Book of abstracts

SpeakVar 2026 Workshop

A workshop on intraspeaker and interspeaker variability

Budapest, 25–27 November 2026

Book of abstracts

SpeakVar 2026 Workshop

A workshop on intraspeaker and interspeaker variability

Budapest, 25–27 November 2026

Editors

Andrea Deme · Tekla Etelka Gráczi · Julianna Jankovics · Kornélia Juhász · Alexandra Markó · Sarolta Murányi · Katalin Pirsel · Veronika Svindt · Zsuzsa Szánthó

https://doi.org/10.18135/SpeakVar.2026

ISBN 978-615-6678-20-1

The workshop and the compilation of the book of abstracts were supported by the Lendület (Momentum) Programme of the Hungarian Academy of Sciences.

ELTE Research Centre for Linguistics

ELTE Nyelvtudományi Kutatóközpont

Budapest

2026



SpeakVar Workshop: A workshop on intraspeaker and interspeaker variability

25–27 November, 2026

ELTE Hungarian Research Centre for Linguistics
33 Benczúr Street, Budapest 1068, Hungary
Auditorium

Program

Wednesday, 25 November

9.00–9.15 Opening

9.15–10.15 Keynote speech – chair: Andrea DEME

Anne HERMES: Coordinative structures in speech: Linguistic, age- and disease-related factors

10.15–10.30 Coffee break

10.30–12.30 Oral session 1 – chair: Anne HERMES

Cécile FOUGERON: From chaos to patterns: Unraveling inter-speaker variability in articulation rate across adulthood

Bernadett DAM: Intraspeaker variability in speech characteristics before and after stroke-induced dysarthria

Ioannis PAPAKYRITSIS, Emilia MICHOU: Intraspeaker variability of nasalance scores: Data from Greek

12.30–14.00 Lunch break

14.00–15.30 Oral session 2 – chair: Cécile FOUGERON

Katelyn TAYLOR, Kirsty McDOUGALL: /s/ and /ʃ/ variation across gender-diverse speakers: Implications for forensic casework involving minority speaker populations

Teja REBERNIK, Katharina POLSTERER, Valentine LUCQUIAULT, Valerie QUERNER, Martijn WIELING, Anne HERMES: Effects of speech rate manipulation in Dutch word-initial clusters: age-related and individual variability

Andrea DEME, Zsuzsa SZÁNTHÓ, Tekla Etelka GRÁCZI: Production variability in segmental contrasts – Long and short vowels in speakers with Down syndrome

15.30–16.00 Coffee break

16.00–17.30 Poster session – chair: László FEJES

Tekla Etelka GRÁCZI, Sarolta MURÁNYI, Katalin PIRSEL, Zsuzsa SZÁNTHÓ, Ákos BLASKOVICS: Diadochokinetic tasks in Multiple Sclerosis: Acoustic and articulatory aspects

Kornélia JUHÁSZ, Katalin PIRSEL: Finding their voices: Is it more difficult for speakers diagnosed with Down syndrome to initiate voicing?

Beier LIU, Huba BARTOS, Katalin MÁDY: Choose your act by your sound: Prosody, speech acts and utterance-final particles in Mandarin Chinese

Marian OLIVEIRA, Lucrécia de AQUINO SANTOS, Deise SANTANA MAIA: Mid vowel variation in speakers with T21: Evidence from two Brazilian dialects

Lena SENGER, Peter UHRIG: Intra-speaker variation across the lifespan: Grammatical change in a high-density corpus of an American TV personality

Zsuzsa SZÁNTHÓ, Tekla Etelka GRÁCZI: Temporal characteristics of speech breathing in Multiple Sclerosis: A study of expiratory patterns and phonation time in Hungarian speakers

Alina TIMASHOVA: Prosodic variability at the morphological boundary: Stress patterns in Tatar-Russian intra-word code-switching

Marie-Theres WEIßGERBER, Oliver BUNK, Georg LOHFINK: Beyond the Yes-Bias? Multilingual sensitivity to schwa optionality across registers: An acceptability judgement study on variation and norm adherence

Thursday, 26 November

9.00–10.00 Keynote speech – chair: Tekla Etelka GRÁCZI

Stefano CORETTA: Bayesian multilevel modelling and estimation of uncertainty in speaker variation

10.00–10.15 Coffee break

10.15–12.15 Oral session 3 – chair: Ioannis PAPAKYRITSIS

Cécile FOUGERON, Nicolas AUDIBERT, Christine MEUNIER: Do speakers differ in the way they vary?

Ákos BLASKOVICS: Vowel-to-vowel isochrony across different suprasegmental and intervocalic configurations

Viktor KHARLAMOV, Daniel BRENNER, Benjamin V. TUCKER: Acoustic predictor importance and speaker-specific variation in Random Forest models of conversational English fricatives

Henrietta POKK, Péter RÁCZ: Variation and priming in the Hungarian 2SG.SBJV

12.15–13.30 Lunch break

13.30–15.30 Oral session 4 – chair: Viktor KHARLAMOV

Maria MENDES CANTONI, Adelino PINHEIRO SILVA: GLM-based dialect deconfounding for speaker verification across multiple Brazilian Portuguese varieties

Nhat-Nam NGUYEN, Pierre-André VUISSOZ, Yves LAPRIE: Toward speaker-independent acoustic-to-articulatory inversion guided by anatomical grid

Philipp BUECH, Teja REBERNIK, Louise BRUZZO, Anne HERMES: Effects of aging on vowel production in French: An acoustic and articulatory study

Reinhold GREISBACH, Anne HERMES: Typology of overlap in word-initial consonant clusters in German

15.30–16.00 Coffee break

16.00–18.00 Oral session 5 – chair: Kornélia JUHÁSZ

Xuesong LI: Is nonnative English speech more variable? Evidence from intra- and inter-speaker vowel variability

Zsuzsanna BÁRKÁNYI, Zoltán G. KISS: Multiple perceptual paradigms and learner variability in the perception of English aspiration by advanced L2 learners

Mónika VARGA: Intraspeaker and interspeaker variation in speech-related registers: Intensifiers in middle Hungarian

Giorgia ZANTEI: Melk or Melek? A maximum entropy grammar account for schwa epenthesis in Dutch

November 27. péntek

8.30–9.00 Regisztráció

9.00–10.30 üléselnök: MARKÓ Alexandra

TAR Éva, LÉNÁRT Zoltán: A beszéd, a beszédszervi mozgás és a felső végtagi mozgás variabilitásának sajátosságai perzisztáló beszédhanghibák esetén – egy pilótakutatás tapasztalatai

DEME Andrea, GOSZTOLYA Gábor, TÓTH László: A magyar magánhangzós hosszúsági kontraszt megvalósulása kisagyi stroke után jelentkező dizartriában – esettanulmány

FŐZŐ Eszter: A spontán beszédben megjelenő nyelvhasználati sajátosságok generációs különbségei

10.30–11.00 kávészünet

11.00–12.30 üléselnök: TAR Éva

KOHÁRI Anna: Jelölik-e zöngeminőséggel a prominenciát a magyar anyanyelvű beszélők?

TATÁR Zoltán: Automatikus életkorbecslés a Phonexia Speech Platform alkalmazásával

OSZKÓ Beatrix, VARGHA Fruzsina Sára, VÁRNAI Zsuzsa: Mondhatjuk-e, hogy „a magyarban" megkülönböztetünk hosszú és rövid felső nyelvállású magánhangzókat?

12.30–13.30 ebédszünet

13.30–15.30 üléselnök: VARGHA Fruzsina Sára

REBRUS Péter, TÓTH Szilárd: Variáció a szerbhorvát e-konjugációs igék jelen idejű alakjaiban

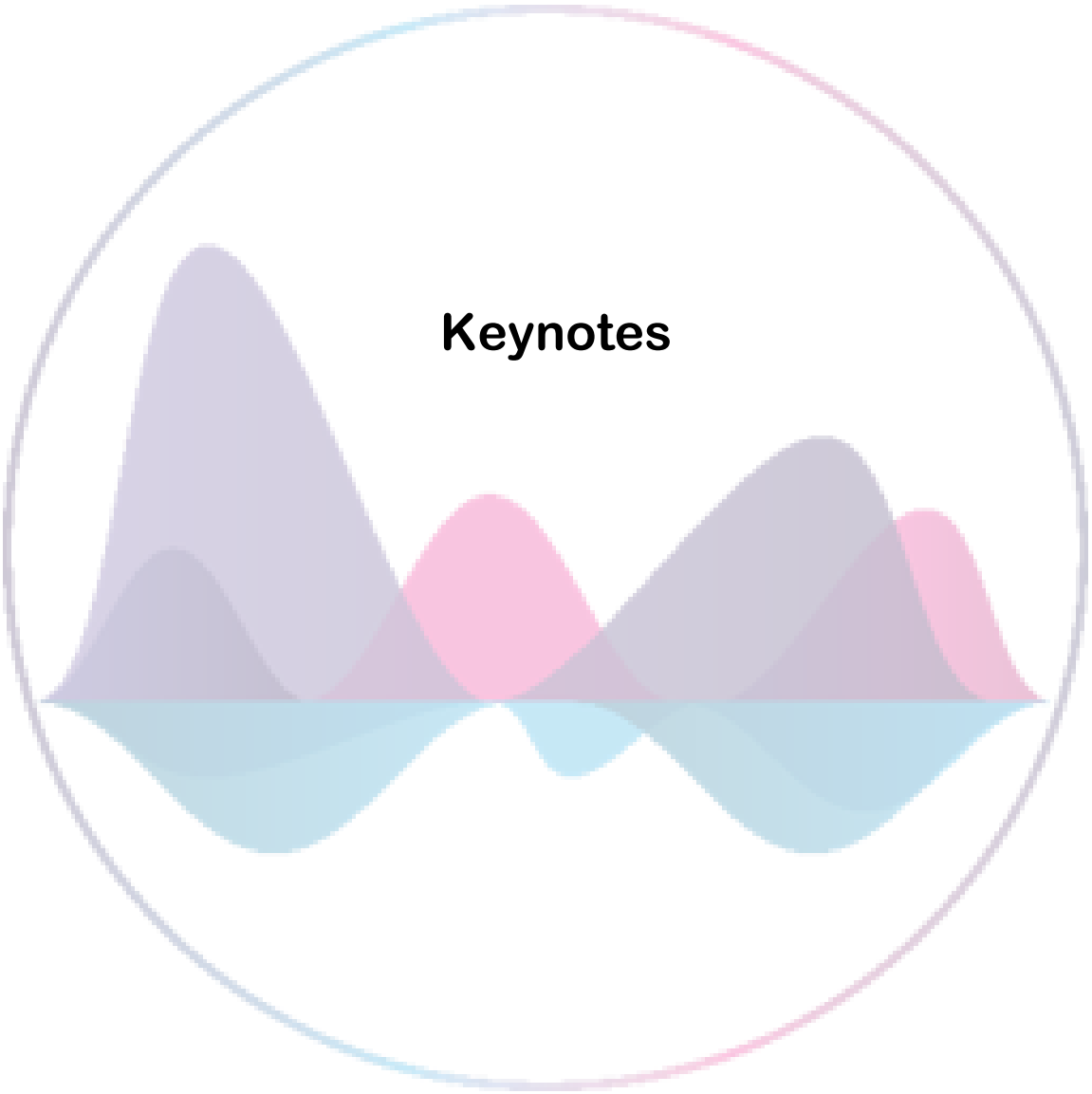
VARGA Zoltán, TATÁR Zoltán: A Phonexia beszédtechnológiai platform fejlődésének összehasonlító elemzése (2018–2026)

KOHÁRI Anna, CSÉBFALVI Fruzsina: Beszélők közötti és beszélőn belüli variabilitás a frázisvégi zöngeminőségben

MARKÓ Alexandra: Sztrájk a svájci nejlongyárban. Hat évtized alakváltozatai a filmhíradók korpuszában

15.30 Zárszó

Keynotes



BAYESIAN MULTILEVEL MODELLING AND ESTIMATION OF UNCERTAINTY IN SPEAKER VARIATION

Stefano Coretta

University of Edinburgh

s.coretta@ed.ac.uk

Background

Bayesian methods are being increasingly adopted in quantitative linguistics, although their uptake remains relatively limited. In this talk, I introduce some of the core principles of Bayesian inference and discuss their practical implications for the study of variation among speakers. In particular, I focus on Bayesian multilevel (or hierarchical/mixed-effects) models and on how they can be used not only to account for speaker-level variability, but also to make that variability itself an explicit target of inference.

Case study

The discussion centres on a case study of the voicing effect in Mancunian English, using vowel-duration data from 15 speakers from Manchester. Vowels tend to be longer before voiced than before voiceless consonants, but the magnitude of this durational difference need not be identical across speakers. This raises a broader methodological question: what does it mean to model speaker variation? Rather than treating the speakers in a sample as a collection of unrelated individuals, multilevel models allow us to conceptualise them as observations from a population of speakers and to model the distribution of individual differences within that population.

Using the voicing effect as a running example, I show how a regression model can be extended to include speaker-specific adjustments to both the intercept and the effect of consonant voicing. These adjustments are themselves modelled as arising from a population-level distribution. The parameters describing this distribution, the multilevel hyperparameters, provide information about the amount of between-speaker variation in baseline vowel duration and in the magnitude of the voicing effect, as well as about the relationship between these two sources of variation.

A Bayesian formulation makes uncertainty at each of these levels explicit. Rather than producing single estimates for the population-average voicing effect, individual speakers, or the degree of between-speaker variability, the model yields posterior distributions representing uncertainty about each quantity. I use these posterior distributions to illustrate the distinction between inference about an average effect across observations and inference about the distribution of effects across a population of speakers.

Conclusion

The broader aim of the talk is to encourage a shift in how speaker variation is approached statistically. In much current practice, multilevel structures primarily serve to account for non-independence while attention remains focused on population-level regression coefficients. I argue that the multilevel component can instead be substantively informative: its parameters allow us to ask directly how much speakers vary in an effect of interest, and with what degree of uncertainty. Characterising this variation can also provide information for the design of future studies, including prospective power analyses. Bayesian multilevel modelling therefore offers a coherent framework for moving from modelling linguistic effects in a sample of speakers to making explicit inferences about variation within a population of speakers.

COORDINATIVE STRUCTURES IN SPEECH: LINGUISTIC, AGE- AND DISEASE-RELATED FACTORS

Anne Hermes

Laboratoire de Phonétique et Phonologie (CNRS, Université Sorbonne Nouvelle), Paris, France

anne.hermes@sorbonne-nouvelle.fr

Introduction

Speech coordination can be understood as "orchestrating speech" (Hermes 2025). In this analogy, the brain acts as the conductor, governing the timing, planning, and execution of speech, while the articulators (e.g., tongue, lips, jaw) function as instruments, each contributing to the acoustic output. Phonological rules serve as the score, specifying how sounds are organized in terms of syllable structure, stress, and intonation. This orchestral view captures the complexity and multi-level coordination required for speech production and aligns with the concept of coordinative structures, i.e., self-organized collections of components functioning as coherent units (Bernstein 1967; Turvey 1990; Kelso 2014). Time and timing play a pivotal role in such structures, and this temporal patterning constitutes the fundamental basis of speech coordination (Kelso 2014; Saltzman & Kelso 1987). Uncovering coordinative structures across domains, i.e., articulatory coordination, segmental characteristics, speech breathing patterns, and phrasing, is therefore essential for understanding how linguistic systems operate, what variability reveals about their underlying organization (Browman & Goldstein 1992; Goldstein & Fowler 2003), and how external factors, such as biological (aging) or pathological (disease) ones, interact with and reshape them (Goldstein, Byrd & Saltzman 2006).

Methods

The talk draws on acoustic and kinematic data (articulation and speech breathing) collected across different languages (e.g., Italian, Polish, Tashlhiyt Amazigh, Georgian) and age groups (20–90 years), and from individuals with complete facial paralysis (Moebius syndrome). Coordination patterns are examined using Electromagnetic Articulography (EMA), which tracks both (i) the movements of speech articulators (lips, tongue, jaw) for intra- and inter-gestural timing, and (ii) chest movement, enabling measurement of inhalation and exhalation phases during speech breathing (Thies, Buech & Hermes 2026). To assess *linguistic* factors, we examine how coordination patterns in word-initial clusters vary depending on their phonological syllable status (i.e., branching vs. non-branching onsets). *Biological* factors are assessed by analyzing speech tasks (diadochokinesis, sentence production, text reading) in the acoustic, articulatory, and speech breathing domain across the lifespan. *Disease*-related effects are explored through kinematic data from individuals with Moebius syndrome (lip paralysis), focusing on compensatory articulatory strategies.

Results

The findings show that syllable structure is reflected in the temporal relationships between articulatory gestures, supporting the view of the syllable as a unit of articulatory organization across languages. Age-related changes affect the coordination of respiratory and articulatory systems, introducing variability and challenging the balance between stability and flexibility in speech patterns. Notably, aging does not uniformly impact the speech motor control system. In individuals with Moebius syndrome, intelligible speech is achieved despite the absence of labial function, through compensatory strategies such as enhanced lingual gestures and reorganization of articulatory timing.

Conclusion

Across all domains investigated, speech emerges not from isolated components but from coordinative structures whose elements self-organize into coherent functional units (Bernstein

1967; Turvey 1990; Kelso 2014). Variability within these structures is not mere noise but provides insight into their organization, revealing how linguistic systems differ and how different biological factors such as aging and disease reshape coordination. These findings underscore that speech coordination is a flexible, multi-level property of the speech motor system, which is constrained by linguistic structure, (re-)shaped by biological aging, and adaptable even under severe constraints such as complete facial paralysis.

Acknowledgements
A substantial part of the work on aging presented here has been carried out within the framework of the ANR-23-CE28-0017 INSPECT project supported by the French National Research Agency.

References

Bernstein, N. A. 1967. *The coordination and regulation of movement*s. Oxford: Pergamon Press.

Browman, C. P. – Goldstein, L. 1992. Articulatory phonology: An overview. *Phonetica* 49/3–4. 155–180.

Goldstein, L. – Byrd, D. – Saltzman, E. 2006. The role of vocal tract gestural action units in understanding the evolution of language. In: Arbib, M. (ed.): *Action to language via the mirror neuron system*. Cambridge: Cambridge University Press. 215–249.

Goldstein, L. – Fowler, C. A. 2003. Articulatory phonology: A phonology for public language use. In: Schiller, N. O. – Meyer, A. S. (eds.): *Phonetics and phonology in language comprehension and production: Differences and similarities*. Berlin, New York: De Gruyter Mouton. 159–208.

Hermes, A. unpublished (HDR, June 2025). *Orchestrating speech: The interplay of linguistic, age- and disease-related factors in speech coordination*. Université Paris Cité. Unpublished.

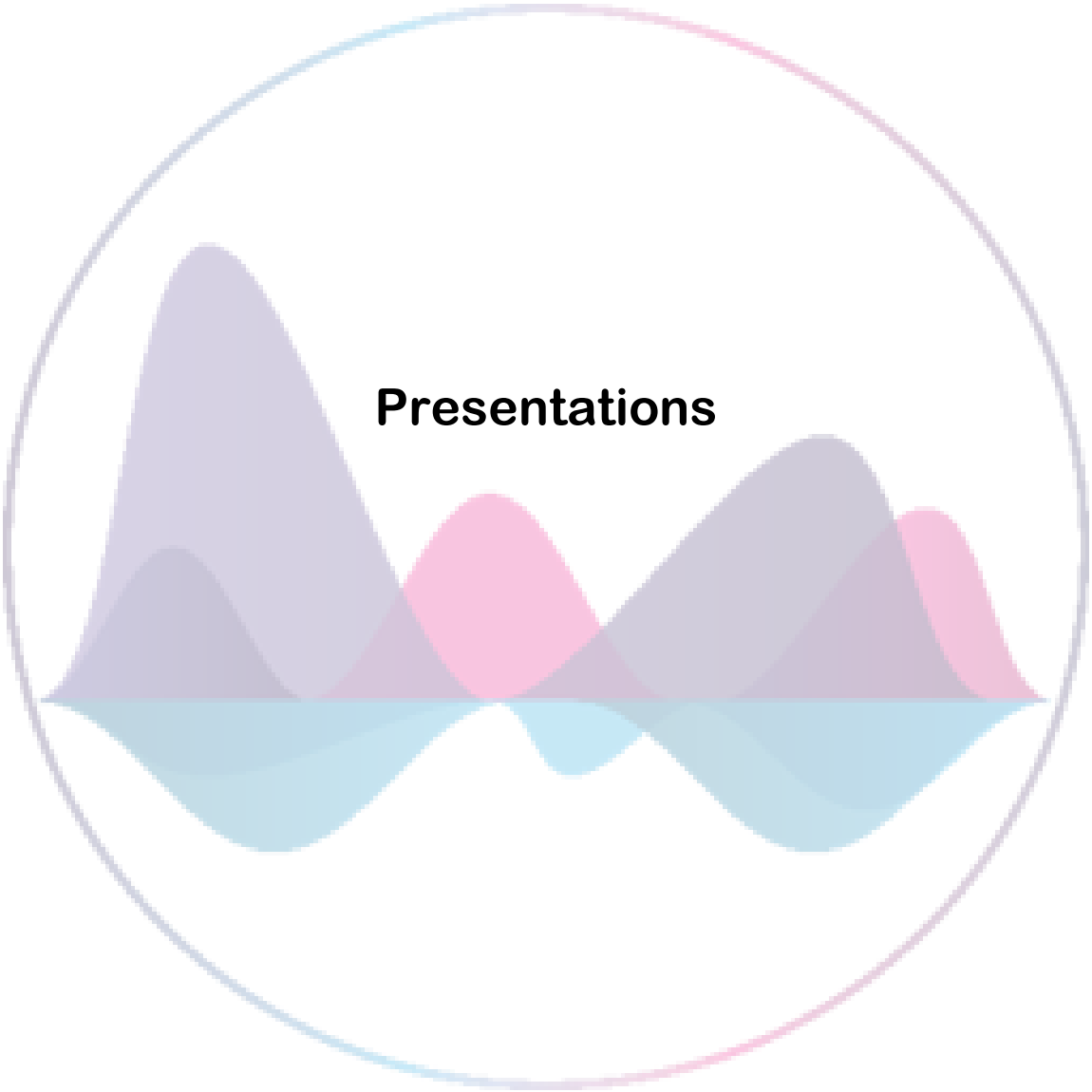
Kelso, J. A. S. 2014. The dynamic brain in action: Coordinative structures, criticality, and coordination dynamics. In: Plenz, D. – Niebur, E. (eds.): *Criticality in Neural Systems*, 67–104.

Saltzman, E. – Kelso, J. A. S. 1987. Skilled actions: A task-dynamic approach. *Psychological Review* 94/1. 84–106.

Thies, T. – Buech, P. – Hermes, A. 2026. A comparison of capturing speech breathing using respiratory inductive plethysmography and electromagnetic articulography. *PLoS One* 21/5. e0348123.

Turvey, M. T. 1990. Coordination. *American Psychologist* 45/8. 938–953.

Presentations



MULTIPLE PERCEPTUAL PARADIGMS AND LEARNER VARIABILITY IN THE PERCEPTION OF ENGLISH ASPIRATION BY ADVANCED L2 LEARNERS

Zsuzsanna Bárkányi[1] – Zoltán G. Kiss[2]

[1] The Open University, UK; ELTE Research Centre of Linguistics
[2] Eötvös Loránd University, English Linguistics Department

barkanyi.zsuzsa@nytud.hun-ren.hu, gkiss.zoltan@btk.elte.hu

A central question in second language speech learning is not simply whether learners can hear fine phonetic detail, but whether they use that detail to build target-like phonetic categories. English word-initial voiceless stops provide a useful test case because /p t k/ are typically realised with long-lag aspiration, whereas in "true voice" languages such as Hungarian (Beckman et al. 2013), voiceless stops are generally produced with short VOT. For L1 Hungarian learners, English aspiration is therefore neither fully unfamiliar nor phonologically contrastive in the same way as in English: it is a subtle, non-contrastive phonetic dimension that may be perceptible without necessarily being integrated into category structure.

