquids are influenced not only by linguistic environment, but also aspects of speaker sociolinguistic identity and stylistic goals. While variation is most prevalent among male speakers and in coda or medial positions, the VOX party exhibits unique rates of liquid modification, combining regionally salient forms with nationally normative ones. These liquid variants function as part of a larger indexical field of meaning (Silverstein, 2003), associated with local identity, and are markers of solidarity salient to constituents. This aligns with broader political trends positioning VOX as ideologically

distinct from both mainstream parties, particularly by using phonetic variation as part of a broader effort to index outsider authenticity and cultural nationalism.

## 5. References

Bradley, T. G. (2006). Phonetic realizations of /sr/ clusters in Latin American Spanish. Díaz-Campos, M. (Ed.) Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology (1–13). Somerville, MA: Cascadilla Proceedings Project.

Coupland, N. (2001). Language, situation, and the relational self: theorizing dialect-style in sociolinguistics. In Eckert, Penelope, Rickford, John R. (Eds.), Style and Sociolinguistic Variation, pp. 185–210. Blackwell, Malden/Oxford.

Cruz-Ortiz, R. (2022). Sociofonética andaluza. Caracterización lingüística de los presidentes y ministros de Andalucía en el Gobierno de España (1923–2011). Berlin: De Gruyter.

Eckert, P. (2008). Variation and the indexical field. Journal of Sociolinguistics 12, 453–76.

Fernández de Molina Ortés, E. 2021. Estudio sociolingüístico de las intervenciones políticas y públicas de los presidentes del gobierno extremeño. Cultura, Lenguaje y Representación 26, 167–87.

Henriksen, N. (2014). Sociophonetic analysis of phonemic trill variation in two sub-varieties of Peninsular Spanish. Journal of Linguistic Geography 2, 4–24

Henriksen, N. & Willis, E. W. (2010). Acoustic characterization of phonemic trill production in Jerezano Andalusian Spanish. In Marta Ortega-Llebaria (ed.), Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology (115–127). Somerville, MA: Cascadilla Proceedings Project.

Hernández-Campoy, J. M. & Cutillas-Espinosa, J. A. (2013). The effects of public and individual language attitudes on intra-speaker variation: A case study of style-shifting. Multilingua 31(1), 79–101.

Herrero de Haro, A. (2017). Four mid back vowels in Eastern Andalusian Spanish. Zeitschrift fur romanische Philologie 133(1), 82–114.

Parkinson, S. (1988). Portuguese. In M. Harris & N. Vincent (eds.), The Romance Languages (131–169). New York: Oxford University Press.

Pollock, M. & Wheeler, J. A. (2022). Coda /s/ and intervocalic /d/ elision in Andalusia: The Formation of Susana Díaz's Regional Identity in Political Discourse. Language and Communication 87, 191–204.

Pollock, M. (2023). Toeing the Party Line: Indexing Party Identity through Dialectal Phonetic Features in Spanish Political Discourse. Languages 8(3), 196.

Rama, J., Zanotti, L., Turnbull-Dugarte, S. J. & Santana, A. (2021). VOX: The Rise of the Spanish Populist Radical Right. Routledge.

Ruiz-Peña, M. E. (2013). "Alma" o "arma", evidencia de la neutralización /l/ /r/ en la variedad dialectal andaluza de Sevilla. Unpublished MA Thesis, University of Western Ontario.

Silverstein, M. (2003). Indexical Order and The Dialectics of Sociolinguistic life. Language & Communication 23, 193–229.

Willis, E. W. (2006). Trill Variation in Dominican Spanish: An Acoustic Examination and Comparative Analysis. In Sagarra, N. & Toribio, A. J., Selected Proceedings of the 9th Hispanic Linguistics Symposium (121–131). Somerville, MA: Cascadilla Proceedings Project.

Zahler, S. & Daidone, D. (2014). A variationist account of trill /r/ usage in the Spanish of Málaga. IULC Working Papers, 14(2). Online.

# The 'evanescent /l/' in Venetan: a preliminary EMA study

Angelo Dian, Francesco Burroni
*University of Oxford, LMU Munich*
angelo.dian@ling-phil.ox.ac.uk, francesco.burroni@phonetik.uni-muenchen.de

## Abstract

This study is the first articulatory investigation of the so-called 'evanescent /l/' in Venetan. We propose that this sound be transcribed as a non-syllabic [ɛ̯] or [ə̯] in the variety of interest.

**Keywords:** EMA, Venetan, 'evanescent /l/'

## 1. Introduction

Venetan is an Italo-Romance language spoken in northeastern Italy alongside Italian, especially in the Veneto region, as well as in diaspora communities established through migration (Cordin, 2021). Though often referred to as a "dialect" of Italy due to its unofficial status (e.g., Maiden & Parry, 1997), Venetan is not a dialect of Italian: it developed independently from Proto-Romance and differs from Italian in phonology, morphosyntax, and lexicon.

A characteristic feature of several Venetan varieties—particularly those of Venice and central Veneto—is the elle evanescente ('evanescent /l/'): a process in which intervocalic /l/ weakens between non-front vowels and deletes when either flanking vowel is front. For example, /ˈpalo/ 'pole' has been transcribed as [ˈpae̯o], and /ˈbɛlo/ 'beautiful (m sg)' as [ˈbɛo].

Although widely discussed, existing accounts of this process remain mostly impressionistic and diverge in their proposed articulatory interpretations. Lepschy (1962) describes a "relaxed", non-lateral sound produced with a central groove in the raised tongue dorsum and lateral contact with the upper teeth. Zamboni (1974) and Tomasin (2010) transcribe the segment as a non-syllabic [e̯] with dorsopalatal articulation. Mafera & Roman (2006) report a fleeting [e]-like vocoid with tongue-tip raising but no alveolar contact. Other accounts describe a pre-velar approximant (Canepari, 1976, 1979) or fully vocalised realisations such as [e] or [i] (Belloni, 2009; Pellegrini, 1977). This variation in impressionistic descriptions highlights the need for empirical articulatory data. The present study offers the first such analysis using Electromagnetic Articulography (EMA).

## 2. Methods

Two male speakers of central Venetan, aged 24 (S1) and 36 (S2), from the neighbouring villages of Loria and Cassola (provinces of Treviso and Vicenza), were recorded. The corpus included disyllabic paroxytone Venetan words with /l/, /n/, and /j/ in intervocalic and initial positions within carrier phrases (see Appendix). EMA data and synchronized audio were collected using a Carstens AG501 system. Sensors were attached with oral adhesive. Four reference sensors (left/right mastoid, nasion, upper incisor) enabled head correction. Five tongue sensors were placed on the tip (TT), dorsum (TD), body (TB), and parasagittal (lateral) edges (PARAL, PARAR). Additional sensors tracked the jaw (below lower left incisor) and lips (upper/lower vermilion borders). Stimuli were presented via MATLAB; a bite plane was recorded post-session. Head movement was corrected offline using an orthogonal Procrustes transformation and spatially aligned based on the occlusion and upper incisor sensors.

## 3. Results and discussion

Figure 1 displays average /VCV/ vertical tongue trajectories over time for S1 (top) and S2 (bottom). On the left, TB movements are shown for /VlV/ intervals (red) and /VjV/ intervals (blue); on the right, TT movements are shown for /VlV/ (red) and /VnV/ (blue) intervals. The segments /j/ and /n/ serve as articulatory references for TB and TT, respectively, as their

primary constriction involves these tongue regions. Trajectories for TD, PARAL, and PARAR are omitted here, as they did not exhibit notably different movement patterns.