This study investigates how advanced L1 Hungarian learners of English perceive English word-initial aspirated /p/. While previous research has extensively examined L2 VOT production and the contribution of VOT to perceived foreign accent, fewer studies have directly tested how L2 listeners *perceive* aspiration itself. Even fewer have compared multiple perceptual paradigms in the same learner group. The present exploratory study therefore asks whether advanced Hungarian learners are sensitive to VOT differences in English /p/, whether they show evidence of emerging English-like category organisation, and whether different perceptual tasks converge in measuring this sensitivity.

Sixteen Hungarian university students majoring in English language and literature participated in three perception experiments. All participants had at least B2-level English proficiency, and most had studied English phonetics and phonology. The experiments focused on word-initial pre-vocalic /p/ and used stimuli produced by a male native speaker of British English, subsequently manipulated in Praat. Three complementary tasks were administered in a fixed order. Experiment 1 used an accentedness-rating task: participants heard three versions of a short story differing only in the manipulated VOT values of voiceless stops and rated how native-like the speaker sounded on a five-point scale. Experiment 2 used a category-goodness rating task: participants rated how well the initial sound of six /p/-initial English words was pronounced along a six-step VOT continuum ranging from 3 ms to 78 ms. Experiment 3 used an AXB oddity task to test fine-grained auditory discrimination of VOT differences ranging from 15 ms to 75 ms. Cumulative link mixed models were used for the ordinal rating tasks, while a generalised logistic mixed-effects model was fitted to the AXB accuracy data.

The accentedness-rating task showed that longer VOT values increased perceived nativeness. The shortest VOT condition received lower ratings than the English-like long-VOT condition, although ratings were generally compressed toward the high end of the scale. Model-based probabilities indicated that the likelihood of a maximum rating increased from 0.56 for the short-VOT condition to 0.78 for the long-VOT condition. A significant positive linear trend confirmed that higher VOT values were associated with higher nativeness ratings, but individual participants differed substantially in how consistently they used the scale.

The category-goodness task revealed stronger and more systematic sensitivity to VOT. Ratings increased monotonically as VOT length increased, with low-VOT stimuli receiving poorer ratings and high-VOT stimuli receiving better ratings. All adjacent VOT levels differed significantly except the two longest values, 63 ms and 78 ms, indicating that ratings reached a ceiling around the English-like VOT range. The predicted probability of the highest rating rose sharply across the continuum, suggesting that learners had developed an emerging prototype for English /p/ aspiration near the 63 ms value. However, random slopes by participant improved model fit, showing substantial individual variation in category organisation, scale use, and the point at which ceiling effects appeared.

The AXB oddity task showed that participants could discriminate VOT differences above chance, but accuracy depended strongly on acoustic distance. A 15 ms difference was identified correctly in 60.6% of cases, while 60 ms and 75 ms differences yielded robust performance at 89.4% and 92.5%, respectively. The logistic mixed-effects model showed a highly significant positive effect of VOT difference: each additional 15 ms step increased the odds of correct identification by approximately 74%. Individual precision scores nevertheless varied widely, and performance in the AXB task did not straightforwardly predict performance in the rating tasks.

Taken together, the findings indicate that advanced L1 Hungarian learners of English are sensitive to aspiration in English word-initial /p/, but that this sensitivity is task-dependent. The AXB task captured low-level auditory precision, the category-goodness task reflected emerging phonetic category organisation, and the accentedness task measured a more holistic evaluative response. These dimensions were related but not reducible to a single perceptual ability. Some learners who discriminated small VOT differences accurately did not necessarily use that detail in category-goodness or accentedness judgments, while others showed graded rating patterns despite weaker discrimination at small acoustic distances. This dissociation suggests that access to phonetic detail does not automatically entail its use in phonological categorisation or accent evaluation.

Within the revised Speech Learning Model framework (Flege & Bohn, 2021), the results support a mixed interpretation. The weak differentiation in the holistic accentedness task suggests that learners may under-weight VOT as a linguistically salient cue, consistent with aspiration functioning as a “similar” category for Hungarian learners. At the same time, the category-goodness task points to partial restructuring toward an English-like long-lag prototype. Learners did not simply fail to perceive aspiration; rather, they showed gradient reorganisation with broad tolerance for intermediate VOT values and ceiling effects once English-like aspiration was reached. Overall, the study argues that L2 perception of non-contrastive phonetic detail is a multi-level phenomenon involving partially independent layers of auditory precision, category structure, and evaluative judgment.

References

Patel, R. – Grigos, M. I. 2006. Acoustic characterization of the question–statement contrast in 4, 7 and 11 year-old children. *Speech Communication* 48/10. 1308–1308.

Beckman, J. – Jessen, M. – Ringen, C. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49/2. 259–284.

Flege, J. E. – Bohn O.-S. 2021. The Revised Speech Learning Model (SLM r). In: Wayland, R. (ed.): *Second language speech learning: Theoretical and empirical progress*. Cambridge: Cambridge University Press. 3–83.

VOWEL-TO-VOWEL ISOCHRONY ACROSS DIFFERENT SUPRASEGMENTAL AND INTERVOCALIC CONFIGURATIONS

Ákos Blaskovics

ELTE Research Centre for Linguistics, BME Doctoral School of Psychology

blaskovics.akos@nytud.hu

Introduction

This study investigates vowel-to-vowel (V-to-V) isochrony in polysyllabic British English words: does the temporal interval between successive vowels remain constant regardless of varying segmental and prosodic contexts? The question is situated at the intersection of two different strands of research. Studies on isochrony of rhythmical units have long debated whether temporal regularity is exhibited at the level of stress feet (Jassem, Hill & Witten 1984), and Fowler (1983) proposed that perceptually isochronous syllables may reflect isochronously produced vowel nuclei. The second strand concerns the articulatory organisation of syllable onsets: the C-centre hypothesis of Browman & Goldstein (1988) holds that onset consonants are temporally centred on the following vowel, such that V-to-V timing could in principle remain stable despite varying onset complexity. The present study asks whether these two strands converge: do V-to-V time lags remain invariant across (1) simple (stop) vs. complex (stop+rhotic) second-syllable onsets, (2) voiced vs. voiceless intervocalic stops, and (3) within-foot vs. across-feet vowel pairs?

Methods

Twenty-two native speakers of British English varieties were recorded in an ultrasound tongue imaging (UTI) experiment. Participants produced 84 target words, matched for stress pattern, syllable count, and the first two vowel qualities (/æ–ə/, /ə–æ/, /æ–ɪ/, /ə–ɪ/), varying only in the place and voicing of the intervocalic stop and the presence of a following rhotic (e.g. *academic* ~ *acrobatic*, *Achilles* ~ *acrylic*). Words were produced in a carrier sentence across two repetitions. Midsagittal tongue contour data were recorded using the Articulate Assistant Advanced (AAA) software (Articulate Instruments Ltd. 2025). Based on the tongue contour data, which were extracted using DeepLabCut (DLC; Wrench & Balch-Tomes 2022), kinematic landmarks were obtained for each vowel: gestural onset, opening peak velocity, and maximum displacement. The duration between the respective landmarks of the first two vowels (time lags) was modelled as the outcome variable using Bayesian mixed-effects regression models (brms; Bürkner 2017) with lognormal response distribution, and with foot affiliation, onset complexity, and stop voicing as fixed-effect predictors and by-participant random intercepts.

Results

Kinematic results present a mixed picture. For gestural onset time lags, within-foot conditions show posteriors that cluster closely across onset complexity conditions, consistent with isochrony at this landmark. However, overall uncertainty is high: 80% credible intervals span 48–70 ms, and landmarks could be extracted for only ~22% of ultrasound tokens. The clearest departures from isochrony is shown by voiced within-foot stops (complex-onset lags longer by ~20–32 ms). Acoustic results are somewhat similar: V-to-V durations increase consistently with a complex onset and with a voiceless intervocalic stop across all foot affiliation and voicing conditions, with no sign of isochrony in any configuration.

Conclusion

The articulatory and acoustic results converge only in some of the conditions. Gestural onset may be the kinematic landmark at which V-to-V isochrony is most plausible, particularly within the stress foot, echoing Öhman's (1966) proposal of a continuous vocalic substrate underlying consonantal gestures. That acoustic durations show no such invariance is consistent with the view that the isochronous signal, if present, is obscured by the known effects of consonant

voicing on vowel duration (Coretta 2020) and C-centre-governed gestural overlap. These findings again highlight that acoustic measures alone are insufficient to evaluate articulatory isochrony hypotheses. Future work should address the low amount of obtained landmarks, and incorporate formal C-centre stability analysis (Marin & Pouplier 2010) to test syllabic affiliation assumptions directly.

References

Articulate Instruments Ltd. (2025). *Articulate assistant advanced user guide*. [Computer software manual]. Retrieved from https://www.articulateinstruments.com/downloads/ (Version 221.5.11)

Browman, C. P. – Goldstein, L. 1988. Some notes on syllable structure in articulatory phonology. Phonetica 45/2–4. 140–155.

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

Coretta, S. 2020. Vowel duration and consonant voicing: A production study. PhD dissertation, University of Manchester.

Fowler, C. A. 1983. Converging sources of evidence on spoken and perceived rhythms of speech: cyclic production of vowels in monosyllabic stress feet. Journal of Experimental Psychology: General 112/3. 386–412.

Jassem, W. – Hill, D. R. – Witten, I. H. 1984. Isochrony in English speech: its statistical validity and linguistic relevance. In: Gibbon, D. – Richter, H. (eds.): Intonation, Accent and Rhythm. Berlin: de Gruyter. 203–225.

Marin, S. – Pouplier, M. 2010. Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. Motor Control 14/3. 380–407.

Öhman, S. E. G. 1966. Coarticulation in VCV utterances: Spectrographic measurements. Journal of the Acoustical Society of America 39/1. 151–168.

Wrench, A. – Balch-Tomes, J. 2022. Beyond the edge: Markerless pose estimation of speech articulators from ultrasound and camera images using DeepLabCut. Sensors 22/3. 1133.

EFFECTS OF AGING ON VOWEL PRODUCTION IN FRENCH: AN ACOUSTIC AND ARTICULATORY STUDY

Philipp Buech[1] - Teja Rebernik[1] - Louise Bruzzo[1]- Anne Hermes[1]

[1]Laboratoire de Phonetique et Phonologie (CNRS & Universite Sorbonne Nouvelle), Paris, France

{philipp.buech, louise.wohmann, anne.hermes}@sorbonne-nouvelle.fr, teja.rebernik@cnrs.fr

l. Introduction

Aging is known to induce various physiological changes, also the speech motor system. The study of the effect of aging on speech has become a growing subject of investigation that sheds light on changes in various acoustic and articulatory features across the lifespan, such as vowel duration, fundamental frequency (f0), formants (F1, F2) or the internal structure of articulatory gestures. For example, the production of vowels is reported to be longer with increasing age (e.g., Albuquerque et al., 2020; Fougeron et al., 2021; Hermes et al., 2023; Ramig, 1983), but the impact of age on other vowel-related acoustic characteristics such as f0, F1 and F2 as well as the vowel space size may vary depending on other variables, such as speaker sex and/or vowel quality (e.g., Gahl & Baayen, 2019; Kong et al., 2024; Reubold et al., 2010; Tykalova et al., 2021). In terms of articulation, aging is reported to affect the internal structure of gestures especially regarding the tongue body in vocalic gestures and prosodic implementation (e.g., Hermes et al., 2018; Mücke et al., 2023). This raises the question whether aging has an influence not only on fully phonologized contrasts, but also on more subtle, linguistic subsystems. One example is word-level accentuation in French, that is part of the prosody of the language albeit not being distinctive (Fougeron & Smith, 1993; Vaissiere, 1991). We present here preliminary results on vowel productions contrasting younger speakers in their 20s and older speakers in their 70s, comparing accented and unaccented vowels in French. This production study focuses not only on vowel acoustics in multi-syllabic words, but also on their articulation using Electromagnetic Articulography (EMA).

2. Methods

The data presented here is part of a larger corpus including a variety of speech tasks collected in the ANR INSPECT project for the study of the impact of aging on speech. It consists of simultaneous acoustic and articulatory recordings using the EMA AGS01 (Carstens Medizinelektronik GmbH) with sensors attached to jaw, the upper and lower lips, the tongue tip, the tongue's midpoint and the tongue body. The data presented here is from two repetitions of a text reading task comprised of 416 words each. Six younger speakers in their twenties were selected (three females, three males – age range: 21-26) and five older speakers in their seventies (three females, two males – age range: 70 - 76). The acoustic recordings were automatically annotated using the Montreal Forced Aligner and vowel boundaries were manually corrected afterwards. Only oral vowels (/a, e, i, o u, y, J, E/) in multi-syllabic words with varying accentuation (unaccented, accented) in C<u>V</u>C, C<u>V</u>#, #<u>V</u>C contexts were selected. We extracted the vowel duration, as well as the mean f0, F1 and F2 (in Bark) and the horizontal and vertical tongue body position (in mm) within a window of S0% of the vowels' duration centered around the vowels' midpoint. The formants and tongue body positions were centered at the vowel space centroids. In addition, the acoustic and articulatory vowel space area (VSA) was calculated from the convex hull of the centroids of all vowels per speaker and accent type. A total of 4386 tokens went into the analysis.

3. Results

Table 1 summarizes the mean duration and f0 by sex, age group, and accent. As can be seen, accented vowels were longer in duration regardless of sex or age group. Overall, older speakers produced longer vowels than younger speakers, with the difference between age

groups was twice as high in accented vowels (female: 1S ms, male: 16 ms) than in unaccented vowels (female: 7 ms , male: 6 ms). Also, the difference in duration regarding accentuation was higher in older speakers (female: 26 ms, male: 30 ms) than in younger speakers (female: 19 ms, male: 20 ms). Regarding f0, accented vowels had a higher f0 than unaccented vowels in general (see Table 1). Older female speakers had a clearly lowered f0 than younger speakers in both accented and unaccented syllables, while f0 was similar in younger and older male speakers. The f0 difference regarding accentuation was similar between younger and older speakers (female: 16 Hz vs. 14 Hz; male: both 1S Hz).

Figure 1 displays the VSA in both acoustics and articulation by sex and accentuation, with the means and standard deviations given in the annotations (blue: younger speakers, red: older speakers). The acoustic and the articulatory VSA expands in accented vowels compared to unaccented vowels, regardless of sex and age group. In general, older speakers had a more expanded VSA than younger speakers in accented and unaccented vowels and in both acoustics and articulation, except for the articulatory VSA in older female speakers who displayed a 4 mm2 smaller but more variable VSA in unaccented vowels than younger female speakers.

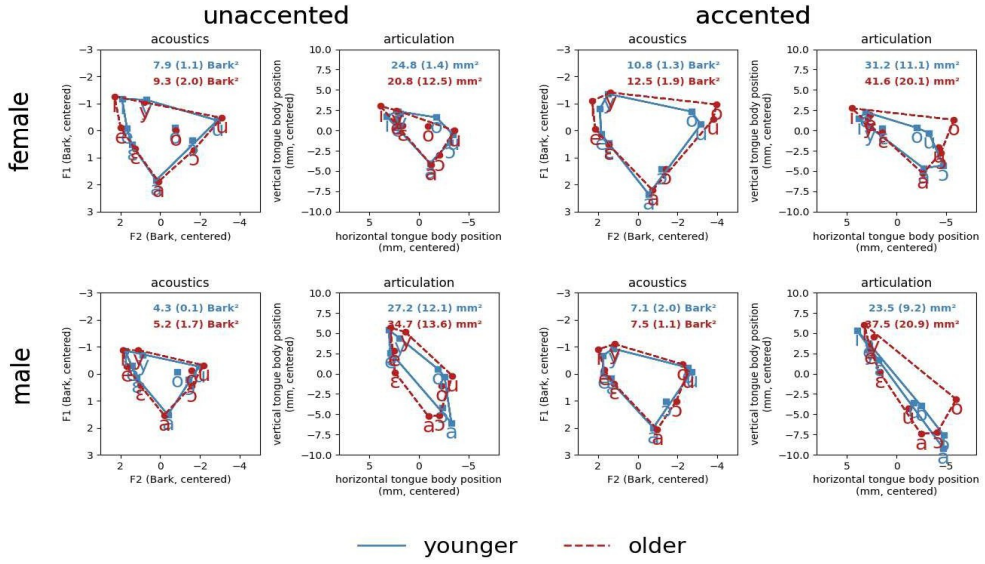


Figure l. *Acoustic (in Bark2) and articulatory (in mm2) vowel space area by sex, age group and accentuation. Means and standard deviations for the respective metrics are given within the plots for younger (blue) and older (red) speakers.*

4. Conclusion

These preliminary results confirm previous findings in the literature, i.e., that older speakers produce longer vowels than younger speakers (e.g., Fougeron et al., 2021; Mücke et al., 2023). In addition, the magnitude of the unaccented-accented duration distinction is more pronounced in older than in younger speakers. The f0 in vowels with respect to accentuation varies between the sexes: While productions of older female speakers tend to exhibit a decreased f0 compared to younger speakers, older male speakers displayed a similar f0 across accents. In line with e.g., Santiago and Mairano (2019), accented vowels had a larger acoustic VSA than unaccented vowels, and older speaker produced larger VSAs than younger ones (e.g., Fougeron et al., 2021). This finding is also reflected in the articulatory VSA. The present findings provide initial evidence that aging modulates not only vowel acoustics in general but also the acoustic and articulatory marking of prosodic accentuation in French. Older adults appear to amplify the accentual contrast in duration and vowel space, while sex-differentiated f0 patterns highlight the importance of controlling for speaker sex in aging

studies. These results contribute to our understanding of lifespan changes in the speech motor system and the resilience of prosodic subsystems in healthy aging. For the workshop, we will provide a comprehensive overview of the acoustics and articulation of unaccented and accented vowels in French in more detail and for more speakers, including speaker-variability and vowel- and accent-specific patterns.

		duration (ms)		f0 (Hz)	
sex	age group	unaccented	accented	unaccented	accented
female	younger	73 (19)	92 (33)	230 (30)	246 (48)
	older	80 (24)	106 (38)	204 (S2)	218 (63)
male	younger	6S (19)	8S (3S)	132 (23)	147 (32)
	older	71 (18)	101 (39)	134 (24)	149 (41)

Table l. *Mean duration and f0 (standard deviation in brackets) by sex, age group and accent.*

Acknowledgements
This work was funded by the ANR JCJC INSPECT project (ANR-23-CE28-0017).

References

Albuquerque, L. - Oliveira, C. - Teixeira, A. - Sa-Couto, P. - Figueiredo, D. (2020). A Comprehensive Analysis of Age and Gender Effects in European Portuguese Oral Vowels. *Journal of Voice, 37*, 143.e13-143.e29.

Fougeron, C. - Guitard-Ivent, F. - Delvaux, V. (2021). Multi-Dimensional Variation in Adult Speech as a Function of Age. *Languages, 6*, 1-2S.

Fougeron, C. - Smith, C. L. 1993. French. *Journal of the International Phonetic Association*, (23), 73-76.

Gahl, S. - Baayen, R. H. 2019. Twenty-eight years of vowels: Tracking phonetic variation through young to middle age adulthood. *Journal of Phonetics, 74*, 42-S4.

Hermes, A. - Audibert, N. - & Bourbon, A. 2023. *Age-related vowel variation in French*. In Skarnitzl, R. - Volin, S. (eds.), *Proceedings of the 20th International Congress of Phonetic Sciences - ICPhS 2023.* 204S-2049.

Hermes, A. - Mertens, J. - Mücke, D. 2018. Age-related effects on sensorimotor control of speech production. *Proc. Interspeech 2018*, 1S26-1S30.

Kong, C. - Long, X. - Liu, J. 2024. An acoustic study on age-related changes in vowel production of Chinese. *Journal of Phonetics*, *104*. 101324.

Mücke, D. - Thies, T. - Buech, P. - Furlani, N. - Schreen, J. - Hermes, A. 2023. Variability in space, compensation in time: Effects of aging on tongue body kinematics with respect to prosodic focus marking. In Skarnitzl, R. - Volin, S. (eds.), *Proceedings of the 20th International Congress of Phonetic Sciences - ICPhS 2023.* 203S-2039.

Ramig, L. A. 1983. Effects of physiological aging on speaking and reading rates. *Journal of Communication Disorders*, *16*(3), 217-226.

Reubold, U. - Harrington, J. - Kleber, F. 2010. Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication*, *52*, 638-6S1.

Santiago, F. - Mairano, P. 2019. Prosodic Effects on L2 French vowels: A corpus-based investigation. In: Calhoun, S. - Escudero, P. - Tabain, M. - Warren, P. (eds): *Proceedings of the 19th International Congress of Phonetic Sciences*, 1084-1088.

Tykalova, T. - Skrabal, D. - Boril, T. - Cmejla, R. - Volin, J. - Rusz, J. (2021). Effect of Ageing on Acoustic Characteristics of Voice Pitch and Formants in Czech Vowels. *Journal of Voice*, *35*(6), 931.e21-931.e33.

Vaissiere, J. 1991. Rhythm, accentuation and final lengthening in French. In J. Sundberg - L. Nord - R. Carlson (eds.), *Music, Language, Speech and Brain*. 108-120.

GLM-BASED DIALECT DECONFOUNDING FOR SPEAKER VERIFICATION ACROSS MULTIPLE BRAZILIAN PORTUGUESE VARIETIES

Maria Mendes Cantoni[1] – Adelino Pinheiro Silva[2]

[1] Universidade Federal de Minas Gerais
[2] Academia de Polícia de Minas Gerais

mmcantoni@gmail.com, adelinocpp@gmail.com

Introduction

Variability is inherent to speech and stems from both speaker-related factors (e.g., sociolinguistic and personal) and linguistic factors (e.g., phonetic-phonological and coarticulatory). Dialectal variation introduces changes in acoustic parameters that reflect geographic origin rather than individual identity (Docherty; Mendonza-Denton, 2012). In speech, multiple sources of variability are confounded (Rankinen & Jong 2021) and the modeling of speaker variability is an important development for automatic speaker recognition and verification systems (Wang et al. 2025). This study aims to model speaker variability in continuous speech across different Brazilian Portuguese (BP) dialects. In previous work on a mono-dialectal corpus, the residuals of a Generalized Linear Model (GLM) fitted for phonological context were shown to perform well on a speaker separation task (Silva & Cantoni, 2024; Cantoni & Silva, 2024). The present study extends this methodology to a multi-dialectal scenario to investigate the dialectal decoupling promoted by the GLM.

Methods

The dataset used totals 50 speakers, 5 dialectal varieties, and 36,524 vowel observations and was sampled from CEFALA-1 (Follador Neto, Silva & Yehia, 2019) and CORAA (Candido Junior et al., 2021), with automatic phonetic alignment (McAuliffe et al., 2017). Acoustic variables (AV) include duration, intensity, F_0, formants (F_1–F_4), spectral voice quality measures (H_1H_2, H_2H_4, H_4H_{2k}, $H_{2k}H_{5k}$, CPP, SHR), and coefficients of variation for F0–F4 (Lee, Keating & Kreiman, 2019). For each dependent variable, a Gaussian GLM was fitted with fixed effects for dialect and phonological context (diphthong, stress, vowel height, nasality, word position, preceding and following consonant). The premise is that the model residuals (GLM-RES) represent the voice production component unexplained by context and dialect, which captures, among other information, the individual speaker profile. For comparison, the space of normalised AV and a principal component projection (PCA) were also evaluated. Evaluation used within-speaker bootstrap with 20 iterations, with a 70-30 split into training and test sets per iteration. Bootstrap profiles served as sub-samples for computing within-speaker (SS) and between-speaker (DS) distances. A logistic regression calibrated distances into log-likelihood ratios (LLR) on the training set; evaluation metrics were empirical cross-entropy (ECE), equal error rate (EER), and log-likelihood cost (C_{LLR}). Speaker separability was further quantified by the DS/SS ratio using Mahalanobis distance, which controls for the feature covariance structure.