As previously proposed, /e_V/, /ɛ_V/ tokens (pooled in Figure 1) and /i_V/ tokens consistently showed elision of intervocalic /l/, with tongue gestures transitioning directly from the first to the second vowel or remaining static in the case of /i_i/. For TT (right-hand panels), /VlV/ intervals lack the TT raising characteristic of /VnV/, indicating the absence of an apical gesture for intervocalic /l/. For TB (left-hand panels), trajectories are consistently lower for /VlV/ than for /VjV/, suggesting a reduced degree of tongue body raising in the former. In both cases, differences in tongue movement patterns were also found statistically using linear mixed-effect regression applied to functional principal component scores.
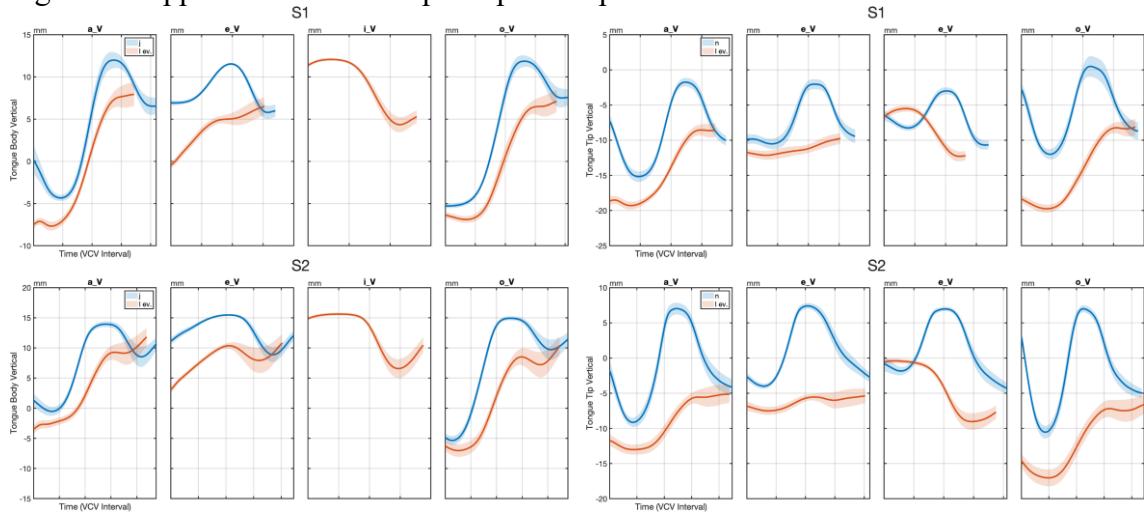


Figure1. Mean TB trajectories for /VlV/ and /VjV/ and TT trajectories for /VlV/ and /VnV/ by speaker.

Figure 2 illustrates average midsagittal trajectories (horizontal × vertical) of TB for /alV/ (red) and /ajV/ (blue) tokens. The /l/ gesture shows greater horizontal displacement, starting further back than /j/ but reaching a similarly front position. Vertically, however, /l/ does not reach the same height as /j/, indicating a reduced degree of dorsopalatal constriction. This articulatory difference is supported by the acoustic data: as shown in Figure 3, non-elided /l/ realisations occupy a region of the vowel space in the mid front area, though clearly lower than [e]. For the /ɔlo/ context there is a shift towards the mid central region.
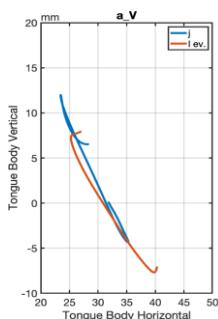


Figure 2. 2D TB trajectories of /j/ and /l/ from S1 (left = front)
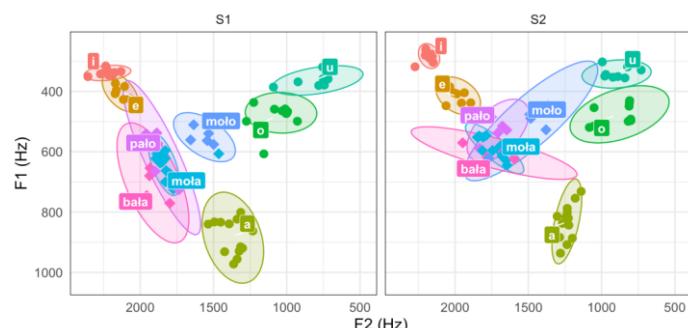
Figure 3. Acoustic vowel space for /i e a o u/ including non-elided intervocalic /l/ realisations (at peak F2 value) by speaker.

Overall, our results suggest that the elle evanescente, when not elided, is articulated as a low-mid front or a mid-central non-syllabic semivowel [ɛ̯] or [ə̯], depending on the vocalic context, characterised primarily by forward tongue movement with no targeted tongue-tip or lateral gestures. This pattern reflects a vocalisation process affecting intervocalic /l/ in syllable-onset

position. Future studies should consider patterns of variation of this sound across Venetan varieties, as previously reported impressionistically.