Results

Table 1 and Figure 1 summarise the performance of the representation spaces. GLM-RES achieved the lowest EER (4.2%), outperforming raw measures (AV: 5.2%) and PCA projection (9.4%). The C_{LLR} of GLM-RES (0.009) is higher than AV's (0.005), possibly reflecting artificial confidence extracted by AV from dialectal clusters in extreme prior regions of the ECE curve. The DS/SS ratio computed with Mahalanobis distance (Figure 1B) shows that GLM-RES has the highest speaker separability (median = 2.94), above AV (2.81) and PCA (2.42): when covariance structure is controlled, the residual space is the most effective at separating individual vocal identities.

Conclusions

Results show that in a multi-dialectal scenario the GLM successfully removed a portion of dialectal variance and improved individual speaker discrimination relative to raw measures. The reduction in EER combined with the higher DS/SS Mahalanobis ratio indicates that the gain is

not an artifact and persists when the geometry of the feature space is properly accounted for. This result is consistent with previous work (Silva, Cantoni, 2024) and generalises the methodology to a scenario of greater complexity. Future work should expand the regional coverage of the corpus.

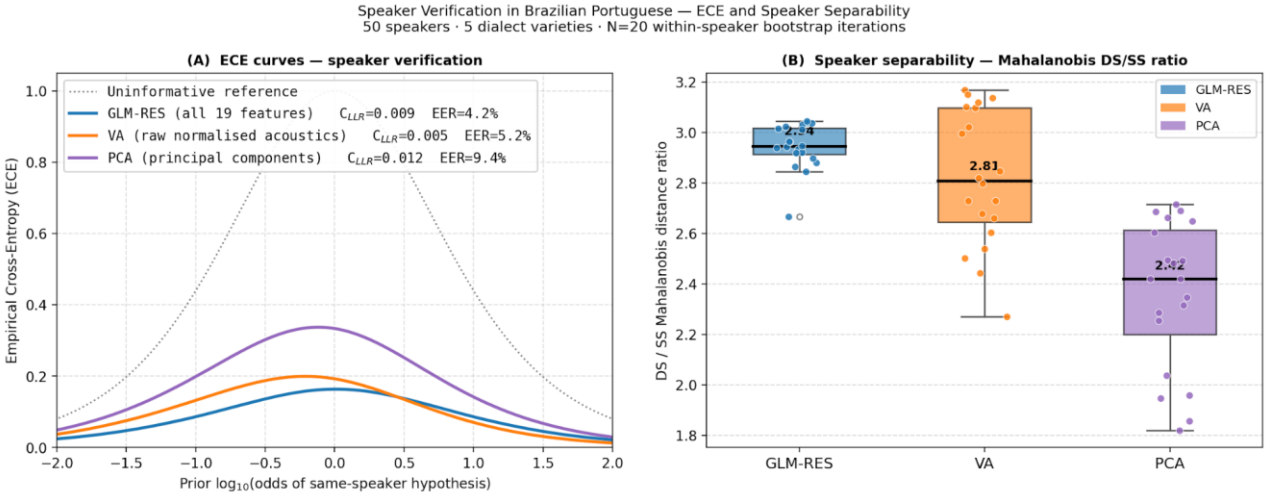


Figure 1. *A) ECE curves for acoustic representation spaces: prior log₁₀(odds) of the same-speaker hypothesis in the x-axis and normalised cross-entropy (lower = better) in the y-axis. (B) DS/SS Mahalanobis distance ratio.*

Space	Description	EER (%) C_{LLR} DS/SS Mahal.
GLM-RES	GLM residuals	4.2 0.009 2.94
AV	Raw normalised voice acoustics	5.2 0.005 2.81
PCA	PCA projection of AV	9.4 0.012 2.42

Table 1. *Speaker verification performance. EER (lower = better), C_{LLR} (lower = better), DS/SS Mahal. (higher = better).*

References

Candido Junior, A. et al. 2021. CORAA: a large corpus of spontaneous and prepared speech manually validated for speech recognition in Brazilian Portuguese. arXiv preprint arXiv: 2110.15731. doi: 10.48550/arXiv.2110.15731

Cantoni, M. M. – Silva, A. P. 2024. Articulatory and vocal speaker variability in connected speech. In: 13th International Seminar on Speech Production.

Lee, Y. – Keating, P. – Kreiman, J. 2019. Acoustic voice variation within and between speakers. The Journal of the Acoustical Society of America 146/3. 1568–1579. Docherty, G. – Mendoza-Denton, N. 2012. Speaker-Related Variation–Sociophonetic Factors. In: Cohn, A.; Fougeron, C.; Huffman, M. K. (eds). The Oxford Handbook of Laboratory Phonology.

McAuliffe, M. et al. 2017. Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In: Interspeech 2017. 498–502.

Neto, A. F. – Silva, A. P. – Yehia, H. C. 2019. Corpus CEFALA-1: base de dados audiovisual de locutores para estudos de biometria, fonética e fonologia. Revista de Estudos da Linguagem 27/1. 191–212.

Rankinen W. – de Jong K. 2021. The Entanglement of Dialectal Variation and Speaker Normalization. Language and Speech. 64/1, 181-202. doi: 10.1177/0023830920929379.

Silva, A. P. – Cantoni, M. M. 2024. Comparação Forense de Locutor por Modelos Lineares Generalizados de Variabilidade Articulatória e Vocal na Fala Encadeada. Avante: Revista Acadêmica da Polícia de Minas Gerais 7/1. doi: 10.70365/2764-0779.2024.101.

Wang, W. – Zhao, H. – Yang, Y. – Yang, Y. 2025 A review of speaker verification: Methods, network architectures, tasks and challenges. Engineering Applications of Artificial Intelligence, 162, 112351, 10.1016/j.engappai.2025.112351.

INTRASPEAKER VARIABILITY IN SPEECH CHARACTERISTICS BEFORE AND AFTER STROKE-INDUCED DYSARTHRIA

Bernadett Dam[1,2]

[1]Doctoral School of Linguistics, University of Szeged
[2]Institute of Applied Pedagogy, University of Szeged
dam.bernadett@szte.hu

1. Introduction

Dysarthria is a motor speech disorder that can affect multiple aspects of speech, ultimately resulting in reduced speech intelligibility (Duffy, 2013). It is always caused by neurological impairment, however, we see great variation in the affected components of speech based on the localization and severity of the impairment. For example, in stroke-induced dysarthria, speech disturbances can include imprecise articulation, changes in voice quality and pitch, resonatory problems, and changes in temporal characteristics (Ackermann et al., 2010). To give a detailed account of these changes in individual speakers, one could compare speech samples recorded before and after the impairment occurred.

This study aims to analyze several speech characteristics of a dysarthric speaker by comparing pre- and post-stroke speech samples. The characteristics analyzed include speech and articulation rate, vowel formant measures and F0 measures.

2. Methods

The participant is a male Hungarian native speaker in his fifties. The first speech sample was recorded pre-stroke, the second sample 19 months after the stroke, and a third sample was recorded two months later. All recordings contain the same eight sentences read aloud by the speaker. Measurements were conducted in Praat (Boersma, Weenink & Shchupak, 2026), manual labeling was followed by using scripts (Elvira-García, 2017; Lennes, 2017; Arantes, 2019).

Global speech and articulation rate was measured by manually labeling utterances, syllables and silences. The length of each segment was calculated using a script written by Lennes (2017). Formant frequency measurements were conducted on four different vowels: [ɒ], [ɛ], [i] and [u]. After labelling each vowel, the first three formant frequencies were extracted based on the mean of the mid 50% of the vowel (Elvira-García, 2017). Minimum and maximum frequencies were also calculated. F0 measures were conducted similarly, measuring the mean, median, standard deviation and range of F0 after manual labeling (Arantes, 2019).

3. Results

3.1. Speech and articulation rate

The global speech and articulation rates of the three samples are shown in Tabe 1. As expected, the participant's speech became notably slower post-stroke, his speech rate being more than twice as high pre-stroke. We can also notice a slight increase in speech tempo in the second dysarthric sample, which may suggest that his speech gradually increases over the course of recovery (see also Fig.1). A similar tendency can be seen in individual vowel durations. The average duration of all analyzed vowels in the healthy sample was 54.12 ms, while in the first and second dysarthric sample it was 101.82 ms and 99.27 ms, respectively.

	Speech rate	Articulation rate
healthy	6.05	6.08
dys_1	2.67	2.90
dys_2	2.86	3.29

Table 1. *Speech and articulation rate (syllables/second)*

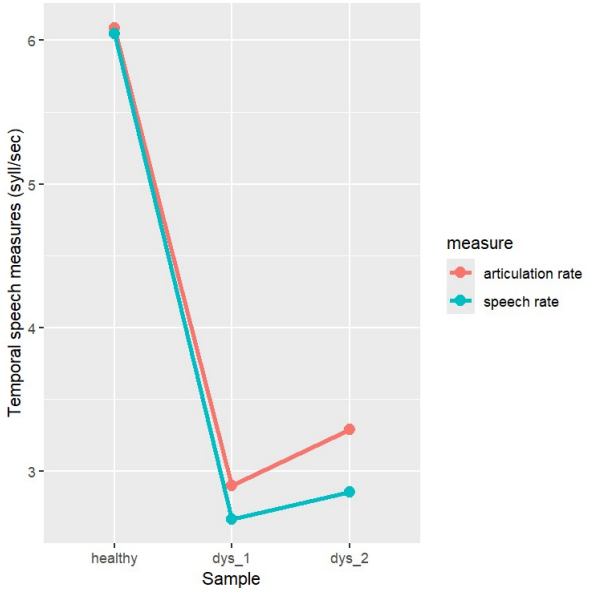


Figure 1. *Speech and articulation rate changes across the three speech samples*

3.2. Vowel formant measures

Figure 2 shows the average formant values of the selected vowels ([ɒ], [ɛ], [i] and [u]) across the three speech samples. Besides averages, standard deviation proved to be an informative measure. In both dysarthric samples, a higher standard deviation was measured in all four vowels across F1 and F2 (healthy sample: $SD_{F1} = 36.34$; dysarthric samples: $SD_{F1} = 53.68$, $SD_{F1} = 49.16$). This suggests less consistent articulation for these vowels which can be connected to changes in articulatory control in dysarthria.

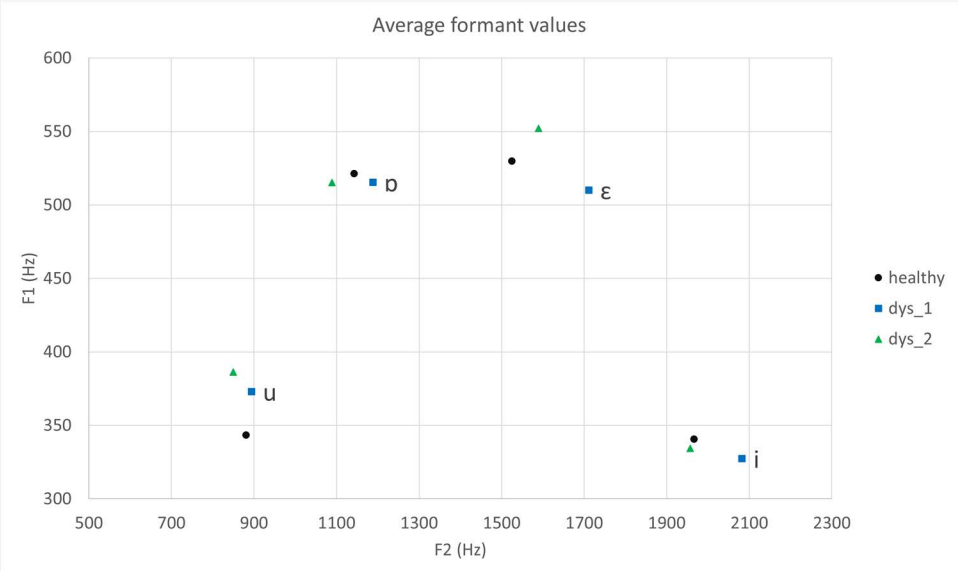


Figure 2. *Average formant values across the three speech samples*

3.3. F0 measures

As for F0, a higher average frequency was measured post-stroke, which is unexpected (healthy: 122.71 Hz, dysarthric: 136.56 Hz and 144.46 Hz). Standard deviation and range measures were more in line with expectations, both showing a higher value in the first dysarthric sample in comparison to the healthy one, and becoming slightly lower in the second dysarthric sample (healthy: SD = 23.3, range: 17.95; dysarthric: SD = 27.33 and 25.75, range: 21.74 and 19.19). This also reflects reduced control, in this case phonatory control.

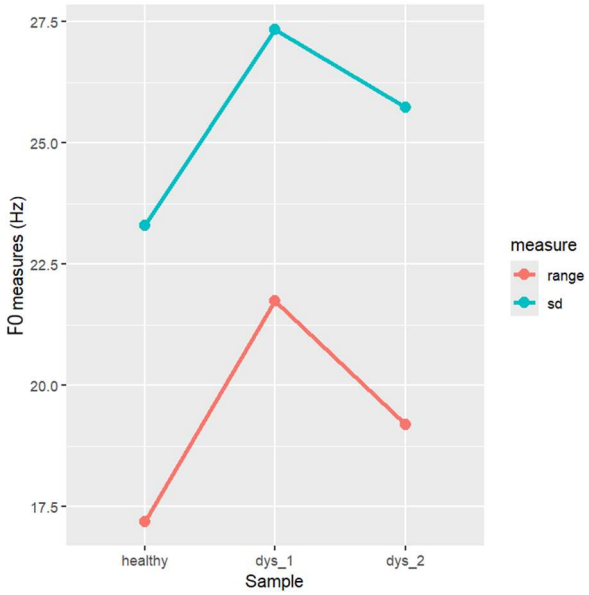


Figure 3. *Range and standard deviation of F0 across the three speech samples*

4. Conclusion

The selected measures illustrate how stroke-induced dysarthria and the following recovery (even in the span of two months) affect several speech characteristics. We see less consistency in formant structures and F0 measures in the dysarthric samples than in healthy speech. Reduced articulatory control is also illustrated by the reduced speech rate and articulation rate. Future work could explore how these measures change over a longer span of time, giving a more detailed account of the speaker's recovery.

References

Ackermann, H. – Hertrich, I. – Ziegler, W. 2010. Dysarthria. In J. S. Damico – N. Müller – M. J. Ball (eds.): *The Handbook of Language and Speech Disorders*. Chichester, West Sussex: Wiley-Blackwell. 362–390

Arantes, P. 2019. parantes/f0-measures: Version 1.1.1 (v1.1.1). Zenodo. Retrieved 22 May 2026 from https://doi.org/10.5281/zenodo.3470115

Boersma, P. – Weenink, D. – Shchupak, A. 2026. Praat: doing phonetics by computer [Computer program]. Version 6.4.67, retrieved 22 May 2026 from https://praat.org

Duffy, J. R. 2013. *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*. St. Louis, MO: Elsevier.

Elvira-García, W. 2017. wendyelviragarcia/vowels. GitHub. Retrieved 22 May 2026 from https://github.com/wendyelviragarcia/vowels

Lennes, M. 2017. SpeCT - Speech Corpus Toolkit for Praat (v1.0.0). First release on GitHub (1.0.0). Zenodo. Retrieved 22 May 2026 from https://doi.org/10.5281/zenodo.375923

A MAGYAR MAGÁNHANGZÓS HOSSZÚSÁGI KONTRASZT MEGVALÓSULÁSA KISAGYI STROKE UTÁN JELENTKEZŐ DIZARTRIÁBAN – ESETTANULMÁNY

Deme Andrea[1,2,3] – Gosztolya Gábor[4,5] – Tóth László[4]

[1]ELTE Eötvös Loránd Tudományegyetem; [2]MTA – ELTE NYTK Lendület Neurofonetikai Kutatócsoport; [3]ELTE Nyelvtudományi Kutatóközpont; [4]Szegedi Tudományegyetem; [5]HUN-REN–SZTE Mesterséges Intelligencia Kutatócsoport

deme.andrea@btk.elte.hu, ggabor@inf.u-szeged.hu, tothl@inf.u-szeged.hu

1. Bevezetés

A dizartria olyan artikulációs zavar, amelynek hátterében a beszéd motorikus megvalósítását érintő, változatos okokra visszavezethető agyi-idegi problémák állnak fenn ép mentális működések mellett (Horváth–Hirschberg 2013). A dizartria tünetei változatosak lehetnek. A logopédiai irodalom jellemzőkként írja le a lelassult (esetleg lokálisan felgyorsult) beszédet, az egyes beszédhangok hosszabb ejtését, az elmosódott artikulációt, a fonáció zavarait, valamint a légzés és a hangadás közti koordinációs zavarokat (Horváth–Hirschberg 2013). Fonetikai elemzések jellemzően azt vizsgálják, hogy mi az összefüggés a dizartriás beszéd érthetősége és a beszédben mért különféle paraméterek, például a magánhangzókat jellemző formánsok, illetve az ezekből deriválható mérőszámok, például a magánhangzótér mérete között. Ezek a kutatások javában azt mutatják, hogy a beszéd érthetősége (különböző mértékben, de) összefügg a magánhangzótér méretével, a magánhangzók közötti akusztikai távolságokkal, illetve a magánhangzó nyíltságát leképező F_1 formáns változatosságával úgy, hogy a kisebb magánhangzótér, illetve a nagyobb átfedések és változatosság nehezebben érthető beszéddel járnak együtt (Kim et al. 2011). Emellett arra is irányulnak elemzések, hogy milyen hatása lehet a mért beszédsajátosságokra és ezzel összefüggésben az érthetőségre a különféle beszédmódoknak, amelyek akár potenciális terápiákként is megjelenhetnek: ilyen a beszédsebesség szándékos lassítása vagy a hangerő fokozása. Az ehhez kapcsolódó eredmények szerint bizonyos mértékig mindkét manipuláció hathat az érthetőségre, ugyanakkor hatásuk a mérhető akusztikai paramétereken nem egyértelmű (eltérést találtak például a szklerózis-multiplex- és a Parkinson-kór-eredetű dizartriák beszédsajátosságai között) (Tjaden–Wilding 2004). Értelemszerűen ritka az, hogy egyazon beszélőtől elérhetők legyenek tipikus beszédállapotot tükröző (a dizartriás tünetek megjelenése előtt készült) és a dizartria tüneteit mutató beszédminták, így azt, hogy a dizartria pontosan hogyan is változtat ugyanazon beszélő beszédének objektíven mérhető sajátosságain, jóval kevésbé ismerjük. Ami a beszéd időviszonyait, és kifejezetten a magánhangzók időtartamát illeti, dizartriás tüneteket mutató beszédet nemben és életkorban illesztett kontroll beszélők ejtésével vetett össze Turner munkatársaival (1995). A szerzők egyfelől azt találták, hogy a dizartriás tüneteket mutató beszédben az átlagosan a kontroll beszélőkben találtaknál lassabb beszédsebesség mellett közel kétszer olyan hosszúak voltak a magánhangzók, mint a neurológiailag ép beszélők esetében. Továbbá azt is megfigyelték, hogy a sebesség gyorsításakor és lassításakor a dizartriás beszélők általánosságban jóval kisebb mértékben tudták változtatni a sebességet, illetve a beszédhangok tartamát, mint a kontrollcsoport tagjai.

A magyar nyelvben a magánhangzók tulajdonságai között megjelenik egy, a világ nyelveiben ritkább kontraszt, a hosszúság. Ennek a tulajdonságnak a mentén a magyar magánhangzó-rendszer 7 rövid-hosszú párra tagolódik: a zárt/felső nyelvállású és félig zárt/középső nyelvállású hangok között elsődlegesen a hangoztatás időtartama képzi a hangzásbeli különbséget (pl. [i]-[iː] vagy [u]-[uː]), míg a félig nyílt/alsó és nyílt/legalsó magánhangzópárok tagjai (az [ɒ] és [aː], valamint az [ɛ] és [eː]) hangszínükben is eltérnek (Siptár–Törkenczy 2007). A korábbi kutatások szerint az időtartambeli eltérés nagyobb az utóbbi két magánhangzócsoportban, mint az előbbi kettőben, továbbá a beszéd sebességének gyorsítása és lassítása másként hat a rövid és hosszú hangok tartamára, ugyanis a hosszúak nagyobb mértékben nyúlnak vagy rövidülnek, mint a rövidek (Magdics 1969). Ennek ellenére az újabb

vizsgálati eredmények azt mutatják, hogy a tipikus fejlődésű beszélők esetében a hosszúsági kontraszt még „extrém” körülmények között is megmarad, tehát az akusztikus jellemzők leképezik a hangok szembenállását: például egy speciális, erőteljesen lelassított beszédregiszterben, a dajkanyelvben (Deme et al. 2019), vagy a beszélő által természetes úton maximálisan felgyorsított beszédben (Deme et al. 2025). Mivel a dizartriás tüneteket mutató beszéd jellemzően lassabb, de a lassulás várhatóan nem minden beszédhangot érint ugyanolyan mértékben, továbbá a motoros kontroll is korlátozott dizartria esetén, felmerül a kérdés, hogy mi történik a nagymértékben az időtartammal kifejezett hosszúsági magánhangzókontraszttal dizartriás tüneteket mutató beszédben. Feltételezhető, hogy a dizartria tünetei csökkentik a rövid és hosszú beszédhangok elkülönítésének mértékét, illetve ez az elkülönítés nem rendszerszerű, de ezt a feltevést alátámasztó korábbi objektív mérési eredményekről nincs tudomásunk. Más okból beszédmotoros nehezítettséget mutató (Down Szindrómával diagnosztizált) beszélők esetében ugyanakkor nem látszik egyértelműen a kérdéses kontraszt kifejeződésének mérséklődése (Deme et al. 2026).

2. Módszerek

A fenti hipotézist egy olyan férfi beszélő ejtésében vizsgáljuk meg, akitől a dizartriás tünetek megjelenése előtti és utáni időszakból is rendelkezünk egymással összehasonlítható hangfelvételekkel, 7 év különbséggel. A felvételeken 200-200 felolvasott mondat szerepel úgy, hogy a második alkalommal kisagyi stroke, illetve a kisagy egy részének eltávolítása után (logopédiai szakvélemény alapján) súlyos fokú ataxiás dizartria tünetei jelentek meg a beszélő beszédében, (a szakvélemény szerint) lelassult, monoton és hipotóniás sajátosságokkal. A jelen vizsgálatban a beszéd általános jellemzésére elemezzük a szünetek számát és tartamát, illetve a mondatok és szavak ejtésének tartamát. A hosszúsági kontraszt elemzéséhez 4 magánhangzópárban, az [i]-[iː], [u]-[uː], [ɒ]-[aː] és [ɛ]-[eː] esetében vizsgáljuk meg 1) a beszédhangok tartamát, illetve a rövid és hosszú hangok elkülönítését a tartamok arányával, valamint 2) a beszédhangok hangszínét és hangszínbeli elkülönítését az F_1 és F_2 formáns frekvenciája, és az azok mentén számolt euklideszi távolságok szerint. A magánhangzók száma és kontextusa az egyes korpuszokban kiegyenlítetlen, illetve kontrollálatlan, de a két korpuszban az egyező mondatok miatt lényegében megegyezik.

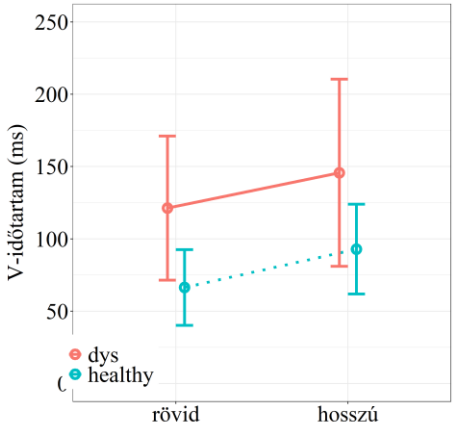
3. Eredmények

Előzetes eredményeink a beszédfolyamatosságra és beszédsebességre, illetve a magánhangzók tartamára vonatkoznak. A vizsgált beszélő a dizartriás tünetek megjelenése után több, mint másfélszer annyi szünetet produkált (58 vs. 136 db), melyek átlagos időtartama több, mint háromszorosa volt az első felvételeken megfigyelhetőknek, jelentősebb szórással (60,26±34,52 ms vs. 216, 31±102,60 ms), a beszéd tehát jelentősen töredezettebbé vált. Emellett az átlagos mondathosszok szintén másfélszeresükre nőttek megnövekedett szórással (3100,00±553,30 ms vs. 4953,46±5867,95 ms), az átlagos szóidőtartamok pedig kétszeresek lettek a második felvétel idejére, szintén jelentősebb szóródással (500,75±364,46 ms vs. 954,95±665,23 ms), tehát az artikuláció sebessége a felére csökkent. A magánhangzók időtartama átlagosan szintén közel kétszeresére nőtt a dizartriás tünetek megjelenésével, jelentősebb szórással (72,47±36,96 ms vs. 126,98±67,19 ms). Míg a rövid magánhangzók 182%-ukra nyúltak a dizartriás felvételre, addig a hosszúak 157%-ukra (1. ábra), tehát a hosszú és rövid magánhangzók tartamaránya lecsökkent (2. ábra), ugyanakkor semelyik párban sem érte el az 1-et (nem találtunk teljes produkciós semlegesedést). A legnagyobb elkülönítést mindkét felvételen az [ɒ]-[aː] és [u]-[uː], a legkisebbet az [ɛ]-[eː] és [i]-[iː] mutatta. A spektrális adatok elemzése jelenleg zajlik.

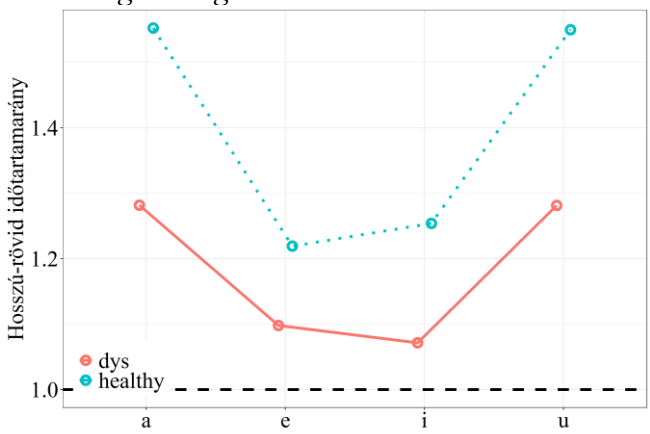
4. Következtetések

Az eredmények szerint a dizartria tüneteit mutató beszéd lassabb és töredezettebb volt, több és hosszabb szünettel, és az időzítés mentén minden paraméterben nagyobb változatosságot (szórást) mutatott. Az időtartamok elemzése alapján a dizartria tünetei csökkentették a rövid és hosszú beszédhangok elkülönítésének mértékét, amit korábban sem mesterségesen lelassított

neurotipikus beszédben (Deme et al. 2019, Deme et al. 2025), sem DS-val diagnosztizáltak beszédében (más eredetű motoros és/vagy fiziológiai nehezítettségben) (Deme et al. 2026) sem tapasztaltunk.



1. ábra *A magánhangzók időtartama a két beszédmintában*



2. ábra *A hosszú és rövid magánhangzók időtartamaránya a két beszédmintában*

Irodalom

Deme A. et al. 2019. A magánhangzós hosszúsági fonológiai kontraszt a dajkanyelvben a csecsemő életkorának függvényében. Beszédkutatás 2019. 221–242.

Deme A. et al. 2025. A mássalhangzók és magánhangzók időtartama, valamint a magánhangzós fonológiai hosszúsági kontraszt kényelmes tempójú és felgyorsított beszédben. Alkalmazott Nyelvtudomány 25: 1-20.

Deme A. – Szánthó Zs. – Gráczi T. E. 2026. A magyar rövid és hosszú magánhangzók időtartama Down-szindrómával diagnosztizált beszélőkben. Előadás: MFFLT Kongreszus, Budapest, 2026 május 30.

Horváth Sz. – Hirschberg J. 2013. Diszartria/diszartrofónia (dysarthria/dysartrophonia). In Hirschberg Jenő – Hacki Tamás – Mészáros Krisztina (szerk.): Foniátria és társtudományok. II. kötet. Budapest: ELTE Eötvös Kiadó. 80–86.

Kim, H. et al. 2011. Vowel contrast and speech intelligibility in dysarthria. Folia Phoniatria et Logopaedica 63: 187–194.

Tjaden, K.– Wilding, G. E. 2004. Rate and loudness manipulations in dysarthria: acoustic and perceptual findings. Journnal of Speech, Language, and Hearing Research. Vol 47. 766–783

Turner, Greg S. – Tjaden, Kris – Weismer, Gary 1995. The inflence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. Journal of Speech and Hearing Research, Volume 38, 1001-1013.

Siptár, P., Törkenczy, M. 2007. The Phonology of Hungarian. Oxford University Press.

Köszönetnyilvánítás

A kutatás a Bolyai János Kutatási Ösztöndíj szakmai támogatásával készült.

PRODUCTION VARIABILITY IN SEGMENTAL CONTRASTS – LONG AND SHORT VOWELS IN SPEAKERS WITH DOWN SYNDROME

Andrea Deme[1,2,3] – Zsuzsa Szánthó [1,2,3] – Tekla Etelka Gráczi[2,3]

[1] ELTE Eötvös Loránd University; [2] MTA–HUN/ELTE-REN NYTK Lendület "Momentum" Neurophonetics Research Group; [3] ELTE Hungarian Research Centre for Linguistics

deme.andrea@btk.elte.hu, szantho.zsuzsa@btk.elte.hu, graczi.tekla.etelka@nytud.hu

Introduction

Down Syndrome (DS) is a genetic condition characterized by atypical intellectual, neurological, anatomical, and physiological development (Antonarikis et al. 2021). Common features include mild cognitive impairment, midface hypoplasia combined with an average sized tongue, and low muscle tone in the orofacial muscles (Bunton & Leddy 2011). These factors can affect speech production, often resulting in reduced intelligibility of DS speech, as described in clinical literature based on perceptual evaluations and screenings (e.g., van Borsel 1996). However, relatively little research has examined instrumentally measurable acoustic parameters of DS speech that function as linguistically relevant perceptual cues, particularly for vowels. Studies on English showed that vowels in DS speech tend to be longer, and the vowel space is reduced—that is, vowels are less differentiated than in typically developing speakers (e.g., Bunton & Leddy 2011; Vorperian et al. 2023). In contrast, findings from other languages, such as French, have reported greater vowel differentiation in speakers with DS (Rochet-Capellan & Dohen 2015).

Hungarian features a phonemic vowel length (quantity) contrast: the 14 vowels of the system form 7 long and short pairs. Close and close-mid vowels (e.g., [i]-[iː] and [u]-[uː]), differ primarily in duration, while open-mid and open vowels ([ɒ]-[aː] and [ɛ]-[eː]), differ in both duration and quality (Siptár & Törkenczy 2007). Earlier studies showed that this length contrast is more marked in duration in low vowels and that speech rate affects long and short vowels differently: in fast speech, long vowels undergo greater durational reduction than short vowels (Magdics 1969). Recent research demonstrated that despite these asymmetries, the length contrast is preserved even under "extreme" conditions – both in slow/exaggerated (infant directed) speech (Deme et al. 2019), and in fast speech (Deme et al. 2024) in typically developing speakers. Our preliminary results on speech of Hungarian speakers with DS revealed that their speech rate was 1.2-1.5 times slower than that of controls (words were produced 1.2 times longer and vowels 1.5 times longer). Although duration ratios did not fall below 1 in any vowel pair (which would indicate complete production neutralization of the contrast), all vowels were uniformly longer in DS speech. As a result, durational differentiation of [i]-[iː], [u]-[uː], and [ɛ]-[eː] was comparable across speaker groups, while [ɒ]-[aː] showed reduced differentiation in speakers with DS (Deme et al. 2026) (Figure 1.).

The present study further investigates whether the vowel length contrast is acoustically expressed in DS speech, where reduced motor control and anatomical differences may hinder its production. We extend our previous analyses on duration by investigating spectral differentiation of long and short vowels. We hypothesize that both durational and spectral correlates of the vowel length opposition distinguish vowel pairs to a lesser extent in DS speech.

Methods

We analyze 4 long–short vowel pairs in 6 speakers with DS (aged 23–40) and 10 control speakers (aged 19–22), using 4 linguistic conditions (real words vs. nonwords × in minimal pair vs. not in minimal pair). In the present analysis, we do not investigate the effect of linguistic condition; however, it should be noted that vowel pairs cannot be directly compared, as systematically different consonantal contexts were used to construct the 4 conditions. We measured vowel duration and calculated duration ratios of long–short vowel pairs to assess durational correlates of the vowel length contrast in production. In addition, F_1 and F_2 frequencies were estimated in the middle 20% of the same vocalic segments and (after manual

correction) Euclidean distances of the long and short pairs were calculated to evaluate spectral differentiation. Statistical analyses were conducted using linear mixed effects modelling, Tukey's post hoc tests, and one sample *t*-tests.

Results

Preliminary results indicate that all vowels were uniformly longer in DS speech. Durational differentiation of [i]-[iː], [u]-[uː], and [ɛ]-[eː] were similar across the two speaker groups, whereas [ɒ]-[aː] showed reduced differentiation in speakers with DS (Figure 1.). All duration ratios exceeded 1, indicating that, on the group level, complete durational neutralization did not occur in production in either group. However, both vowel duration and duration ratios exhibited greater variability in the DS group, due to substantial amount of interspeaker variability in this group. As for formant measures, greater variability was also observed in speakers with DS. Nonetheless, vowel spaces constructed from all tokens in each group did not show clear signs of reduction in the DS group (Figure 2.), and variability appeared greater for back vowels that for front vowels. Analysis of Euclidean distances of the long–short vowel pairs revealed no consistent reduction of spectral differentiation in DS speech (Figure 3). The [ɒ]–[aː] pair (in $C_{alveolar}_C_{alveolar}$ context) showed comparable distances across groups, similar to and [u] and [uː] (in $C_{alveolar}_C_{velar}$ context). In contrast, [ɛ] and [eː] (in $C_{velar}_C_{alveolar}$ context) were significantly more differentiated in control speakers than in speakers with DS. Finally, contrary to expectations, [i]–[iː] (in $_C_{alveolar}C_{alveolar}$ context) exhibited significantly greater differentiation in the DS group than in the control group.

Discussion

To summarize, we found no across the board reduction of the vowel length contrast in speakers with DS; however, durational and spectral correlates of the vowel length contrast exhibited different patterns. Durational distinction appeared least stable in [ɒ]–[aː], where the length contrast is expressed by both timbral and durational cues. Spectrally, speakers with DS showed no reduction in the length contrast; in fact an enhanced contrast was observed for high vowels, where traditional accounts suggest primarily durational encoding. Overall, these findings do not indicate a clear group-level deficit in the production of the vowel length contrast in speakers with DS. However, individual data (not reported here) suggest substantial variability within the DS group. We argue that fine-grained phonetic analyses of this kind can contribute to a more nuanced understanding of DS speech production and may also provide valuable insights for clinical speech therapy.

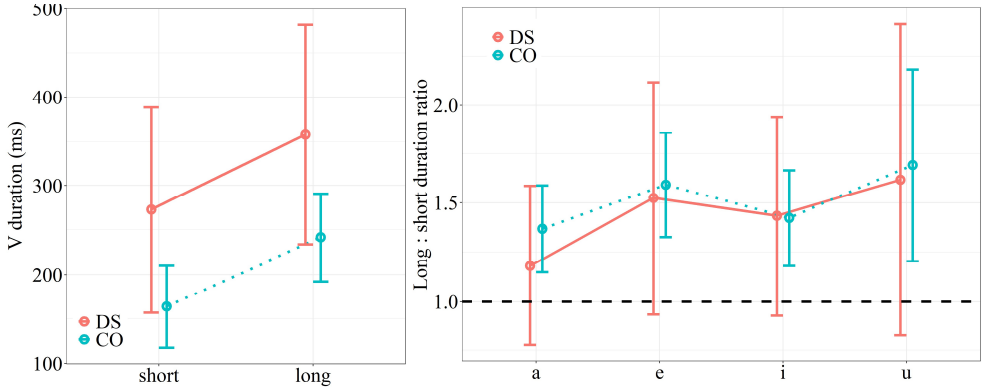


Figure 1. *Duration (left) and duration ratio of long and short vowels (right) in speakers with DS and in the control group*

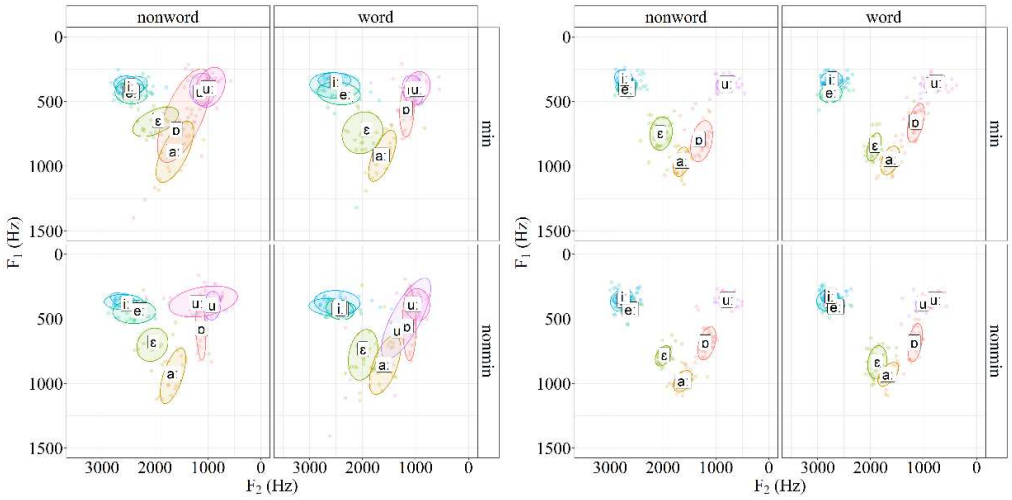


Figure 2. *Vowels in the $F_1 \times F_2$ space in speakers with DS (left) and in the control group (right)*

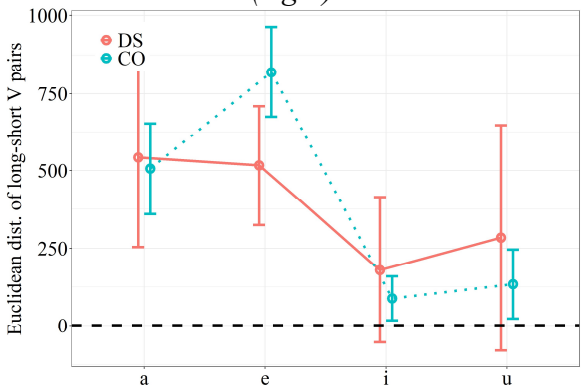


Figure 3. *Euclidean distances of long and short vowels in speakers with DS and in the control group*

References

Antonarakis, S. E. et al. 2021. Down syndrome. Nat Rev Dis Primers. 6/1, 9. DOI:10.1038/s41572-019-0143-7.

Deme, A. et al. 2024. German and Hungarian long and short vowels in fast speech. ISAPh 2024, Tartu, Estonia. 5–10. DOI: 0.21437/ISAPh.2024-2

Deme A. et al. 2026. A magyar rövid és hosszú magánhangzók időtartama Down-szindrómával diagnosztizált beszélőkben [Hungarian long and short vowels duration in speakers with DS]. Talk presented at MFFLT Congress, Budapest, 30. May 2026.

Magdics, K. 1969. The duration of Hungarian speech sounds in calm and fast speech. (in Hungarian). Nytud. Ért. 67. 45–63.

van Borsel, J. 1996. Articulation in Down's syndrome adolescents and adults. European Journal of Disorders of Communication 31. 415–444.

Rochet-Capellan, A., & Dohen, M. 2015. Acoustic characterization of vowel production by young adults with Down syndrome. In M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. Scobbie (Eds.), 18th ICPhS, 1–5.

Siptár, P., Törkenczy, M. 2007. The Phonology of Hungarian. Oxford University Press.

Vorperian et al. 2023 Vowel production in children and adults with down syndrome: Fundamental and formant frequencies of the corner vowels. Journal of Speech, Language, and Hearing Research 66/4. Pages 1208–1239.

Acknowledgements

The research was supported by the Bolyai János Research Scholarship of the Hungarian Academy of Sciences (A.D.), and the EKÖP-25 University Research Scholarship Program of the Ministry for Culture and Innovation (Zs. Sz.).

FROM CHAOS TO PATTERNS: UNRAVELING INTER-SPEAKER VARIABILITY IN ARTICULATION RATE ACROSS ADULTHOOD

Cécile Fougeron

Laboratoire de Phonétique et Phonologie, CNRS, Univ. Sorbonne Nouvelle, Paris, France

cecile.fougeron@sorbonne-nouvelle.fr

Introduction

When we look at variability in large corpora, we are struck by what look like an apparent chaos. When it comes to inter-speaker variability in large cohorts, examining the distribution of individual data points is essential to transform the apparent chaos into identifiable global trends, subgroup profiles, and individual singularities.

In the present study, we address this issue by analyzing a large database investigating how individual and group patterns may inform our knowledge on the evolution of speech across adulthood.

Method

252 French-speaking individuals aged 15 to 90 years were recorded using the same protocol, originally designed for the assessment of speech disorders ([1, 2]) and comprising multiple tasks.

Here, we examine individual variation in articulation rate across two tasks:

1. A text-reading task, involving the reading of a short story containing 8 sentences and a total of 500 words.
2. A diadochokinetic (DDK) task, where speakers must continuously repeat, as quickly and accurately as possible, sequences of either identical syllables (either /bababa/, /dedede/, or /gogogo/) or mixed syllables (e.g., /badego/). This task is commonly used to assess motor control in the rapid repetition of alternating movements in clinical settings, and as such is a performance task.

Articulation rate is measured in phonemes per second (ph/s) in the text-reading task across 23 to 49 inter-pausal units (IPUs) per speaker, and in syllables per second (syll/s) over 4 seconds of uninterrupted repetitions in the DDK tasks

Results

Results, illustrated in Figure 1, reveal striking variability across and within age groups:

- Ages 15–25: There is high inter-individual heterogeneity in reading rate, with suboptimal DDK performance, suggesting that the speech production system is still maturing and follows diverse individual trajectories.
- Ages 25–35: A distinct, homogeneous subgroup emerges, characterized by a faster articulation rate in the reading task and excellent DDK performance. These speakers appear to represent a peak in temporal control of articulation.
- Ages 35+: Articulation rate decreases progressively with age, with a more pronounced slowing observed from around age 70, replicating earlier findings [3, 4 among others].
- Ages 70+: Crucially, while reading becomes slower, more than half of the speakers perform as well as middle-aged adults in DDK tasks, indicating preserved articulatory motor control despite aging.

These findings highlight that within-group variability is as informative as average trends:

- Before age 25, speech maturation follows highly individualized paths.
- Between 25 and 35, a window of optimization emerges for some speakers.
- After 70, only a subset of speakers exhibits signs of attrition, while others retain strong control over articulation rate.

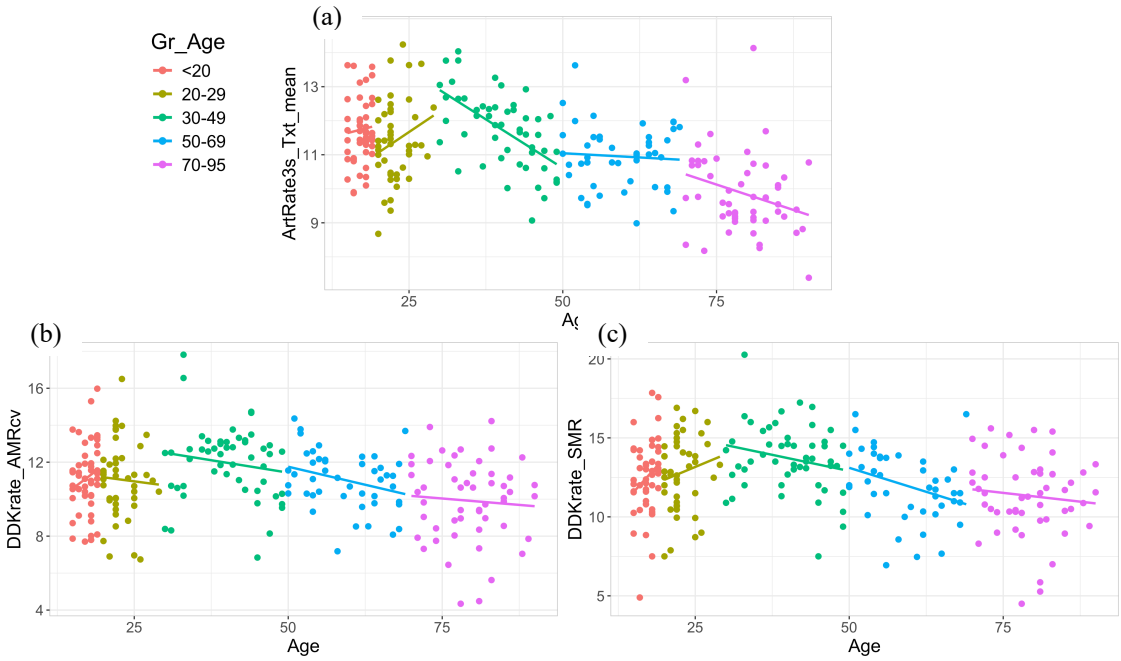


Figure 1. *Per speaker (N=252) ordered by age: (a) articulation rate (ph/s) in text reading task, (b) DDK rate (syll/sec) for identical syllables (mean of /bababa/, /dedede/, and /gogogo/, DDKrate_AMRcv) and (c) mixed syllable (/badego/ sequence, DDKrate_SMR). For ease of visualization, speakers have been artificially grouped in the graphs into 5 equal-sized age-groups of approximately 50 speakers each, but trends are analyzed across the entire age range (15 to 90 years).*

This work is supported by the Swiss National Science Foundation Sinergia Grant (number: CRSII5_202228).

References

[1] FOUGERON, C., DELVAUX, V., MENARD, L., & LAGANARO, M. (2018). "The MonPaGe_HA database for the documentation of spoken French throughout adulthood." *Proceedings of LREC 2018*, Miyazaki, Japan, pp. 4301-4306

[2] LAGANARO, M., et al. (2021). "Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: the MonPaGe-screening protocol." *Clinical Linguistics & Phonetics* 5 (11), pp.1060-1075

[3] TUCKER, B.V., FORD C., & HEDGES S. (2021) "Speech aging: Production and perception." *WIREs Cogn Sci.* 2021; 12:e1557.

[4] FOUGERON, C., GUITARD-IVENT, F., & DELVAUX, V. (2021). "Multi-Dimensional Variation in Adult Speech as a Function of Age." *Languages* 6: 176.

DO SPEAKERS DIFFER IN THE WAY THEY VARY?

Cécile Fougeron[1], Nicolas Audibert[1], Christine Meunier[2]

[1]Laboratoire de Phonétique et Phonologie, CNRS, Univ. Sorbonne Nouvelle, Paris, France
[2]Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

{cecile.fougeron, nicolas.audibert}@sorbonne-nouvelle.fr, christine.meunier@univ-amu.fr

Introduction

Variability is inherent to the speech production process but it however occurs between certain bounds, so that it seldomly compromises communication between speakers sharing the same code, nor extends beyond the blurred zone when speech turns to atypical. It is also clear that listeners can make sense of (some of) this variation as a source of information on the linguistic content and on the speaker's identity and condition, for instance (Kreiman & Lee, 2025, among many others).