## 4. References

Belloni, Silvano (1991). Grammatica veneta. Padua: Esedra.

Canepari, Luciano (1976). The dialect of Venice. Journal of the International Phonetic Association, 6, 67-76.

Canepari, Luciano (1979). I suoni dialettali e il problema della loro trascrizione. In Manlio Cortelazzo (ed.), Guida ai dialetti veneti I (45-81). Padua: Cleup.

Cordin, Patrizia (2021). Italo-Romance: Venetan. In Oxford Research Encyclopedia of Linguistics. Oxford: Oxford University Press. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-724. Accessed on 31/05/2025.

Lepschy, Giulio C. (1962). Fonematica veneziana. L'Italia dialettale, 25, 1-22.

Mafera, Giovanni & Roman, Giovanni (2006). Saggi minimi di dialettologia veneta. Silea: Piazza.

Maiden, Martin & Parry, Mair (1997). The dialects of Italy. Oxford: Oxford University Press.

Pellegrini, Gian Battista (1977). Studi di dialettologia e filologia veneta. Pisa: Pacini

Tomasin, Lorenzo (2010). La cosiddetta "elle evanescente" del veneziano: fra dialettologia e storia linguistica. In Giovanni Ruffino & Mari D'Agostino (eds.), Storia della lingua italiana e dialettologia (729-751). Palermo: Centro di studi filologici e linguistici siciliani.

Zamboni, Alberto (1974). Veneto. In Manlio Cortelazzo (ed.), Profilo dei dialetti italiani. Pisa: Pacini.

Appendix

| V_V | /l/ | /j/ | /n/ | | Initial /l/ |
|-----|-----|-----|-----|---|-------------|
| a_a | bała 'ball' | maja 'sweater' | tana 'burrow' | | late 'milk' |
| a_e | bałe 'balls' | maje 'sweaters' | tane 'burrows' | | lenjo 'wood' |
| a_i | pałi 'poles' | taji 'cuts' | cani 'dogs' | | lima 'file' |
| a_o | pało 'pole' | tajo 'cut' | dano 'damage' | | longo 'long' |
| e_a | beła 'beautiful (f sg)' | sveja 'awake (f sg)' | sena 'dinner' | | luni 'Monday' |
| e_e | bełe 'beautiful (f pl)' | sveje 'awake (f pl)' | sene 'dinners' | | |
| e_i | bełi 'beautiful (m pl)' | sveji 'awake (m pl)' | beni 'goods' | | |
| e_o | beło 'beautiful (m sg)' | svejo 'awake (m sg)' | meno 'I drive (tr)' | | |
| i_a | fiła 'row' | | mina 'mine' | | |
| i_e | fiłe 'rows' | | mine 'mines' | | |
| i_i | fiłi 'threads' | | pini 'pine trees' | | |
| i_o | fiło 'thread' | | Dino 'given name' | | |
| o_a | moła 'soft (f sg)' | moja 'wet (f sg)' | nona 'nana' | | |
| o_e | mołe 'soft (f pl)' | moje 'wet (f pl)' | none 'nanas' | | |
| o_i | mołi 'soft (m pl)' | moji 'wet (m pl)' | noni 'grandparents' | | |
| o_o | moło 'soft (m sg)' | mojo 'wet (m sg)' | nono 'granddad' | | |

Table A. List of intervocalic (left) and phrase-initial (right) target words. 'e' and 'o' graphemes here may represent high-mid or low-mid vowels, depending on the word and/or Venetan variety, e.g. /ɔ/ for moło and /o/ for longo. The initial /l/ words shown are typically not produced with an evanescent /l/ in absolute initial position in central Venetan.

**Carrier phrases:**

Intervocalic: A digo ___ gran poco ('I say ___ very seldom').

Phrase initial: ___ zè quel che digo ('___ is what I say').