However, one thing we do not fully apprehend is how within- and between-talker variation relates to each other (Eatock & Mason, 1994). Since within-speaker variation can make the speech of a speaker resemble or dissemble the speech of others, one needs to know how individual acoustic spaces relates to the acoustic space of other talkers. In this study, we will question how within- and between-talker variability profiles can be a trait characterizing speakers.

Method

To answer this question, we looked at the variability of the acoustic properties of vowels produced by 23 French speakers selected from the PTSVOX database (Chanclu et al., 2020) including 11 female and 12 male speakers between 18 and 24 years-of-age. All were recorded on a minimum of two sessions in two speech tasks: while reading three short texts (Read speech) and during an interview about their studies and occupations (Spontaneous speech).

A total of 385,724 vowel tokens, corresponding to 6 oral and 3 nasal vowel categories of French, were selected. Their acoustic properties were measured with 12 MFCCs (extracted on a 15ms frame centered in the middle of the vowel), which produce a richer representation of the acoustics of the vowels than formants and allow the inclusion of nasal vowels.

Individual differences in terms of variability in the within- and between-speaker dimensions is measured for each speaker by 2 indices:

- The 'Distance2others' (***d2o***) index captures (in both read and spontaneous speech tasks) the acoustic distance between the speaker's vowel tokens and that of the other speakers of the same sex (with a control of vowels identity and consonantal contexts). As such, these indices approximate how the acoustic space of an individual, in a given task, differs from that of the other speakers, i.e. how this speaker is singular and distinct for the others.
- The 'Within-Speaker Variability' (***wsVar***) index captures the acoustic distance between the speaker's vowel tokens in the read speech task and in the spontaneous task. As such, this index approximates the size of the acoustic space of each individual in terms of variability across speech style. High *wsVar* indices characterize speakers with a large acoustic space in which vowels tokens differ in acoustic realization, while speakers with a low *wsVar* have more stable acoustic realization of their vowels across speech styles.

Results

Figure 1 illustrates how the 23 speakers vary according to *wsVar* (i.e. within-talker variability) and to *d2o* (acoustic distance from other speakers), in the reading task (left) and in the spontaneous task (right).

First, individual differences appear in both dimensions: (a) along the *wsVar* dimension we see that speakers differ in how much variability they have in the acoustic realization of their vowels; (b) along the *d2o* dimension we see that speakers differ in how distinct or similar they are to the other speakers in the pool. When comparing *d2o* indices in the two speech tasks show that speakers are globally more distinct from one-another in the reading task than in spontaneous speech. We hypothesize that acoustic reduction in spontaneous speech makes talkers sound more similar, while clearer speech in read speech favors singular traits to surface.
The relation between the two dimensions is not straightforward with an overall weak positive correlation between the two dimensions (r=.5 in read and .6 in spontaneous speech). Three profiles rather explain the distribution of the speakers in the two dimensions (i) in the upper-right of each plots, a handful of 'atypical and inconsistent' speakers shows large acoustic distance between their own vowels (*wsVar>90)* and their large acoustic vowel space also make them distinct from the other speakers in the pool (*d2o* >115); (ii) in the lower left corner, another group of 'conventional and regular' speakers show stable acoustic realization of their vowels (*wsVar<90)* and their acoustic space is shared by other speakers in the pool, from which they show little distance (*d2o* <120 in read and 120 in spontaneous speech); (iii) the third profile concern the remaining speakers who are also stable in the acoustic realization of their vowels (*wsVar<90)* but present individual traits that make them distinct from the other speakers in the pool (d2o>120 in read speech and >115 in spontaneous speech).
While these results need to be confirmed on a larger database (considering the fact that the 'distance to others' index obviously depends on the number of speakers in the pool), the implications of these three different speaker's profiles on speech perception and speaker recognition remain to be further tested in follow-up studies.

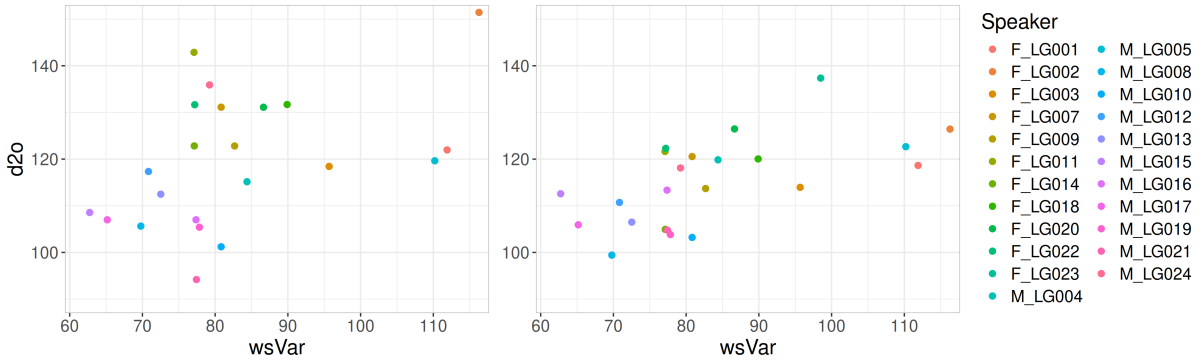


Figure 1. *For each speaker, relationship between his/her own variability (wsVar) in the vowel acoustic properties across speech style and his/her distance to the other 22 speakers' acoustic space in the reading task (left) and in the spontaneous task (right).*

References

Kreiman J. and Lee Y. (2025) Biological, linguistic and individual factors govern voice quality *J. Acoust. Soc. Am.* 157, 482–492 (2025)

Eatock, J.P., and Mason, J.S. 1994. A quantitative assessment of the relative speaker discriminating properties of phonemes, in Acoustics, Speech, and Signal Processing-ICASSP- 94, vol. 1, pp. I–133.1994

Chanclu, A., Georgeton, L., Fredouille, C., Bonastre, J.F. 2020. PTSVOX : une base de données pour la comparaison de voix dans le cadre judiciaire. 6e conférence conjointe JEP-TALN, Nancy, France. pp.73-81.

A SPONTÁN BESZÉDBEN MEGJELENŐ NYELVHASZNÁLATI SAJÁTOSSÁGOK GENERÁCIÓS KÜLÖNBSÉGEI

Főző Eszter

NBSZ Szakértői Intézet

fozo.eszter@nbsz.gov.hu

Bevezetés

A kriminalisztikai nyelvészek bizonyos esetekben az írott szövegfeldolgozás mellett hangzószöveg-elemzéssel is foglalkoznak, amikor a hangtechnikai szakértők az akusztikai, fonetikai és/vagy a biometrikus beszélőazonosítási vizsgálatokat nem vagy korlátozottan tudják elvégezni (például a hanganyag minősége nem megfelelő, a beszélő szándékosan torzítja a hangját, vagy a hangfelvételen hallható beszéd DeepFake technológiával készült). Ezekben az esetekben, ha a beszédproduktum érthető, akkor a beszédleirat szövegnyelvészeti módszerekkel feldolgozható (Leemann et al. 2025).

A nyelvi profilalkotás hazai szakirodalma meglehetősen hiányos, kevés a nem, életkor, iskolázottság stb. fogalmazói csoporttulajdonságok behatárolását célzó kutatás írott és hangzó (beszédleirat) szövegeken egyaránt (vö. Markó et al. 2022, Szabó et al. 2017). Az életkor behatárolásához kapcsolódó beszélőiprofil-alkotási irodalom a hangzó beszéd vizsgálatát részesíti előnyben (pl. Tatár 2013, Beke 2018), ugyanakkor a beszédleiratok szövegnyelvészeti feldolgozása a nemzetközi szakértői gyakorlatban a beszédelemzés része (vö. Morrison et al. 2018, Gold–French 2019).

Módszer

A nyelvi profilalkotás és az egyéni beszédsajátosságok témakörében a BEA (Gósy et al. 2012) és a Longitudinális (Gráczi et al. 2020) adatbázisokkal végzett, korábbi, eredményes longitudinális vizsgálataink alapján (pl. Főző 2023, Markó et al. 2024) jelen kutatáshoz egy nagyobb beszélőszámú, ugyanakkor a profiljegyek tekintetében a lehető legkevesebb változót tartalmazó korpuszon végeztünk elemzéseket. 200 különböző beszélő (100-100 férfi és nő) spontán beszédsajátosságait elemeztük. A korpusz az NBSZ hangszakértői adatbázisa, melyben valamennyi beszélő magyar anyanyelvű, közép- vagy felsőfokú végzettségű. Az adatbázis spontán beszédet tartalmaz (interjú), az egy beszélőre vágott felvételek hosszúsága 2–5 perc (beszélőnként 117–2039 token (= szövegszó) hosszúságú beszédleirat). A beszélők 100%-a a kriminalisztikai szempontból releváns életkori csoportba tartozik (Skarnitzl–Vaňková 2017), 25 és 49 év közöttiek. Mivel jelen vizsgálat elsősorban az életkorhoz tartozó beszédsajátosságok feltárására irányul, nem pusztán a hangtechnikai szakértői módszertanban alkalmazott 10 évenként sávozott életkori (pl. Tatár et al. 2021), hanem a nyelvészetben használt generációs (pl. Nemes 2019, Steigervald 2022) intervallumot is figyelembe vettük. A beszélők az Y (genY) és az X (genX) generációba tartoznak. A teljes korpusz 223667 tokent tartalmaz. A korpusz mérete mind a nemi (nők: 107016, férfiak: 116651 token), mind az életkori csoportosításban (genY: 107624, genX: 116043 token) kiegyensúlyozott.

A hanganyagokat gépi beszédleiratozóval (Phonexia 2026) jegyeztük le. A beszédleiratokat manuális és szoftveres (pl. SZIRA, Vincze et al. 2022; AntConc 4.3.1, Lawrence 2024) úton dolgoztuk fel. A vizsgálat egyik részében az egyes beszédleiratokon lefuttattuk a gépi elemzést, mely különböző morfológiai, szófaji, szemantikai, pragmatikai szinteken abszolút (tényleges (darabszám)) és relatív előfordulás (a beszédleirat szószámához viszonyított előfordulás (arány)) előfordulási adatokat szolgáltatott. Az X és Y generációs beszélők relatív előfordulási

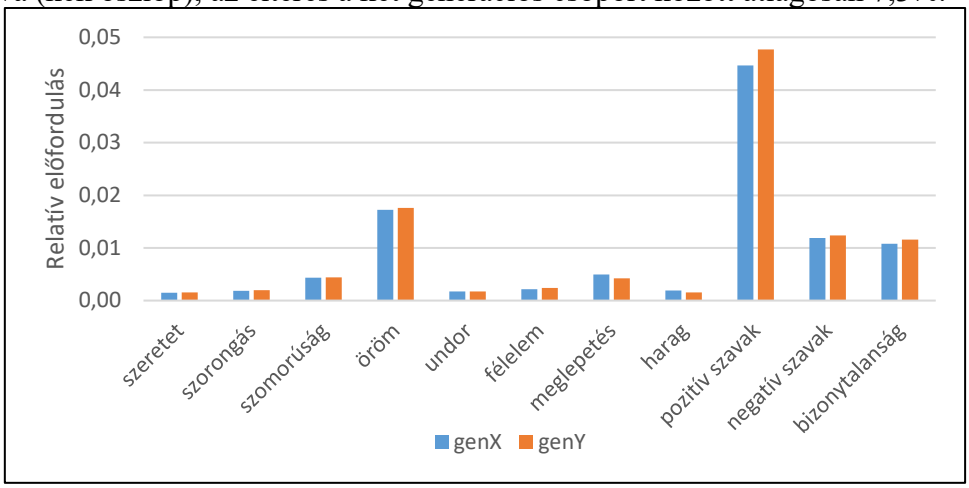
értékeit átlagoltuk, majd összevetettük egymással a hasonlóságok és különbségek megállapításához. A vizsgálat második részében a szókészlet vizsgálatához a teljes korpusz 500 leggyakoribb szavának és azok kollokációinak gyakoriságát elemeztük és hasonlítottuk össze a generációs csoportoknál.

A beszélő életkora a felvétel idején	A beszélő születési dátuma	A beszélő generációs besorolása	A beszélők száma				Tokenszám	
			Nők		Férfiak		Nők	Férfiak
20–29 év	1987–1991	genY	6	49	5	47	53200	54424
30–39 év	1977–1986		43		42			
40–49 év	1967–1976	genX	51		53		53816	62227

1. táblázat *A korpusz paraméterei*

Eredmények

Az első eredmények alapján a vizsgált több mint 80 nyelvi jellemző tekintetében alig van különbség az X és Y generációs csoportok relatív előfordulásai között (átlagosan 11% a szimmetrikus eltérés). Például az érzelemkategóriában a vizsgált 11 jellemző közül 9-ben a genY-beszélőknél (narancs oszlop) volt nagyobb a relatív előfordulás a genX-beszélőkhöz viszonyítva (kék oszlop), az eltérés a két generációs csoport között átlagosan 7,3%.



1. ábra *Az érzelemszavak relatív előfordulása az X és Y generáció beszélőinél*

Következtetések

A vizsgálat alapján a beszédleirat-alapú szövegnyelvészeti elemzés életkorbecslésre (életkori korcsoportba sorolásra) korlátozottan alkalmazható; feltehetően egyéb csoportjellemzők azonossága (pl. az iskolázottság szintje) is hatással van a nyelvhasználati jegyekre, így elmosódnak a generációs különbségek.

Köszönetnyilvánítás

A kutatást az NKFIH az FK128814-es és az Európai Unió az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében támogatta.

Irodalom

Beke A. 2018. Kriminalisztikai alapú beszélőiprofil-alkotás. In: Berend G. – Gosztolya G. – Vincze V. (szerk.): *XIV. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Intézet. 231–244.

Főző E. 2023. A beszédsajátosságok változása 10 év elteltével – a beszédleiratok lexikai vizsgálata és az eredmények felhasználhatósága a beszélőazonosítás műveletében. In:

Gráczi, T. E. (szerk.): *SpeakVar Workshop: A workshop on intraspeaker and interspeaker variability*. Budapest: Hungarian Research Centre for Linguistics. 66–68. https://speakvar.nytud.hu/wp-content/uploads/2023/10/SpeakVar-Workshop-1.pdf

Gold, E. – French, J. P. 2019. International Practices in Forensic Speakers Comparisons: Second Survey. *Internatinal Journal of Speech, Language and the Law*. 26/1. 1–20. doi: 10.1558/ijsll.38028

Gósy M. – Gyarmathy D. – Horváth V. – Gráczi T. E. – Beke A. – Neuberger T. – Nikléczy P. 2012. BEA: Beszélt nyelvi adatbázis. In: Gósy M. (szerk.): *Beszéd, adatbázis, kutatások.* Budapest: Akadémiai Kiadó. 9–24.

Gráczi T. E. – Huszár A. – Krepsz V. – Száraz B. – Damásdi N. – Markó A. 2020. Longitudinális korpusz magyar felnőtt adatközlőkről. In: Berend G. – Gosztolya G. – Vincze V. (szerk.): *XVI. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Intézet. 103–115.

Lawrence, A. 2024. AntConc (Version 4.3.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.lauranceanthony.net/software/AntConc

Leemann, A. – Perkins, R. – Buker, G. S. – Foulkes, P. 2025. *An Introduction to Forensic Phonetics and Forensic Linguistics*. London – New York: Routledge.

Markó A. – Főző E. – Gráczi T. E. (2022) Alkalmazható-e a diskurzusjelölők vizsgálata a beszélői profilalkotásban? *Alkalmazott Nyelvtudomány 2022/1. különszám*. 74–93. https://doi.org/10.18460/ANY.K.2022.006

Markó A. – Főző E. – Gráczi T. E. 2024. *Hát ugye egyébként szerintem,* avagy diskurzusjelölő-társulások mint egyéni beszédjellemzők a kriminalisztikai azonosításban. In: Fogarasi K. – Ittzés D. – Mány D. (szerk.): *A MANYE Kongresszusok Előadásai 14/1. XXIX. Magyar Alkalmazott Nyelvészeti Kongresszus.* Budapest: Akadémiai Kiadó. https://mersz.hu/dokumentum/m1262tia__102/

Morrison, G. S. – Enzinger, E. – Zhang, C. 2018. Forensic Speech Science. In: Freckelton, I. – Selby, H. (eds.): *Expert Evidence: Law, Practice, Procedure and Advocacy* (Ch. 99). Sydney, Australia: Thompson Reuters.

Nemes O. 2019. *Generációs mítoszok*. Budapest: HVG-Kiadó Zrt.

Phonexia 2026. *Speech Recognition and Voice Biometrics.* Phonexia. https://www.phonexia.com/

Skarnitzl, R. – Vaňková, J. 2017. Fundamental frequency statistics for male speakers of common Czech. *Acta Universitatis Carolinae: Philologica 3. Phonetica Pragensia* 2017/3. 7–17. https://doi.org/10.14712/24646830.2017.29

Steigervald K. 2022. *Generációk harca. Hogyan értsük meg egymást?* Budapest: Partvonal Kiadó.

Szabó M. K. – Lázár B. – Nyíri Zs. – Morvay G. 2017. A negatív emotív elemek vizsgálata a nemek közötti nyelvhasználati különbségek szempontjából. In: Gráczi T.E – Ludányi Zs. (szerk.): *XI. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. MTA Nyelvtudományi Intézet, Budapest. 161–176.

Tatár Z. 2013. Beszélőprofil-alkotás lehetőségei a kriminalisztikai fonetikában. *Alkalmazott Nyelvtudomány* XIII/1–2. 121–130.

Tatár Z. – Varga Z. – Főző E. 2021. Beszélőazonosítás a kriminalisztikában. In: Markó A. (szerk.): *Tanulmányok a beszédtudomány alkalmazásainak köréből*. Budapest: ELTE Eötvös Kiadó. 229–251.

Vincze V. – Főző E. – Kicsi A. – Vidács L. 2022. SZIRA: Szövegfeldolgozó Információs Rendszer és Adattár a szerzőazonosítás szolgálatában. *Alkalmazott Nyelvtudomány 2022/1 különszám* 52–73. DOI: https://doi.org/10.18460/ANY.K.2022.005

DIADOCHOKINETIC TASKS IN MULTIPLE SCLEROSIS: ACOUSTIC AND ARTICULATORY ASPECTS

Tekla Etelka Gráczi[1, 2] - Sarolta Murányi[3, 1, 2] - Katalin Pirsel[3, 1, 2] - Zsuzsa Szánthó[3, 1, 2] - Ákos Blaskovics[1, 4]

[1]ELTE Research Centre for Linguistics
[2]MTA−ELTE Lendület "Momentum" Neurophonetics Research Group
[3]ELTE Eötvös Loránd University
[4]Budapest University of Technology and Economics, Faculty of Natural Sciences, PhD School in Psychology

graczi.tekla.etelka@nytud.hu, muranyi.sarolta@btk.elte.hu, pirsel.katalin@nytud.hu, szantho.zsuzsa@btk.elte.hu, blaskovics.akos@nytud.hu

Introduction

Multiple Sclerosis (MS) is a neurodegenerative disease leading to motor and cognitive changes (Ismail–Saqr 2022). As the appearance of demyelination is diverse, the symptoms vary largely across the patients including dysarthria, apraxia and dysphonia in roughly 50% (Kieling et al. 2023). Both tongue and jaw movements were found to affected (Mefferd et al. 2019). Temporal and variability measures of diadochokinetic (DDK) tasks (repeating 1–3 syllable (nonsense) words as fast and as precise one can) (Duffy 2013, Rozenstocks et al. 2020) is a usual method in speech pathology to examine speech motor disorders, that we combined with articulatory measures (ultrasound tongue imaging (UTI) analysis in order to investigate the tongue movement, and compare the results with the gross and manual fine motor (walk and drawing). Our hypotheses were that (i) speakers with MS will show larger within speaker variation in both acoustic and articulatory measurements in the nonsense words produced in DDK tasks than the control speakers, (ii) and that this variation will be related to their gross and fine motor skills.

Methods

12 speakers with MS but officially no dysarthria, apraxia, or dysphonia diagnosis, and 12 control speakers participated in a recording protocol including various speech tasks. We observed the dysarthria symptoms in their reading, and spontaneous speech. The patients' walk and manual fine motor skills (using MOCA's 1−3rd subtests, Nassreddine et al. 2005, Volosin et al. 2013) were also rated. DDK tasks, pronunciation of *papapa* [ˈpɒpɒpɒ] (as a three syllable word) and *pataka* [ˈpɒtɒkɒ] were recorded once only wearing a head-mounted microphone as a practice, later 3 times with UTI (Articulate Instruments Ltd 2014) using a microphone clipped to the textile part of the helmet. We asked the speaker to try to be faster with each trial. The 3 trials of the DDK tasks recorded with UTI were annotated manually into speech segments using the 2nd to 9th repetition in each trial by each speaker. The duration of the word, syllables, sounds, as well as the realization type (frication, approximation) of the consonants were analysed. Both the annotation and measurements were carried out in Praat (Boersma et al. 2026). The participants filled out a questionnaire of their health status. The tongue kinematics of these repetitions were also analysed in *pataka* for each CV- and VC-sequence. The 11 tongue points of the UTI recordings were observed by speaker in order to select the point showing the largest movement connected to the change of the place of articulation of the segments. The start of the gesture onset and the plateau onset and offset were detected using rticulate (Coretta 2025) using the 20% distance from the maximum tongue position. The gesture and plateau duration, as well the variation of maximal displacement were calculated. The articulatory measurements, and the statistical analyses were run in R (R Core Team 2026; LMM: Kuznetsova et al. 2017, and emmeans: Lenth–Piaskowski 2026). The analysis of the speeded repetitions and the articulatory measures are in progress.

Preliminary Results

Most of the duration patterns of the first trials of two non-sense words did not show any significant differences between the two groups, while some MS speakers were the slowest in all three trials per word. The frication of /t/ appeared in both groups with higher frequency in the MS group, and the frication and approximation of /k/ in the slowest trials appeared only in the MS group. The approximation of /k/ also means that the segment appeared as voiced, i.e., both the vocal fold and the lingual articulation were challenged. The variation of the tongue position in the consonants of *pataka* was higher in the MS group. The analysis of the two faster trials and the gesture phases data are in progress.

Conclusions

The first results showed larger acoustic variability and more frequent mispronunciations of the consonants in the MS group although none of them has speech related diagnoses.

Acknowledgements

Á. B., Zs. Sz. and K. P. were supported by the EKÖP-26 University Research Scholarship Program of the Ministry of Science and Technology from the source of the National Research, Development and Innovation Fund.

References

Articulate Instruments Ltd. 2014. *Articulate assistant advanced ultrasound module user manual,* revision 2.14. Articulate Instruments.

Boersma, P. – Weenink, W. – Shchupak, A. 2026. *Praat: doing phonetics by computer* [Computer program]. Version 6.4.67, retrieved 21 May 2026 from https://praat.org

Coretta, S. 2025. *rticulate: Articulatory Data Processing in R.* doi:10.32614/CRAN.package.rticulate. https://doi.org/10.32614/CRAN.package.rticulate, R package version 2.2.0, https://CRAN.R-project.org/package=rticulate.

Duffy, J. R. 2013. *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management,* 3rd ed. St. Louis, MO, USA: Elsevier.

Ismail, I. I. – Saqr, M. 2022. A quantitative synthesis of eight decades of global multiple sclerosis research using bibliometrics. *Frontiers of Neurology* 13. https:// doi.org/ 10. 3389/ fneur. 2022. 845539.

Kieling, M.L.M. – Finkelsztejn, A. – Konzen, V. R. – Brzoskowski Dos Santos, V. – Ayres, A. – Klein, I. – Rothe-Neves, R. – Rozenfeld Olchik, M. 2023. Articulatory speech measures can be related to the severity of multiple sclerosis. *Frontiers of Neurology* 14:1075736. doi: 10.3389/fneur.2023.1075736

Kuznetsova, A. – Brockhoff, P. B. – Christensen, R. H. B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82/13, 1–26. doi:10.18637/jss.v082.i13.

Lenth, R. – Piaskowski, J. 2026. *emmeans: Estimated Marginal Means, aka Least-Squares Means.* doi:10.32614/CRAN.package.emmeans, R package version 2.0.3, https://CRAN.R-project.org/package=emmeans.

Mefferd, A. S. – Lai, A. – Bagnato, F. 2019. A First Investigation of Tongue, Lip, and Jaw Movements in Persons with Dysarthria due to Multiple Sclerosis. *Multiple Sclerosis Related Disorders* 27. 188–194. DOI:10.1016/j.msard.2018.10.116.

Nassreddine Z. S. – Phillips N. A. – Bédirian V. – Charbonneau S. – Whitehead V. – Collin, I. – Cummings, J. L. – Chertkow, H. 2005. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairmert. *Journal of American Geriatrics Society,* 53. 695–699. DOI: 10.1111/j.1532-5415.2005.53221.x

R Core Team 2026. R: *A Language and Environment for Statistical Computing. R Foundation for Statistical Computing,* Vienna, Austria. https://www.R-project.org/.

Rozenstoks, K. – Novotny, M. – Horakova, D. – Rusz J. 2020. Automated Assessment of Oral Diadochokinesis in Multiple Sclerosis Using a Neural Network Approach: Effect of Different Syllable Repetition Paradigms. IEEE Transactions on Neural Systems and Rehabilitation Engineering 28 (3), 767–767.

Volosin, M. – Janacsek, K. – Németh, D. 2013. A Montreal Kognitív Felmérés (MoCa) magyar nyelvű adaptálása egészséges, enyhe kognitív zavarban és demenciában szenvedő idős személyek körében. *Psychiatria Hungarica*, 28/4. 370–392.

TYPOLOGY OF OVERLAP IN WORD-INITIAL CONSONANT CLUSTERS IN GERMAN

Reinhold Greisbach[1] – Anne Hermes[2]

[1]IfL Phonetics, University of Cologne, Germany
[2]Laboratoire de Phonétique et Phonologie (CNRS & Université Sorbonne Nouvelle), Paris, France

reinhold.greisbach@uni-koeln.de, anne.hermes@sorbonne-nouvelle.fr

Introduction

Segmenting consonant clusters using systematic phonetic criteria is challenging because of substantial inter-speaker and intra-speaker variability. A primary source of this variability is in the temporal overlap of articulatory gestures (Browman & Goldstein 1986). In plosive–liquid sequences, at least two consonantal gestures interact (Bombien ea. 2010, Bombien ea. 2013), while an additional glottal activation gesture may also contribute to the acoustic realization. The extent and timing of gestural overlap vary considerably across speakers and tokens, resulting in highly heterogeneous acoustic patterns that complicate segmentation. The present study adopts a fine-grained, microscopic approach on consonant cluster realization in order to identify robust segmentation criteria and to develop a typology of gestural overlap patterns. The analysis focuses exclusively on plosive-/r/ sequences in Standard German. Unlike many previous studies, the investigation does not center on the voiced/voiceless opposition, but instead examines contrasts in place of articulation, namely labial, coronal, and dorsal plosives.

Corpus

The corpus consists of Standard German words forming minimal pair oppositions involving /p–t–k/ and /b–d–g/, supplemented by the rarer contrasts /k–g/, /t–d/, and /p–b/. These additional contrasts were included because true voiced–voiceless oppositions in word-initial consonant clusters are relatively uncommon in German lexical items. The dataset contains 7 paired contrasts and 3 triple contrasts, all of them represented by two syllabic words of German with the contrast in the accented first syllable. These words were embedded in the carrier sentence “Mehr ___ bitte!” (“More ___ please!”). This sentence frame provided a consistent phonetic environment and enabled a standardized segmentation strategy for the clusters under investigation.

Method of Analysis

The material was produced by 10 German native speakers in randomized order 6 times together with distractor items. The recordings were manually segmented in Praat, where the plosive-/r/ sequences were annotated in TextGrid files. Based on these annotations, a Python script generated duration-normalized waveform and spectrogram displays within a fixed-width analysis window. The visually salient structures, primarily observable in the spectrogram, were subsequently annotated manually according to five acoustic features, including voicing and frication noise (see Figure 1). The analysis revealed extremely high levels of both inter-speaker and intra-speaker variation. To facilitate systematic comparison, the analysis window was divided into four zones representing the closure phase (Zone -1), the release phase (Zone 0), and up to two realization zones of the liquid /r/ (Zone 1 and 2). Different combinations of the acoustic features will produce different possible signal structures in each zone. For example, the closure phase (Zone -1) could consist solely of a signal pause, or of a voice offset followed by a pause, or be fully voiced, or feature a pause followed by the onset of the voice before the plosion (negative VOT) – to name just some possibilities.

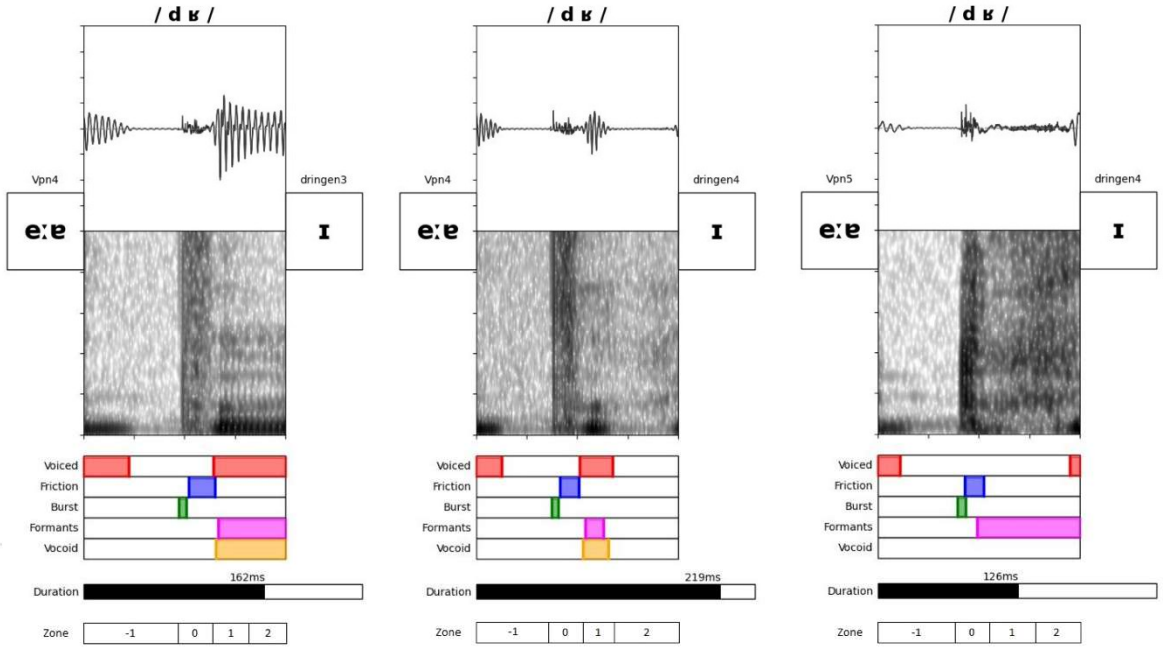


Figure 1. *Oscillogramm (normalized time), spectrogramm (0-10 kHz), tiers for manual classification, duration of segment (in ms), classifying zones. Different realisations of the cluster /dr/ produced by speaker 4 (left and center) and speaker 5 (right). We refer to the formant-like structure visible on the right in the absence of voicing as pseudoformants. The duration of the cluster (in ms) is depicted on the bottom as a black bar relative to a maximum of 250 ms (in this example).*

Results

Our results show that place of articulation of initial plosives in German systematically affect the visible acoustic structures of a cluster. At the same time, speaker-specific realization patterns contribute substantially to the observed variability, with inter-speaker differences proving greater than intra-speaker variation. Figure 2 shows the patterns that occurred most frequently for each speaker across the six repetitions for the first 5 speakers.

ZONE	/br/	-1	0	1	2		/dr/	-1	0	1	2		/gr/	-1	0	1	2
Vpn1	Brillen						drillen						Grillen				
Vpn2	Brillen						drillen						Grillen				
Vpn3	Brillen						drillen						Grillen				
Vpn4	Brillen						drillen						Grillen				
Vpn5	Brillen						drillen						Grillen				

Figure 2. *Color schemes representing different articulations in the proposed zones for the 5 speakers (VpnX) for the minimal triplet "Brillen" (glasses) – "drillen" (to drill) – "Grillen" (crickets).* ***Zone -1****: orange: voice offset before pause;* ***Zone 0:*** *yellow: burst before friction, red: friction before burst;* ***Zone 1 and 2****: yellow: pause; orange: voiceless+pseudoformants; red: voice bar+pseudoformants; magenta: vocalic formants).*

For the first speaker (Vpn1), the place of articulation does not affect the acoustic pattern. In contrast, for the other speakers, there are differences. Remarkably for all of these speakers the coronal plosive /d/ exhibits frication noise already before the release burst. Furthermore, these speakers frequently produce what appears to be an intrusive [ə] after the initial labial /b/ and

coronal /d/ as well, visible as the magenta region in Zone 1 followed by a differently colored region in Zone 2.

Conclusion

Methodologically, the study demonstrates that reliable segmentation of plosive-/r/ clusters requires inspection of frequency ranges extending to at least 10 kHz, since important acoustic cues such as pseudo-formants may occur in higher spectral regions. Furthermore, our findings suggest that, for a more precise classification of acoustic patterns, it may be beneficial for acoustic analyses to extend into the subsegmental domain. The resulting finer-grained classification of acoustic patterns also indicates that caution is warranted with the common statistical practice of averaging measurements across speakers, since this approach may obscure distinct acoustic patterns and the underlying articulatory configurations associated with them.

References

Bombien, L. – Mooshammer, Ch. – Hoole, P. 2013. Articulatory coordination in word-initial clusters of German. *Journal of Phonetics* 41/6. 546-561.

Bombien, L. – Mooshammer, Ch. – Hoole, P. – Kühnert, B. 2010. Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics* 38/3. 388-403.

Browman, C. – Goldstein L. 1986. Towards an Articulatory Phonology. *Phonology Yearbook* 3. 219-252.

FINDING THEIR VOICES: IS IT MORE DIFFICULT FOR SPEAKERS DIAGNOSED WITH DOWN SYNDROME TO INITIATE VOICING?

Kornélia Juhász[1,2] – Katalin Pirsel[1,2,3]

[1]MTA–ELTE "Momentum" Neurophonetics Research Group
[2]ELTE Research Centre for Linguistics
[3]ELTE Eötvös Loránd University

juhasz.kornelia@nytud.elte.hu, pirsel.katalin@nytud.elte.hu

Introduction

This acoustic analysis aims to shed light on the challenges speakers diagnosed with Down syndrome (DS) need to cope with during voice initiation. DS is the most common aneuploid disorder nowadays causing challenges of acquiring speech and linguistic functions. DS speech is characterized by reduced efficiency of upper-respiratory function, which also contributes to limited laryngeal control at the physiological level explaining why DS speakers require twice as much effort to initiate phonation than neurotypical ones (Pryce 1994). The relevance of the study is that according to Kent & Vorperian's (2013) report, prosody was one of the least frequent topics for investigation. Consequently, the most crucial phase of laryngeal control is investigated here: the initiation of voicing. Due to the challenges described above the following question arises: Do the acoustic correlates of the voicing signal indicate more unstable or irregular production for DS speakers when voicing is initiated in sustained vowels?

Methods

The production of 6 adult female DS speakers (DS00*X*); aged 23–40 years) was recorded in line with 9 female control speakers (KO0*XX* = 02–10, aged 19–23 years). Speakers were instructed to sustain the vowels [a], [i] and [u] for approximately 3 seconds. Cepstral peak prominence (CPP) was measured, as Heman-Ackah et al. (2014) suggested that CPP is a better measure of periodicity than HNR because it does not rely on pitch tracking and is therefore reliable even for very aperiodic signals. CPP was extracted by each 1% within the normalized duration of the vowel, but this study focuses exclusively on the values within the first 10% with *praatsauce* (Kirby & Rasmus 2025). CPP reflects the periodicity of the acoustic signal: lower CPP suggests more aperiodic, irregular and breathier voicing (Gittelson et al. 2021). As extensive individual variability is expected in the production of DS speakers we plotted absolute mean CPP for each speaker and for each vowel quality (Fig. 1, right). Similarly, as CPP values can only be interpreted in a relative manner, we took the mean CPP calculated for the entire interval and applied it as a reference for the initial values (i.e., we subtracted the mean CPP of the entire interval from the actual CPP value measured in the first 10%, which means that lower relative CPP values signal more irregular voicing, Fig. 2). We also report group-wise means for absolute and relative CPPs, however these comparisons need to be handled carefully, as the number of speakers in the two groups are imbalanced (Fig. 1, left; Fig. 2, left). As our dataset is limited including only one realization per speaker, we exclusively report descriptive statistics.

Results & discussion

Both absolute and relative CPP values indicate substantial inter-speaker variability among both DS speakers and controls, with some interaction with vowel quality (Fig. 1, Fig. 2). Concerning absolute mean CPPs, four DS speakers (DS001–DS004) produced values within the lower range of the control group, whereas DS006 and DS007 occasionally exhibited CPP values that exceeded those of the controls. As for the relative CPP, [u] exhibits the narrowest range of variability, due to aerodynamic reasons leading to the least mean differences between the two speaker groups (Fig. 2). DS speakers' relative initial CPP values for [a] were generally higher, signalling relatively periodic phonation, in contrast to the control speakers' production, where relatively low initial CPP was possibly induced by a more open vocal tract and lower tongue height (compared to DS production), giving rise to a lower f0 and more unstable phonation. This might occur due to differences in articulatory coordination, such as tongue height differences or timing between the two groups. As for [i], both groups exhibit strong inter-speaker variation with similar patterns. These results contribute to indirectly exploring the difficulties of voicing initiation, as well as to the deeper understanding of the phonation characteristics of speakers diagnosed with Down syndrome.

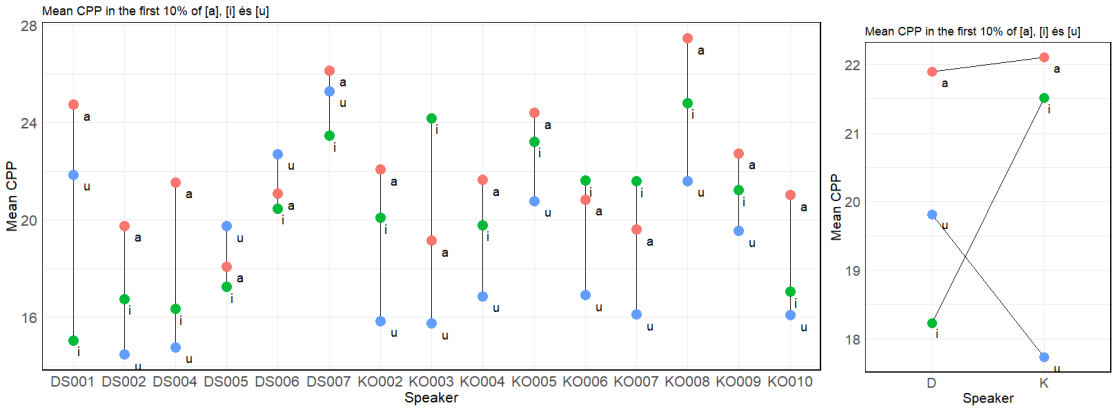


Figure 1. Mean CPP values per speaker and vowel quality separately (left), and group-wise CPP means for the two speaker groups (right, D = DS speakers, K = Controls)

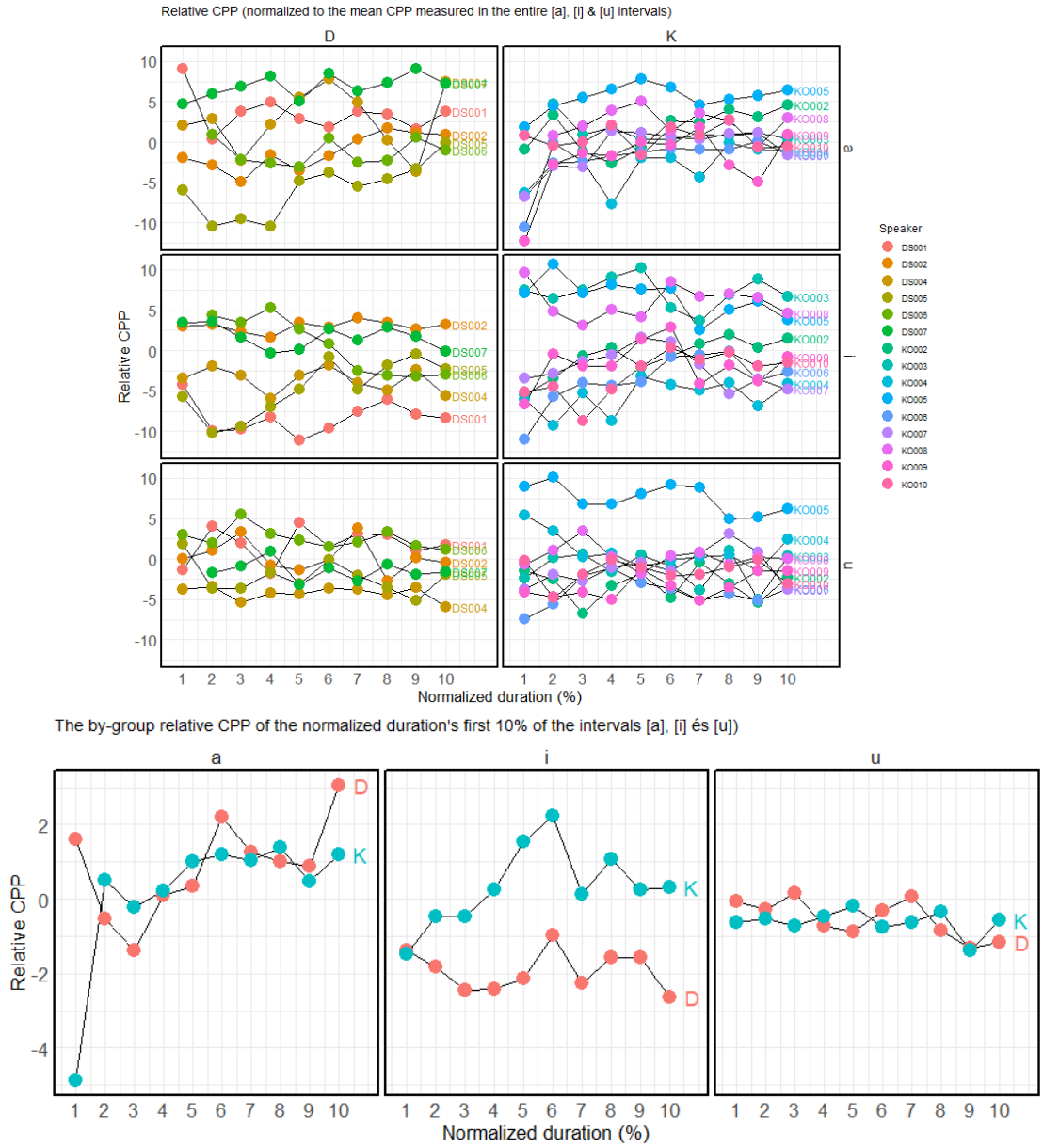


Figure 2. Relative CPP values per speaker (top), and for the two speaker groups (bottom) in the first 10% of the normalized duration

Acknowledgements
The research was supported by the EKÖP-26 University Research Scholarship Programme of the Ministry for Culture and Innovation from the Source of the National Research, Development and Innovation Fund.

References

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. Article 565682.

Heman-Ackah Y. D. – Sataloff R. T. – Laureyns G., Lurie D. – Michael D. D., Heuer R., et al. 2014. Quantifying the cepstral peak prominence, a measure of dysphonia. *Journal of Voice.* 28. 783–788.

Kent, R. D. – Vorperian, H. K. 2013. Speech impairment in Down syndrome: a review. *Journal of Speech, Language, and Hearing Research* 56/1. 178–210.

Kirby, J. – Puggaard-Rode, R. 2025. *PraatSauce: Praat-based tools for spectral processing* [Computer program]. Version 1.0.7.

Pryce, M. 1994. The voice of people with Down syndrome: an EMG biofeedback study. *Down Syndrome Information Network* 2/3.

ACOUSTIC PREDICTOR IMPORTANCE AND SPEAKER-SPECIFIC VARIATION IN RANDOM FOREST MODELS OF CONVERSATIONAL ENGLISH FRICATIVES

Viktor Kharlamov[1] – Daniel Brenner[2] – Benjamin V. Tucker[3]

[1]Florida Atlantic University
[2]University of Arizona
[3]Northern Arizona University

vkharlamov@fau.edu, dbrenner@arizona.edu, benjamin.tucker@nau.edu

Introduction

Prior research on English fricatives has shown that spectral, temporal, and amplitudinal acoustic measures can reliably distinguish fricative categories in statistical modelling studies (e.g., Jongman et al., 2000; McMurray & Jongman, 2011; Kharlamov et al., 2022). However, such findings are typically based on aggregated data, which does not reveal the degree to which the predictive value of individual acoustic measures varies across speakers. This is particularly relevant for research on speaker variability, since phonetic contrasts may be realized differently both across and within speakers (Johnson et al., 1993; Haley et al., 2010).

Methods

The present study examines individual-level patterns of phonetic variability by comparing the relative importance of acoustic predictors in speaker-specific random forest classification models of conversational English fricatives. Random forest modelling was used as it can accommodate non-linear relationships, correlated predictors, and skewed distributions that are commonly found in spontaneous speech data (Breiman, 2001; Tagliamonte & Baayen, 2012). Analysis was based on over 38 thousand tokens of nine English fricatives produced by 24 speakers in a corpus of Canadian English (Wittrock & Tucker, 2019). For each speaker, a separate classification model was trained using commonly employed spectral, durational, and amplitudinal measures. Variable-importance scores were extracted from each model to evaluate which acoustic cues contributed the most to fricative classification at the individual level.

Results

The results show both stability and variability in the importance of acoustic predictors across speaker-specific models. For example, spectral standard deviation measured at the fricative midpoint consistently emerged among the most important predictors, mirroring earlier group-level findings. Its stability across speakers suggests that this measure provides information that is robustly useful for computational classification across speakers. In comparison, measures such as peak frequency, peak power, segment duration, and amplitude-based cues exhibited substantial interspeaker variability in their importance rankings. Some predictors were less reliable at the individual level than the group level, while others were more informative within speakers than across them. In addition, reducing models to a limited set of high-ranking predictors resulted in increased classification error, indicating that lower-ranked measures still contribute to fine-grained distinctions within speakers.

Conclusions

The findings demonstrate that individual-level phonetic variability in the production of fricatives is systematic, and it can affect model performance in distinguishing fricative categories.

The present study is limited to English fricatives, and the findings only concern cue weighting in computational models, not human speech perception. Future research is needed to determine whether similar patterns of predictor importance are observed for other phoneme classes and

to examine the relationship between model-derived patterns of predictor importance and perceptual cue weighting in human listeners.

References

Breiman, L. 2001. Random forests. *Machine Learning* 45/1. 5–32.

Haley, K. – Seelinger, E. – Callahan Mandulak, K. – Zajac, D. 2010. Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. *Journal of Phonetics* 38/4. 548–554.

Jongman, A. – Wayland, R. – Wong, S. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108/3. 1252–1263.

Kharlamov, V. – Brenner, D. – Tucker, B. V. 2022. Temporal and spectral characteristics of conversational versus read fricatives in American English. *Journal of the Acoustical Society of America* 152/4. 2073–2081.

McMurray, B. – Jongman, A. 2011. What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review* 118/2. 219–246.