# Resistance to positional variability in Serbian laterals

Alexei Kochetov
*University of Toronto*
al.kochetov@utorono.ca

**Abstract**
This study uses electropalatography to investigate the production of Serbian laterals /l/ and /ʎ/ in utterance-initial, medial, and final positions. Data obtained from 4 speakers producing these sounds across three vowel contexts showed that the two consonants were clearly differentiated in articulatory measures and exhibited relatively minor positional differences – mainly between final and non-final positions and primarily for the less articulatorily constrained /l/. These results are different from the positional patterns previously reported for the English /l/, but largely similar to those found for Catalan /l/ and /ʎ/, highlighting the importance of phonemic contrasts in resisting positional variability.

**Keywords:** laterals, positional variability, electropalatography, Serbian

## 1. Introduction and background

In many English varieties, the lateral approximant /l/ is produced differently depending on the position within a word or utterance (Giles & Moll 1975; Sproat & Fujimura 1993, among others). In terms of the tongue-palate contact, the 'dark' allophone found in final position is characterized by a weaker tongue tip closure and absence of side contact throughout, compared to the 'clear' initial allophone. The lateral in medial-intervocalic position, on the other hand, tends to show intermediate patterns, indicative of a gradient continuum between the clear and dark endpoints, as illustrated in Fig. 1 for North American English (based on Colantoni et al. 2023a).

Articulatory studies of laterals in other languages have revealed less extensive positional variation, if any at all (e.g., Recasens 2012; Colantoni et al. 2023b), apparently related to the presence of a contrastive lateral in the inventory (e.g. /l/ vs. /ʎ/) and/or the role of articulatory constraints on the production of specific lateral closures (clear /l/ > dark /ɫ/ > /ʎ/). Thus, Recasens and Espinosa (2005) found relatively small positional variation for the Majorcan Catalan /l/ (=[ɫ]), with weaker closures medially-intervocalically than initially and finally. Even less positional variation was observed for the Catalan /ʎ/, which is produced with an extensive linguopalatal contact (Recasens & Espinosa 2006).

This study investigates positional variability in laterals /l/ vs. /ʎ/ in Serbian (e.g. lud 'crazy', jul 'July', ljubav 'love', žulj 'blister'), the language that has so far received limited attention in the articulatory phonetic literature (see Miletić 1933; Liker & Horga 2015 on Croatian).

## 2. Methodology

Four native Serbian speakers (3F) participated in the study, each provided with a custom-made artificial palate equipped with 62 electrodes. They produced 9 repetitions of 18 utterances with /l/ and /ʎ/ in 3 utterance positions (initial, medial, and final) and 3 vowel contexts (/a/, /i/, /u/). Electropalatographic recordings were made using the WinEPG system (Wrench et al. 2002) at a sampling rate of 100 Hz, continuously tracking the tongue-palate contact. Measurements, made at 5 points during the lateral closures, included the amount of contact across the palate (Q) and centre of gravity of the constriction location (CoG). /l/ was expected to show lower Q (less contact) and higher CoG (more anterior), with some positional differences expected for each consonant (cf. Recasens & Espinosa 2005, 2006). Lineal mixed effects models were used to evaluate consonant and position differences at each timepoint.

## 3. Results

Figure 2 illustrates the key linguopalatal patterns based on data from one of the speakers: /l/ was produced with a central denti-alveolar constriction and no side contact (i.e. the lowered tongue sides and dorsum); /ʎ/ was produced with an extensive closure spanning at least the half of the palate with a considerable side contact (i.e. the raised tongue dorsum). Notably, positional differences appeared to be minimal. These are further exemplified in Fig. 3 across 5 timepoints for the entire dataset: both laterals were produced with a somewhat lesser contact (lower Q) finally than initially/medially (yet with lesser differences for /ʎ/), and /l/ was produced with a somewhat more anterior constriction (higher CoG) finally than initially/medially. These differences were confirmed statistically at the $p<0.05$ level).