Tagliamonte, S. A. – Baayen, R. H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24/2. 135–178.

Wittrock, B. J. – Tucker, B. V. 2019. Canadian Prairie dialects: An exploration of Alberta and Saskatchewan vowels. *Canadian Acoustics* 47/3. 117.

JELÖLIK-E ZÖNGEMINŐSÉGGEL A PROMINENCIÁT A MAGYAR ANYANYELVŰ BESZÉLŐK?

Kohári Anna

ELTE Nyelvtudományi Kutatóközpont

kohari.anna@nytud.elte.hu

Bevezetés

Azokban a nyelvekben, ahol a zöngeminőség nem rendelkezik fonológiailag jelentésmegkülönböztető szereppel, a zönge variációi más nyelvi funkciók betöltésében vehetnek részt. Például a zöngeminőség változása jelölheti az intonációs frázisok végét (Garellek, 2022, Markó, 2013). Több nyelvben (pl. amerikai angol, német, holland észt) feltárták, hogy nemcsak a frázisok vége lehet zöngeminőséggel jelölt, hanem a prominens (a környezetéből kiemelkedő) szótagok vagy hangzók is (Garellek, 2022, Bruggeman et al., 2025). Észt nyelvben azt találták, hogy a glottalizáció kevésbé jelenik meg a lexikailag hangsúlyos szótagokon, míg a hangsúlytalanokon jóval gyakrabban fordul elő (Aare et al., 2017). A legújabb megközelítések ugyanakkor a prominencia zöngeminőséggel történő jelölését nem kategorikus keretben tárgyalják. Az utóbbi években egy olyan akusztikai vizsgálati keretrendszert és mérésmódszertant alkottak meg a zöngeminőség tanulmányozására, amelyben a hangzókat egyfelől a periodikus – irreguláris skála mentén, másfelől a hangszalag nyitottsága szerint helyezik el (Garellek 2019). A két skála pedig együttesen jellemzi a zöngeminőséget. Egy mostanában megjelent tanulmány szerint a prominensnek észlelt szótagok hangzóira jellemző volt a nagyfokú periodikusság (HNR, CPP), továbbá a spektrális lejtés mérőszáma (H1*−H2*) alacsonyabb volt prominens szótagok esetén mint nem prominens szótagok esetében (Bruggeman et al., 2025). A magyarban is vizsgálták már korábban prominens szótagok és hangzók zöngeminőségbeli jelöltségét. Markó és munkatársai (2019) megnyilatkozáskezdő pozícióban vizsgáltak magánhangzóval kezdődő álszavakat zöngeminőségkategóriák címkézésével. Eredményeik szerint ebben a pozícióban ejtett magánhangzók mind fókusz, mind topik helyzetben nagy arányban voltak glottalizáltak. Az abszolút megnyilatkozás kezdő helyzet viszont önmagában előidézheti a nemmodális képzést. Következtetéseik szerint ha a kondícióknak lett is volna hatása, elképzelhető, hogy ezt a hatást elmosta volna a megnyilatkozáson és/vagy szón belüli pozíció. Eredményeik azt is megmutatták, hogy a nyíltabb magánhangzók gyakrabban voltak glottalizáltak, mint a felső nyelvállásúak. A magyar beszédben tehát még mindig tisztázatlan az a kérdés, hogy milyen zöngeminőség jellemzi a prominens (a környezetéből prozódiailag kiemelkedő) szótagot a frázisbeli pozíciók kontrollálása mellett, és ez milyen akusztikai sajátosságokkal írható le. Ezért pilotvizsgálatom célja, hogy a prominens szótagok zöngeminőségének akusztikáját nem prominens helyzethez képest különböző pozíciókban (intonációs frázisok belsejében, első és utolsó szavában) vizsgálja. A prominencia zöngeminőséggel történő jelölését nem kategóriák mentén (leheletes, glottalizált zönge) határoztam meg, hanem a zöngeminőséghez kapcsolódó mérőszámok skáláin.

Módszer

Az elemzések egy részét egy nemzetközi mintán alapuló magyar beszédre adaptált, a prominencia vizsgálatára kialakított beszédadatbázison végeztem (Mády et al. 2016). A beszélők képeken szereplő gyümölcsök neveit mondták ki figyelembe véve a gyümölcs nagyságát (kicsi vagy nagy). A méret pedig meghatározza, hogy az első szótag prominens vagy nem prominens módon realizálódik (1. ábra). Az adatbázisban a méret és az észlelhető prominencia között nagyfokú egyezést találtak. A gyümölcsök sorrendben málna, mangó, alma, mandula, mandarin voltak. Az első három gyümölcs szerepelt egy kosárban, és a második kettő egy másik kosárban; tehát a megnyilatkozások két külön fő frázisban realizálódtak. 10 női

beszélő ejtésében vetettem össze a man szótagok magánhangzójának zöngeminőségét a különböző szavakban. Minden szónak volt prominensen és nem prominensen ejtett párja. A beszélők ugyanazt a prominencia mintázatot 5-ször ejtették ki. A hangszintű automatikus lejegyzéseket kézzel ellenőriztem, és a zöngeminőség vizsgálatakor alkalmazott akusztikai méréseket (CPP, HNR05, H1*−H2*) a VoiceSauce programmal végeztem. Lineáris kevert modelleket alkalmaztam R környezetben, amelyekben a zöngeminőség mérőszámai voltak a függő változók, a független változók pedig a prominencia megléte vs. hiánya, továbbá a szavak, amelyek a megnyilatkozás különböző pontjain szerepeltek, és ezen változók interakciója. Az interakciók kiértékeléséhez Tukey-féle post hoc tesztet végeztem.



1. ábra *Egy példa a mutatott tálakra, amelyben a mangó lényegesen nagyobb a többi ételnél*

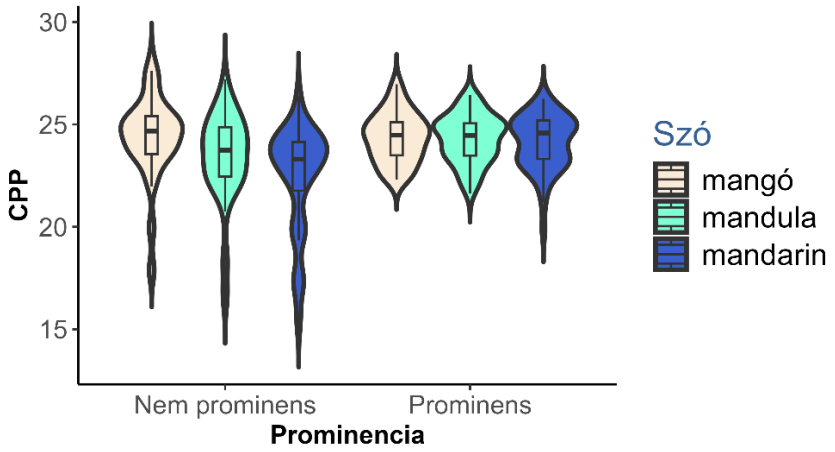
Eredmények

Előzetes eredményeim alapján a periodikusságot vizsgáló mérőszámok (CPP, HNR05) nem mutattak szignifikáns különbséget a prominens és nem prominens szótagok magánhangzójának zöngeminősége között (2. ábra), mindkét helyzetben nagyfokú periodicitás volt megfigyelhető a magánhangzókban. Ugyanakkor a nem prominens szótagok eltérő mintázatot mutattak abban a tekintetben, hogy melyik szóban szerepeltek (2. ábra). Minél inkább a megnyilatkozás végéhez volt közel egy szó, annál inkább jellemző volt rá, hogy első szótagja kisebb CPP és HNR értékekkel rendelkezzen, azaz kevésbé periodikus legyen. Habár az adatok többsége hasonlóan magas CPP és HNR értékeket vett fel a prominenciától függetlenül, a prominensek esetében nem volt jellemző az alacsony érték, a nem prominens szótagok esetében viszont megjelentek ilyen értékek (2. ábra). A spektrális lejtés egyik gyakran vizsgált mérőszáma (H1*−H2*) egyértelmű eltérést mutatott a prominenciától függően (3. ábra). A prominens szótagokban alacsonyabbak voltak ezen mérőszám értékei a nem prominens szótagokban mért értékekhez képest.

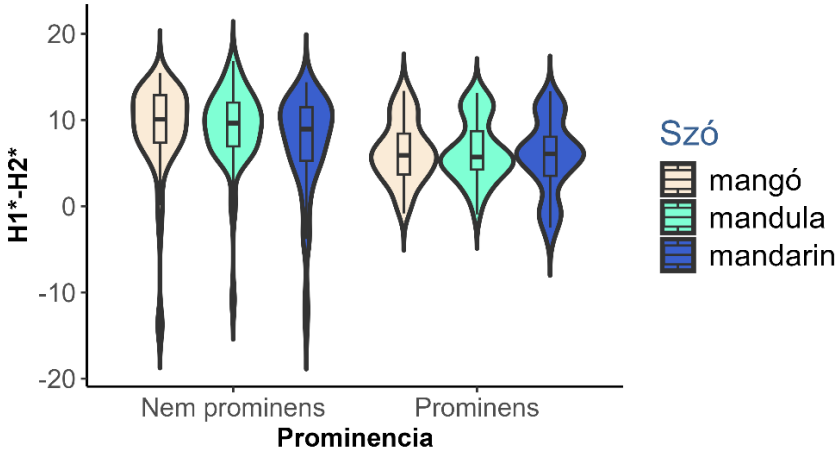
Következtetések

Pilotvizsgálatom eredményei szerint a magyar beszédben is van eltérés a prominens és nem prominens szótagok magánhangzóinak zöngeminőségében. Ezen talált akusztikai eltérések viszont nem feleltethetők meg egyértelműen a klasszikus leheletes-modális-glottalizált zöngetípusok valamelyikének. Helyette azt találtam, hogy a prominens szótagok magánhangzóira is nagyfokú periodikusság jellemző, miközben a prominens szótagok képzését kevésbé meredek spektrális lejtés jellemezheti. Ezek az eredmények összhangban vannak a német beszédben végzett vizsgálatok adataival (Bruggeman et al., 2025). Továbbá az eredmények arra engednek következtetni, hogy minél inkább a megnyilatkozás végéhez volt közel egy szó, annál inkább jellemző volt rá, hogy az első szótagi magánhangzóban nem prominens helyzetben megjelenjen az aperiodikus glottalizáció. Míg prominens szótagok esetében nem volt prototipikus glottalizációra utaló adat. Elképzelhető, hogy a beszélők kifejezetten kerülik ezt a fajta zöngeképzési módot prominens szótagok esetében, míg a frázishatár közelében ennek megjelenése nem prominens szótagokban lehetséges. A vizsgálat limitációja, hogy olyan adatbázison készült, amely csak női beszélők adatait tartalmazta, továbbá nem vizsgáltuk a szókezdő mássalhangzó zöngeminőségét, amelyen akár megjelenhet a prototipikus glottalizáció is. A jelen pilot vizsgálat ugyanakkor megmutatta, hogy a jövőben érdemes a magyar beszédben a prominencia zöngeminőséggel történő jelölését átfogóan is megvizsgálni.

Köszönetnyilvánítás
A kutatást az MTA Bolyai János Kutatási Ösztöndíj (BO/00160/25/1) támogatta.



2. ábra *A CPP mérőszám értékei a prominens és nem prominens szótagokon*



3. ábra *A spektrális lejtés értékei a prominens és nem prominens szótagokon*

Irodalom

Aare, K., Lippus, P., – Simko, J. (2017). Creak as a feature of lexical stress in Estonian. In Interspeech 2017 (pp. 1029-1033). International Speech Communications Association.

Bruggeman, A. – Włodarczak, M. – Wagner, P. (2025). A comparison of discrete and continuous prominence perception methods in German. *Speech Communication*, 168, 103165.

Garellek, M. (2019). The phonetics of voice. In W. Katz – P. Assmann (Eds.), *Handbook of phonetics* (pp. 75–106). Routledge.

Garellek, M. (2022). Theoretical achievements of phonetics in the 21st century: Phonetics of voice quality. *Journal of Phonetics*, 94, 101155.

Markó, A. (2013). *Az irreguláris zönge funkciói a magyar beszédben*. ELTE Eötvös Kiadó.

Markó, A. – Gráczi, T. E. – Deme, A. – Bartók, M. – Csapó, T. G. (2019). Megnyilatkozáskezdő magánhangzók glottális jelöltsége a szintaktikai pozíció és a magánhangzó-minőség függvényében. *Beszédkutatás*, 27(1), 30–53.

BESZÉLŐK KÖZÖTTI ÉS BESZÉLŐN BELÜLI VARIABILITÁS A FRÁZISVÉGI ZÖNGEMINŐSÉGBEN

Kohári Anna[1] – Csébfalvi Fruzsina[2]

[1]ELTE Nyelvtudományi Kutatóközpont
[2] Radboud University

kohari.anna@nytud.elte.hu

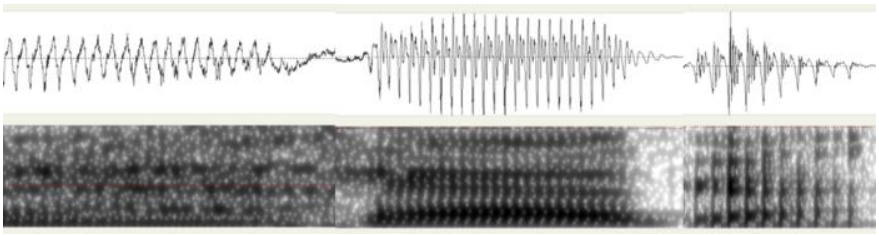
Bevezetés

A zöngeminőség néhány nyelvben fonológiai szempontból jelentésmegkülönböztető szereppel bír (Garellek, 2019), más nyelvekben viszont nem rendelkezik ilyen funkcióval. A legfrissebb percepciós tesztek mégis arra mutatnak, hogy utóbbi nyelvek beszélői is képesek tudatosan megkülönböztetni egymástól a nemmodális zöngeminőséget a modálistól (White et al., 2024), így a meglévő akusztikai variációk felhasználhatók más nyelvi funkciók betöltésére. Ezekben a nyelvekben a nemmodális zöngeképzésnek a beszédben sokféle funkciója lehet, ezek egyike az intonációs frázishatárok jelölése (Garellek, 2019; Davidson, 2021). Több nyelvben, köztük a magyarban is kimutatták, hogy az intonációs frázishatárokon jellemzően glottalizáció jelenik meg, míg a spanyol beszédben ugyanezen pozícióban a leheletes zöngeminőség is igen gyakori lehet (Gonzalez et al., 2022; Markó, 2013). A magyar beszédben a prototipikus megformáltságú, direkt beszédaktusokat tartalmazó különböző mondattípusok (pl. kijelentő vs. felkiáltó mondatok) jellemzően eltérő prozódiával valósulnak meg frázisvégen. Míg a kijelentő és felszólító mondatok tipikusan alacsony határjelző tónussal végződnek, addig a felkiáltó mondatokra inkább a közepes frázisvégi zárótónus jellemző (Gyuris & Mády, 2014). Így a prototipikus glottalizáció, amely irregularitás mellett alacsony alapfrekvenciával is együtt jár, nem valósulhat meg a felkiáltó mondatok végén a magyar beszédben. Újabb akusztikai elemzések feltárták, hogy a beszélők jellemzően inkább modális zöngét ejtenek felkiáltó mondatok végén, és a glottalizáció megjelenése jóval ritkább a kijelentő vagy felszólító mondatok végén produkált zöngeminőséghez képest (Csébfalvi & Kohári, 2025). A beszélők közötti és beszélőkön belüli variabilitás részletesebb elemzése feltárhatja, hogy milyen esetekben ejtenek a beszélők inkább glottalizációt, és milyen körülmények között kerülik el a nemmodális zöngeminőséget frázisvégen. Ezért jelen vizsgálat célja, hogy a feltárja a beszélők közötti és a beszélőkön belüli zöngeminőségbeli különbségeket frázisvégen, és összevesse a realizációkat akusztikai tulajdonságaikkal.

Módszer

10 magyar anyanyelvű beszélő különböző mondattípusokat (kijelentő, felszólító és felkiáltó mondatokat) olvasott fel. A felolvasás előtt nyelvi kontextust mutattunk a kísérletben résztvevőknek, hogy biztosítsuk a konzekvens prozódiai megvalósulást és a nagyon hasonló értelmezést (direkt beszédaktusok ironikus vagy negatív beszélői attitűd nélkül). A kontextus meghatározásához egy olyan átfogó, minden mondattípust elemző, formális elméleti keretet használtunk fel, amely leírja a kommunikációs helyzetben résztvevők szándékát, vélekedését, vágyát, tapasztalatát, és intenzitásskálával jellemzi a különböző attitűdöket (Alberti et al., 2021). Valamennyi vizsgált mondat CVCVCVC szerkezetű célszóval végződött. 10 különböző célszót használtunk fel. Minden célszó előfordult mindhárom mondattípusban, és az összetartozó mondatok azonos szótagszámúak voltak. A megnyilatkozások egy intonációs frázissal valósultak meg. A kísérleti protokollt Dehé és Braun (2020) kutatása inspirálta. A felvételeket manuálisan annotáltuk a Praat programban. Minden célnyelvi mondat utolsó három magánhangzójának megvizsgáltuk a zöngeminőségét, és automatikusan megmértük az alapfrekvenciát, az intenzitást a kifejezetten nem modális zöngeminőségre optimalizált VoiceSauce programmal. Az egyes magánhangzókat perceptuális és vizuális alapon modális, leheletes vagy glottalizáció kategóriába soroltuk. Egy magánhangzó modálisnak számított, ha a zöngeképzés szabályos volt, és nem mutatkozott leheletes vagy glottalizált fonációra utaló jel

(1. ábra). A magánhangzó leheletesnek számított, ha az akusztikai jelben jól hallható és látható zaj jelent meg (1. ábra). Egy magánhangzó glottalizációnak minősült (1. ábra), ha frekvencia- vagy amplitúdóingadozás jelentkezett, illetve ha hallhatóan mutatta a prototipikus glottalizáció jellemzőit (Garellek, 2019). Egy független szakértő az annotációk 10%-át ellenőrizte, és a különböző fonációtípusok jelölései között 96%-os címkézők közötti egyezés mutatkozott. A különböző zöngeminőségbeli kategóriák előfordulási arányainak statisztikai értékeléséhez Fisher-féle egzakt próbát, a beszélők különböző ejtésmódjainak vizsgálatához pedig lineáris kevert modelleket és klaszteranalízist alkalmaztunk R környezetben.



1. ábra *A három zöngeminőség-típus (balról jobbra: leheletes, modális, glottalizált)*

Eredmények

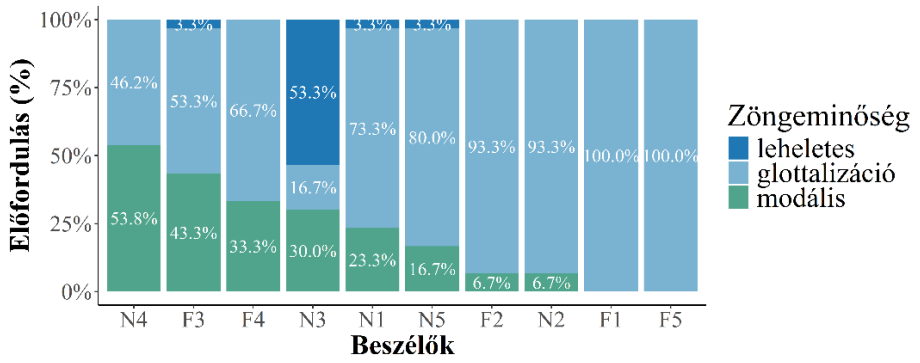
Előzetes eredményeink alapján a kijelentő és felszólító mondatokban jellemzően a beszélők glottalizációt ejtettek frázisvégen, de a modális zönge is megjelenhetett. A felkiáltó mondatokban tipikusan modális zönge fordult elő frázishatáron (2. ábra), míg a glottalizáció jóval kevesebbszer jelent meg ebben a pozícióban (3. ábra). A glottalizáció jellemzően prototipikus formában alacsonyabb f0-lal realizálódott, míg a leheletes zöngével és a modális zöngével megvalósult magánhangzók hasonlóan magasabb alapfrekvencia értékeket mutattak. A glottalizációra alacsonyabb intenzitásértékek voltak jellemzők, mint a modális zöngére. Beszélőkön belül az alapfrekvencia és az intenzitás alapján alapvetően elkülöníthetők voltak a különböző zöngeminőségbeli kategóriák. 10 beszélőből 1-re volt jellemző, hogy leheletes zöngét használt frázisvégen olyan helyzetekben, amikor mások glottalizáltan ejtették a magánhangzót. Egy olyan beszélő volt, aki minden mondattípus végén, így a felkiáltó mondatok végén is csak glottalizált magánhangzókat ejtett, habár a mondat többi részében képes volt modális zöngét produkálni.

Következtetések

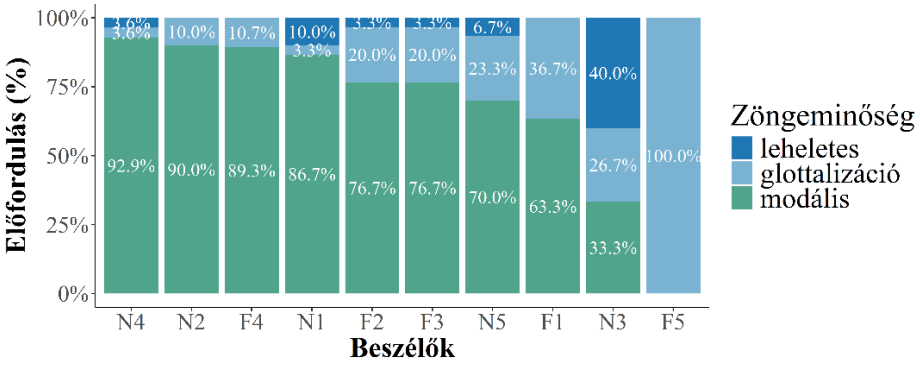
A mondattípus, a frázisvégi zöngeminőség és az alapfrekvencia sajátos összefüggést mutatott. Az alapvetően alacsonyabb alapfrekvenciával megvalósuló kijelentő és felszólító frázisvégeket a beszélők tipikusan glottalizációval valósították meg. Magasabb alapfrekvenciájú kijelentő és felszólító frázisvégeken viszont hajlamosabbak voltak modális zöngét produkálni. A felkiáltó mondatok magasabb alapfrekvencia realizációjának esetében modális, míg alacsonyabb frekvenciák esetében inkább glottalizált magáhangzót ejtettek. Ez arra utalhat, hogy a magyar anyanyelvi beszélők hajlamosak lehetnek elkerülni a magasabb alapfrekvencia esetében a glottalizációt. Habár a glottalizációnak több típusa realizálódhatna, a magyar beszédben frázisvégen jellemzően alacsony alapfrekvenciával és alacsony intenzitással valósult meg. Ez az ejtési preferencia magyarázhatja a beszélőkön belüli zöngeminőségbeli variabilitást. Az eredmények azt is feltárták, hogy a leheletes zönge frázisvégi megvalósulása inkább beszélőre jellemző alternatíva lehet, semmint beszélőn belüli variáció a magyar beszédben.

Köszönetnyilvánítás

A kutatást az MTA Bolyai János Kutatási Ösztöndíj (BO/00160/25/1) támogatta.



2. ábra *A három zöngeminőség-típus felolvasott kijelentő mondatokban*



3. ábra *A három zöngeminőség-típus felolvasott felkiáltó mondatokban*

Irodalom

Alberti, G., Dóla, M., Kárpáti, E., Kleiber, J., Viszket, A., & Szeteli, A. (2021). *Lehetséges lehetséges világaink. Jelentés és Nyelvhasználat*, 8(1), 105–145.