## 4. Discussion and conclusion

Overall, the results showed that Serbian laterals were clearly distinguished from each other by the amount of contact and constriction location (denti-alveolar vs. alveolopalatal). In terms of their linguopalatal profiles, Serbian /l/ was similar to the English dark /l/ allophone, while /ʎ/ showed much more contact than the English clear /l/ allophone (Figs. 1, 2). Furthermore, Serbian laterals were relatively stable across positions, exhibiting only minor differences between final and non-final variants (more so for /l/). These results are consistent with Recasens and Espinosa's (2005, 2006) findings for Majorcan Catalan laterals, with the exception of the medial position (initial, medial > final vs. initial > medial > final). This points to language-particular differences in the coordination of lateral gestures in onsets. The results also highlight the importance of phonemic contrasts in resisting positional variability.
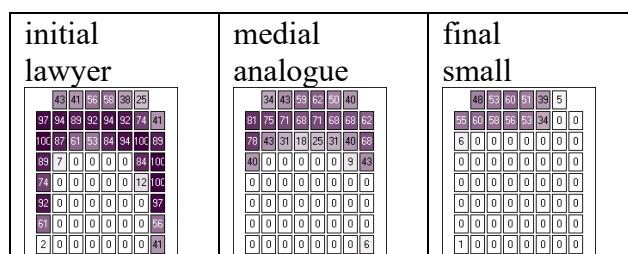


Figure 1 Average linguopalatal contact across the closure of the North American English /l/ produced by a speaker from (Colantoni et al. 2023a) in three utterance positions in selected words, based on 4 repetitions; shades of purple and numbers in the cells indicate the percentage of contact; cells in rows 1-2 = denti-alveolar, rows 3-4 = postalveolar, 5-8 = pre/medio/postpalatal; columns 1-2 and 7-8 = palate sides, columns 3-6 palate centre.
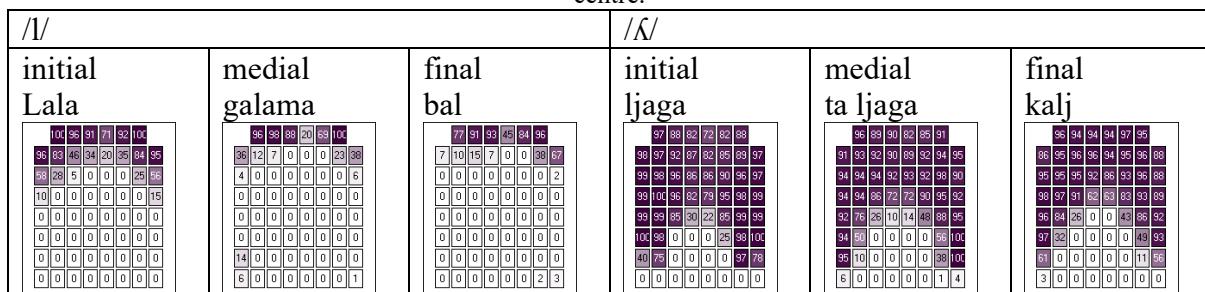


Figure 2 Average linguopalatal contact across the closures of the Serbian /l/ and /ʎ/, as produced by Speaker SR3 in three utterance positions in the context of /a/ (based on 9 repetitions).
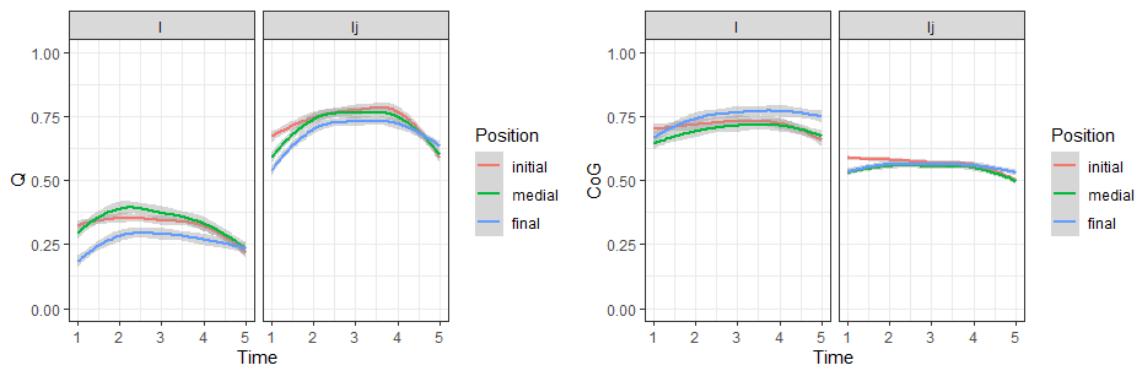
Figure 3 Overall amount of contact and CoG of constriction location for the Serbian /l/ and /ʎ/ ('lj') by utterance-position, taken at 5 timepoints (based on 640 tokens in total)

## 5. References

Colantoni, Laura, Alexei Kochetov, & Jeffrey Steele (2023a). Articulatory Insights into the L2 Acquisition of English /l/ allophony. Language and Speech, 1-33, 00238309231200629.

Colantoni, Laura, Alexei Kochetov, & Jeffrey Steele (2023). Syllable position effects with French and Spanish /l/. In Radek Skarnitzl & Jan Volín (Eds.), Proceedings of the 20th International Congress of Phonetic Sciences (pp. 937–941). Guarant International.

Gick, Bryan, Campbell, Fiona, Oh, Sunyoung & Tamburri-Watt, Linda (2006). Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. Journal of Phonetics, 34(1), 49–72.

Giles, Steven B. & Moll, Kenneth L. (1975). Cinefluorographic study of selected allophones of English /l/. Phonetica, 31, 206–227.

Liker, Marko & Horga, Damir (2015). Electropalatographic analysis of /ɲ/ and /ʎ/ in Croatian. In Proceedings of the International Congress of Phonetic Sciences, 5 pp.

Miletić, Branko (1933). Izgovor srpskohrvatskih glasova. Slovo.

Recasens, Daniel (2012). A cross-language acoustic study of initial and final allophones of /l/. Speech Communication, 54, 368–383.

Recasens, Daniel & Espinosa, Alina (2005). Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: Evidence from two Catalan dialects. Journal of the International Phonetic Association, 35(1), 1-25.

Recasens, Daniel & Espinosa, Alina (2006). Articulatory, positional and contextual characteristics of palatal consonants: Evidence from Majorcan Catalan. Journal of Phonetics, 34(3), 295-318.

Sproat, Richard & Fujimura, Osamu (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. Journal of Phonetics, 21, 291–311.

Wrench, Alan, Gibbon, Fiona, McNeill, Alison M. & Wood, Sara. 2002. An EPG therapy protocol for remediation and assessment of articulation disorders. In John H. L. Hansen & Brian Pellom (eds.), Proceedings of the 7th international conference on spoken language processing (pp. 965–968).

# CONTRIBUTORS

| Surname, Name | Affiliation | Country |
|---|---|---|
| Al-Tamimi, Jalal | Université Paris Cité France | France |
| Arróniz, Santiago | University of Nevada, Reno | USA |
| Audibert, Nicolas | Sorbonne Nouvelle University | France |
| Baran, Michał | University of Szczecin | Poland |
| Bedialauneta Txurruka, Izaro | University of Illinois at Urbana-Champaign | USA |
| Belando, Delia | University of Murcia | Spain |
| Bishop, Jason | City University of New York | USA |
| Boersma, Paul | University of Amsterdam | The Netherlands |
| Bolter, David | Humboldt-Universität zu Berlin | Germany |
| Bravo, Ana | University of Murcia | Spain |
| Broś, Karolina | University of Warsaw | Poland |
| Burroni, Francesco | Ludwig Maximilian University of Munich | Germany |
| Caillol, Coline | Université Paris Cité | France |
| Casquete de la Puente, Maialen | University of Illinois at Urbana-Champaign | USA |
| Čermák, Petr | Charles University | Czech Republic |
| Châu, Hồng Quang | Université Paris Cité | France |
| Chitoran, Ioana | Université Paris Cité | France |
| Colonna, Valentina | University of Granada/Laboratorio di Fonetica Sperimentale "Arturo Genre" | Spain/Italy |
| Cox, Felicity | Macquarie University | Australia |
| Crespo Quesada, Laura | Autonomous University of Madrid | Spain |
| Cruces Rodríguez, María | University of Málaga | Spain |
| Cruz Ortiz, Rocío | University of Granada | Spain |
| de Jong-Lendle, Gea | Philipps-Universität Marburg | Germany |
| De Luca, María | University of Córdoba | Spain |
| Di Caro, Alexandre | Université Jean Moulin Lyon 3 | France |
| Dian, Angelo | University of Oxford | UK |
| Docherty, Gerard | Griffith University | Australia |
| Dong, Suyuan | Sorbonne Nouvelle University | France |
| Du, Xiaojing | University of Cambridge | UK |
| Egurtzegi, Ander | Centre National de la Recherche Scientifique-IKER | France |
| Escamilla, Javier | University of the Basque Country | Spain |
| Ferragne, Emmanuel | Université Paris Cité | France |
| Ferrari, Eric | Université Paris Cité | France |
| Fisher, Rose | Michigan State University | USA |
| Foulkes, Paul | University of York | UK |
| Fröhlich, Andrea | Zurich Forensic Science Institute | Switzerland |
| González, Carolina | Florida State University | USA |
| Henneberger, Inke | Friedrich-Schiller-Universität Jena | Germany |
| Henriksen, Nicholas | University of Michigan | USA |
| Hererdia Mantis, María | University of Granada | Spain |
| Herrero de Haro, Alfredo | University of Granada | Spain |
| Hul, Nielson Sophann | University of Washington | USA |

| | | |
|---|---|---|
| Isgar, Gabrielle | Florida State University | USA |
| Jannedy, Stefanie | Leibniz-Centre General Linguistics | Germany |
| Jiang, Song | University of Toronto | Canada |
| Jiménez Vilches, Raúl | Autonomous University of Madrid | Spain |
| Khattab, Ghada | Newcastle University | UK |
| Ki, Mei-Ying | City University of New York | USA |
| Kochančikaitė, Renata | Lund University | Sweden |
| Kochetov, Alexei | University of Toronto | Canada |
| Krause, Peter A. | CSU Channel Islands | USA |
| Kye, Ted | University of Washington | USA |
| Leemann, Adrian | University of Bern | Switzerland |
| Lo Iacono, Federico | University of Turin/University of Genoa/Laboratorio di Fonetica Sperimentale "Arturo Genre" | Italy |
| Makarov, Yuri | University of Cambridge/Vinogradov Russian Language Institute | UK/Russia |
| Manjón-Cabeza Cruz, Antonio | University of Granada | Spain |
| Meseguer, Gemma | University of the Basque Country | Spain |
| Möller, Sophie | Philipps-Universität Marburg | Germany |
| Noelliste, Erin | University of Northern Colorado | USA |
| O'Neill, Paul | Ludwig-Maximilians-Universität | Germany |
| Oschkinat, Miriam | Leibniz-Centre General Linguistics | Germany |
| Penney, Joshua | Macquarie University | Australia |
| Percival, Maida | University of Oxford/University of Toronto | UK/Canada |
| Pérez Molina, Pablo | University of Málaga/Università di Torino | Spain/Italy |
| Pesantez, Alejandra | Universität Zürich | Switzerland |
| Pollock, Matthew | Louisiana State University Shreveport | USA |
| Ramírez Quesada, Estrella | University of Córdoba | Spain |
| Recasens, Daniel | Autonomous University of Barcelona | Spain |
| Roll, Mikael | Lund University | Sweden |
| Russo, Michela | University of Paris | France |
| Sabb, Kaitlyn | University of Michigan | USA |
| Sáez Rivera, Daniel M. | University of Granada | Spain |
| Saseendran, Ashima | The English and Foreign Languages University | India |
| Sedunova, Sofia | HSE University | Russia |
| Shekoufandeh, Golshid | University of Amsterdam | The Netherlands |
| Shepel, Ekaterina | Université Paris Cité/Saarland University | France/Germany |
| Skarnitzl, Radek | Charles University | Czech Republic |
| Sosinski, Marcin | University of Granada | Spain |
| Thekkadavan, Anupama | The English and Foreign Languages University | India |
| van den Bosch, Antal | University of Amsterdam | The Netherlands |
| Vida Castro, Matilde | University of Málaga | Spain |
| Waluch de la Torre, Waluch | University of Granada | Spain |
| Weiland, Werena | University of Bonn | Germany |
| Weirich, Melanie | Friedrich-Schiller-Universität Jena | Germany |
| White, Hannah | Macquarie University | Australia |
| Willis, Erik W. | Indiana University | USA |
| Yoo, Hiyon | Université Paris Cité | France |
| Yung-Ning, Jhuang | National Taiwan University | Taiwan |