Csébfalvi, F. – Kohári, A. (2025). Voice quality as a marker of phrase-final boundaries in different sentence types. In *Speech Research Conference*, ELTE Research Centre for Linguistics Budapest, 9–10 October 2025 (pp. 27–29).

Davidson, L. (2021). The versatility of creaky phonation: Segmental, prosodic, and sociolinguistic uses in the world's languages. Wiley Interdisciplinary Reviews: *Cognitive Science*, 12(3), e1547

Dehé, N., – Braun, B. (2020). The prosody of rhetorical questions in English. *English Language & Linguistics*, 24(4), 607–635.

Garellek, M. (2019). The phonetics of voice. In W. Katz – P. Assmann (Eds.), *Handbook of phonetics* (pp. 75–106). Routledge.

González, C. – Weissglass, C. – Bates, D. (2022). Creaky voice and prosodic boundaries in Spanish: An acoustic study. *Studies in Hispanic and Lusophone Linguistics*, 15(1), 33–65.

Gyuris, B. – Mády, K. (2014). Approaching the prosody of Hungarian wh-exclamatives. In VLlxx: *Papers in Linguistics*. (pp. 333–349).

Markó, A. (2013). *Az irreguláris zönge funkciói a magyar beszédben*. ELTE Eötvös Kiadó.

White, H. – Penney, J. – Gibson, A. – Szakay, A. – Cox, F. (2024). Influence of pitch and speaker gender on perception of creaky voice. *Journal of Phonetics*, 102, 101293.

IS NONNATIVE ENGLISH SPEECH MORE VARIABLE? EVIDENCE FROM INTRA- AND INTER-SPEAKER VOWEL VARIABILITY

Xuesong Li

Glasgow University Laboratory of Phonetics, University of Glasgow

x.li.15@research.gla.ac.uk

Introduction

The assumption that nonnative speech is more variable than native speech is widespread in second-language (L2) speech research (e.g., Bent et al., 2016; Baese-Berk & Morrill, 2019; Xie & Jaeger, 2020). Greater variability in nonnative speech has also been proposed as one factor contributing to the additional processing difficulty listeners may experience when interacting with nonnative speakers. Nonnative speech is therefore characterised as less stable or more acoustically dispersed (e.g., Romero-Rivas et al., 2015). However, direct empirical evidence for this claim remains limited and inconsistent.

Some studies have reported greater variability in nonnative production: for example, Spanish-accented English vowels have been found to be more acoustically variable and confusable than native vowel production (Wade et al., 2007), while Baese-Berk and Morrill (2015) reported greater variability in speaking rate of nonnative than native English speakers in read speech. Other studies, however, have found lower nonnative variability for particular measures: Baker et al. (2011), for instance, found lower within-speaker word-duration variability in read speech in Chinese- and Korean-accented English than in native American English, while Morrill et al. (2016) reported less speaking-rate variability among nonnative English speakers in spontaneous speech than native speakers. Some studies have even found little or no overall difference: Xie and Jaeger (2020), for example, reported little evidence of greater variability in proficient nonnative English speakers in vowel F1–F2 distributions or durational cues to word-final stops, despite substantial differences in category means and cue relationships. Taken together, these mixed findings suggest that characterising nonnative speech as either more or less variable than native speech in a binary manner seems to oversimplify this complex relationship (Vaughn et al., 2019).

One reason these findings are difficult to reconcile is that "variability" has not always been defined or measured in precise or comparable ways across studies, as Vaughn et al. (2019) note. Existing studies differ in the linguistic properties they examine and measures they use, making direct comparison difficult. A further issue is that variability within individual speakers and variability across speakers are not always examined separately, even though they capture different aspects of speech production: within-speaker variability concerns how consistently an individual produces a phonetic category, whereas between-speaker variability concerns the extent to which category realisations differ across speakers. A systematic comparison of these two components is therefore needed to assess whether nonnative speech is genuinely more variable than native speech.

Given that relatively few studies have explicitly examined variability in nonnative speech, the present study builds on Vaughn et al.'s (2019) distinction between within- and between-speaker variability, examining whether nonnative speakers (NNSs) are more variable than native speakers (NSs) in English monophthong production by decomposing variability into intra- and inter-speaker components — hereafter termed *intra-speaker* and *inter-speaker variability* — and applying a consistent metric to both. This approach allows for a more controlled test of the widely held view that NNS speech is inherently more variable.

Methods

Speech data were collected from 48 NNSs (33 women, 15 men; mean age = 29.6 years), with six participants from each of eight typologically diverse first language (L1) backgrounds, and

6 NSs of Standard Southern British English (5 women, 1 man; mean age = 31.2 years). The NNSs' L1s were Arabic, Bengali, Danish, Estonian, Polish, Punjabi, Russian, and Telugu, selected to vary systematically on dimensions of rhythm class, vowel inventory size, and vowel quantity, for purposes unrelated to the current analysis. The NNSs reported having learnt or spoken English for a median of approximately 20 years (range = 4–61) and having lived in the UK for a median of 2 years (range = 1 month–46 years). They self-reported high overall English-speaking proficiency (mean = 8.46/10). Thirty-four reported a formal English-language qualification, of whom 24 provided IELTS overall scores ranging from 6.0 to 8.5. Most participants (39/48) reported British English as either the only variety or one of several varieties represented in their English-language teaching materials at school, alongside varieties such as American, Indian, and Australian English, although the variety of English they spoke was not independently classified.

In a recording studio, participants read aloud 52 carefully controlled English sentences twice, yielding 5,616 tokens. The sentences were presented one at a time on a computer screen and were designed to elicit short monophthongs, long monophthongs, and diphthongs within both simple and complex foot structures. Recordings were manually segmented in Praat (Boersma & Weenink, 2026), and F1 and F2 were extracted at vowel midpoints. All formant values were Lobanov-normalised to minimise inter-speaker physiological differences and to allow comparison within a shared acoustic space.

The analyses were restricted to monophthongs to provide a stable representation of vowel targets. Variability was quantified as Euclidean distance in the normalised F1–F2 space, measured at two levels. First, intra-speaker variability was defined as each token's Euclidean distance from the corresponding speaker-specific vowel centroid, that is, how consistently a speaker realises a given vowel category. Second, inter-speaker variability was calculated as the Euclidean distance between each speaker's vowel-specific centroid and the corresponding group-level centroid for each group-L1-vowel combination. This measure captures the extent to which speakers within the same L1 background diverge from one another.

Linear mixed-effects models were fitted to each variability measure to compare variability across groups and L1s.

Results

Neither measure showed evidence that NNS monophthong production was systematically more variable than NS production. For intra-speaker variability, the NNS vs. NS group contrast was not significant ($\beta = –0.055$, $t = –0.91$), indicating that NNSs were not less internally consistent than NSs in their production of English monophthong categories. Inter-speaker variability likewise showed no significant group difference ($\beta = –0.016$, $t = –0.26$), suggesting that NNSs were not more dispersed from one another within the same L1 group than NSs in their English monophthong realisations.

L1-specific differences were limited. While Telugu speakers showed greater intra-speaker variability than NSs ($\beta = 0.141$, $t = 3.09$), Estonian speakers showed greater inter-speaker variability ($\beta = 0.126$, $t = 2.97$). However, no comparable differences were observed consistently across the remaining L1 groups or across both measures.

Overall, the results provided no evidence for a general increase in either intra- or inter-speaker variability in NNS English monophthong production. The limited differences observed appeared to be L1-specific.

Conclusion

The present study found no support for generally increased variability in NNS English monophthong production. When variability was decomposed into intra- and inter-speaker components and measured consistently as Euclidean distance in a Lobanov-normalised F1–F2 space, NNSs did not show significantly greater variability than NSs on either dimension.

These findings also highlight a broader methodological issue in L2 speech research: variability has been defined and operationalised in different ways across studies, often with different linguistic properties, measures, and tasks examined. This may partially explain why there are mixed results in the literature. By directly measuring both intra- and inter-speaker variability using a common metric, the present study responds to Vaughn et al.'s (2019) call for a more systematic investigation of variability in nonnative speech. It also supports the need to treat these two aspects of variability as distinct rather than collapsing them into a single property of nonnative speech. The results are also consistent with Baese-Berk and Morrill's (2019) view that patterns of variability are likely to depend on how variability is defined and the linguistic unit under investigation, which echoes Vaughn et al. (2019).

The limited L1-specific effects observed here further suggest that variability in nonnative speech may be shaped by speakers' language backgrounds and the particular L1–L2 pairing examined (Vaughn et al., 2019) rather than by a simple native-nonnative distinction. At least for English monophthongs, greater within-speaker inconsistency and between-speaker dispersion do not appear to be general characteristics of nonnative speech. Future work should therefore examine whether variability is instead structured by L1-specific phonetic patterns, vowel-category relationships, or dynamic properties of vowel realisation, and whether different variability patterns emerge for other phonetic features, elicitation tasks, and measures of variability.

Acknowledgements

This research was funded by the Scottish Graduate School for Arts and Humanities (SGSAH).

References

Baese-Berk, M. M. – Morrill, T. H. 2015. Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America* 138/3. EL223–EL228.

Baese-Berk, M. M. – Morrill, T. H. 2019. Perceptual consequences of variability in native and non-native speech. *Phonetica* 76/2–3. 126–141.

Baker, R. E. – Baese-Berk, M. – Bonnasse-Gahot, L. – Kim, M. – Van Engen, K. J. – Bradlow, A. R. 2011. Word durations in non-native English. *Journal of Phonetics* 39/1. 1–17.

Bent, T. – Atagi, E. – Akbik, A. – Bonifield, E. 2016. Classification of regional dialects, international dialects, and nonnative accents. *Journal of Phonetics* 58. 104–117.

Boersma, P. – Weenink, D. 2026. *Praat: Doing phonetics by computer* [Computer programme]. Version 6.4.48. Available at: http://www.praat.org (Accessed: 15 April 2026).

Morrill, T. – Baese-Berk, M. – Bradlow, A. 2016. Speaking rate consistency and variability in spontaneous speech by native and non-native speakers of English. *Proceedings of Speech Prosody 2016*. 1119–1123.

Romero-Rivas, C. – Martin, C. D. – Costa, A. 2015. Processing changes when listening to foreign accented speech. *Frontiers in Human Neuroscience* 9. 1–15.

Vaughn, C. – Baese-Berk, M. – Idemaru, K. 2019. Re-examining phonetic variability in native and non-native speech. *Phonetica* 76/5. 327–358.

Wade, T. – Jongman, A. – Sereno, J. 2007. Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica* 64/2–3. 122–144.

Xie, X. – Jaeger, T. F. 2020. Comparing non-native and native speech: Are L2 productions more variable? *The Journal of the Acoustical Society of America* 147/5. 3322–3347.

CHOOSE YOUR ACT BY YOUR SOUND: PROSODY-SPEECH ACTS INTERACTION FOR UTTERANCE-FINAL PARTICLES IN MANDARIN CHINESE

Beier Liu[1,2] – Huba Bartos[1,2] – Katalin Mády[1]

[1]Eötvös Loránd University
[2]ELTE Research Centre for Linguistics

liubeier@student.elte.hu, bartos@nytud.hu, mady@nytud.hu

Introduction

The interplay of sentence prosody and particles can modify the speech acts (SAs) performed with the utterance in both tonal and atonal languages by encoding speaker commitment and call on addressee with additional layers (Heim et al. 2016). Particularly, modal particles as metonymic mappings towards SAs (Waltereit 2001) can occupy utterance-final position in Mandarin Chinese (MC) to fulfill certain discourse and relevance functions (Chu 2009). While the phonetic realization of lexical tones is proven to be moderated by the semantic variation of the same word in MC (Chuang et al. 2026), at the sentential level, the mechanism of prosodic modification by SA variation of the same utterance-final particle (UFP) is not yet sufficiently understood, nor is the consistency of such modifications among speakers. This exploratory study therefore investigates the interaction between the MC UFP 吧 *ba* and prosody in the realization of different speech acts.

Methods

To minimize lexical and segmental variation, 6 minimal sentences consisting of a noun, a prototypical labile verb (i.e. 进 *jìn*; enter), the perfective marker 了 *le*, and the UFP 吧 *ba* were constructed, compared to the assertive baseline with Ø in the UFP position. Prototypical labile verbs with both transitive and intransitive readings (Zhang 2017), due to the unmarked causativeness information in MC, can encode different sentence types with the same word order to elicit distinct SAs, namely confirmative assertions with and without expectation of addressee acknowledgement, confirmation-seeking directives with positive and negative previous knowledge, persuasive directives with and without emotional suffering. To preserve a natural speech flow while providing cues to the perceived SA, an additional sentence was given as context after each tested sentence. Five MC L1 speakers produced the materials in a pilot recording experiment via EasyVoiceRecorder (Digipom 2025) over smartphones. The subjects were provided with the entire material before recording without explicit explanation for the research purpose or design and were then instructed to read the target sentences together with the context in a vivid and emotional way. F0, intensity and duration were analyzed acoustically in Praat (Boersma and Weenink 2025). For each of the sound files, the pitch was depicted on the spectrogram with autocorrelation method with speaker-specific range, while the intensity averaged by mean energy with a range between 50 dB to 100 dB. The F0 and intensity contours were compared visually and manually by checking the values of corresponding time points. Data were analyzed by subjective listening perception, together with the judgement of durational information. The analysis process was descriptive without the use of quantitative modelling.

Results

Preliminary analysis indicates that 吧 *ba* utterances share a recurrent prosodic configuration in confirmation-seeking directives requiring a clear addressee acknowledgement with both positive and negative previous knowledge, starting with a flatter noun and falling more drastically with a strongly accented verb, probably due to focus. While the UFP can inherently

provide a meaning of positive knowledge, rendering the negative evidence in the context obsolete, the personal interpretation of stronger questionability can generally be reflected by higher registers compared to the baseline (cf. Juhász et al. 2024) (see Figure 1). However, the UFP tail can only end with a low boundary, unlike the case of questions with 吗 *ma* (optional rising) or Ø (compulsory rising). In contrast, confirmative assertions with or without expectation of minor addressee acknowledgement are both associated with a flatter intonation and reduced accentuation on the verb. They differ, however, in that there is a tail rise optional for the former but seemingly required for the latter. The persuasive directives also have an accented verb with lengthening of either the verb or the UFP, the longer duration of which signifies the higher degree of begging (possibly also the lower authority of the speaker). One sentence of persuasive directive also had readings similar to the baseline, seemingly due to an alternative SA of the context as an assumptive assertion. In all cases the intrusion of the UFP *ba* imposes a shortening of the previous particle *le* as was also shown in Juhász et al. 2025.

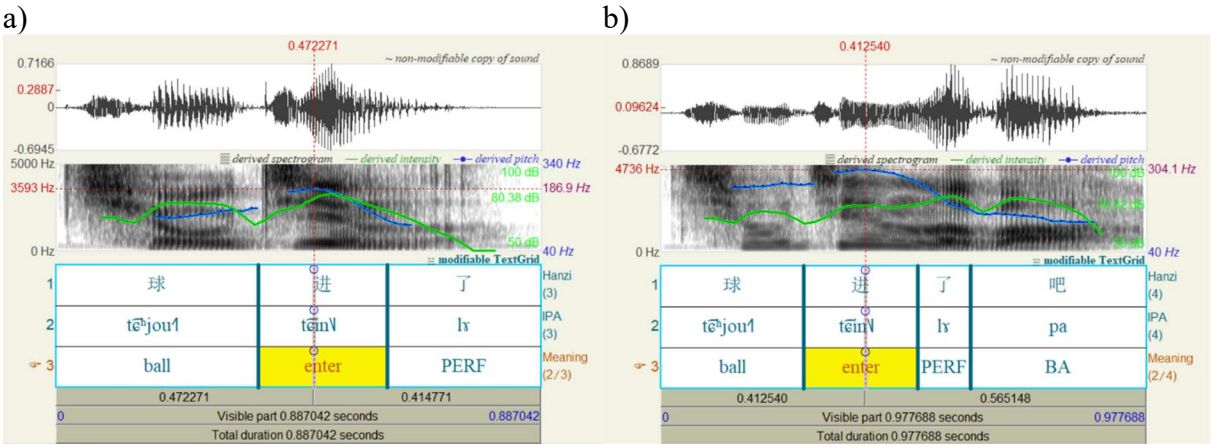


Figure 1. *The pattern of confirmation-seeking directives. a) baseline; b) confirmation-seeking directive with negative previous knowledge from the same speaker. The latter has an elevated and flatter start, a higher verb start position with stronger fall, higher register, and compressed duration of PERF.*

Conclusion

The results provide preliminary acoustic evidence that in MC the SA distinctions are not determined by UFPs alone but emerge from their interaction with prosodic realization, comprising a multimodal construction to clearly encode the semantics and its underlying belief-desire-intention system. The observed speaker variability further suggests that UFP–prosody mappings may constitute a flexible system of prosodic resources rather than fixed intonational realizations, rendering the construction redundantly safe in certain contexts. A larger speaker sample and quantitative normalization of F0 and duration will be required to test these patterns systematically.

Acknowledgements

The research is supported by the China Scholarship Council. Sincere gratitude to the cultivation of ELTE Doctoral School of Linguistics and to the resourceful and wise professors from the ELTE Research Centre for Linguistics.

References

Boersma, P. – Weenink, D. 2025. *Praat: doing phonetics by computer* [Computer program]. Version 6.4.43, retrieved 17 September 2025 from https://www.praat.org.

Chuang, Y.Y. – Bell, M.J. – Tseng, Y.H. – Baayen, R.H. 2026. Word-specific tonal realizations in Mandarin. *Language* 102/1. 1-45.
Chu, C. 2009. Relevance and the Discourse Functions of Mandarin Utterance-Final Modality Particles 1. *Language and Linguistics Compass* 3/1. 282-299.
Digipom Inc. 2026. *Easy Voice Recorder* (Version 1.3.1) [Mobile application software]. Retrieved from App Store.
Heim, J. – Keupdjio, H. – Lam, Z. W. M. – Osa-Gómez – A., Thoma, S. – Wiltschko, M. 2016. Intonation and particles as speech act modifiers: A syntactic analysis. *Studies in Chinese Linguistics* 37/2. 109-129.
Juhász, K. – Gráczi, T. E. – Mády, K. – Hu, F. – Chen, S. – Bartos, H. 2025. Temporal characteristics of Mandarin neutral tones in the production of atonal Hungarian learners of Chinese. In: Gráczi, T. E. – Juhász, K. – Kohári, A. – Mády, K. (eds.): *Speech Research Conference ELTE Research Centre for Linguistics Budapest*. Budapest: ELTE Nyelvtudományi Kutatóközpont. 60-62.
Juhász, K. – Mády, K. – Bartos, H. 2024. The production of Mandarin neutral tone sequences by Hungarian learners. In: Chen, Y. – Chen, A. – Arvaniti, A. (eds.): *Proceedings of Speech Prosody*. Leiden: International Speech Communication Association (ISCA). 547-551.
Waltereit, R. 2001. Modal particles and their functional equivalents: A speech-act-theoretic approach. *Journal of Pragmatics* 33/9. 1391-1417.
Zhang, L. 2017. The typology of labile verbs and Chinese. *Proceedings of the 29th North American Conference on Chinese Linguistics (NACCL-29)* 2. 423-442.

SZTRÁJK A SVÁJCI NEJLONGYÁRBAN. HAT ÉVTIZED ALAKVÁLTOZATAI A FILMHÍRADÓK KORPUSZÁBAN

Markó Alexandra[1,2]

[1]NBSZ Szakértői Intézet, [2]MTA–HUN-REN NYTK Lendület Neurofonetikai Kutatócsoport

marko.alexandra@nbsz.gov.hu

Bevezetés

G. Varga Györgyi 1968-as, *Alakváltozatok a budapesti köznyelvben* című könyvében így ír az idegen szavakban megjelenő *áj-ej* váltakozásról: „Megfigyeléseim szerint a *Svájc* és a *sztrájk* szavak a választékos köznyelvben is körülbelül 50-50%-ban *áj,* illetőleg *ej* megoldásúak" (55). Forró Orsolya 2013-as doktori disszertációjában pedig ezt olvashatjuk a *spájz/spejz* és a hozzá hasonló tövek toldalékolása kapcsán: „Feltételezésem szerint ezeknél a töveknél az antiharmonikus toldalékolás lehetőségének szükséges feltétele a hátsó magánhangzós alakváltozat: elképzelhetőnek tartom, hogy a hasonló szerkezetű *gejl, szejm* és *sztrájk* esetében az ingadozás hiánya a **gájl, *szájm, *sztrejk* változat hiányával is összefügg" (150.). A két forrás között nemcsak a keletkezésük időpontja miatt jelentős a különbség, hanem azért is, mert G. Varga hangzó szövegek, Forró pedig írott korpuszok alapján fogalmazta meg állításait. A helyesírás kodifikáló hatása gyakran elfedi a hangzásban meglévő alakváltozatokat, így az élőbeszéd vizsgálata segíthet ezek feltérképezésében.

A magyar filmhíradók hangzó korpusza elérhető online (https://filmhiradokonline.hu/; Gocsál 2024) 1931-től 1991-ig, és ebben az időszakban a magyar köznyelvi beszéd presztízsváltozatát teszi hozzáférhetővé, viszonylag jó minőségű szöveglejegyzésekkel együtt. Az előadásban három eltérő gyakoriságú szó alakváltozatait, illetve variabilitásuk mintázatait elemezzük a filmhíradók felvételei alapján, ezek a *Svájc/Svejc, sztrájk/sztrejk* és a *nájlon/nejlon.* Az alakváltozatok abban is eltérnek, hogy a mai köznyelvben a *Svájc* és a *sztrájk* esetében az *áj*-t tartalmazó, míg a *nejlon* esetében az *ej*-t tartalmazó változat minősül jelöletlennek.

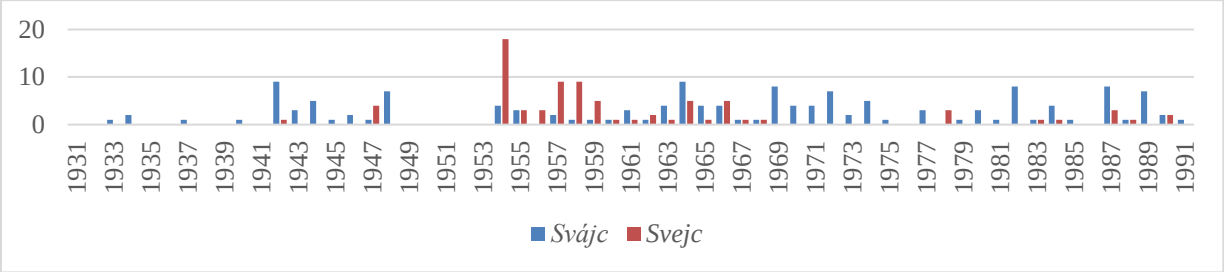
Módszer

A https://filmhiradokonline.hu weboldalon működő kereső segítségével találtunk rá azokra a filmhíradókra, amelyekben a fenti szavak valamelyik alakváltozata szerepel. Dokumentáltuk a filmhíradó címét, felvételi idejét, közvetlen kontextusát, valamint a beszélőről tudható információkat. Ez utóbbi jellemzően a beszélő neme és foglalkozása (bemondó), de néhány esetben az interjúalanyok személyneve és pozíciója (pl. igazgató) is kiderül. Tekintettel arra, hogy a régi filmhíradók zajosak, torzak, és számos esetben háttérzene is megjelenik, objektív akusztikai elemzésre nem volt mód. Így hallás (többszöri meghallgatás és ellenőrzés) alapján rögzítettük az elhangzott alakváltozatot. Elemeztük a szóalakváltozatok előfordulási gyakoriságát a filmhíradó felvételi idejének függvényében. Folyamatban van a beszélővel kapcsolatos információk elemzése, egyelőre a beszélő neméről és státuszáról (bemondó vagy interjúalany) vannak adataink, illetve elsősorban a *Svájc/Svejc* kapcsán tudunk néhány, az egyéni nyelvváltozatra vonatkozó állítást tenni.

Első eredmények

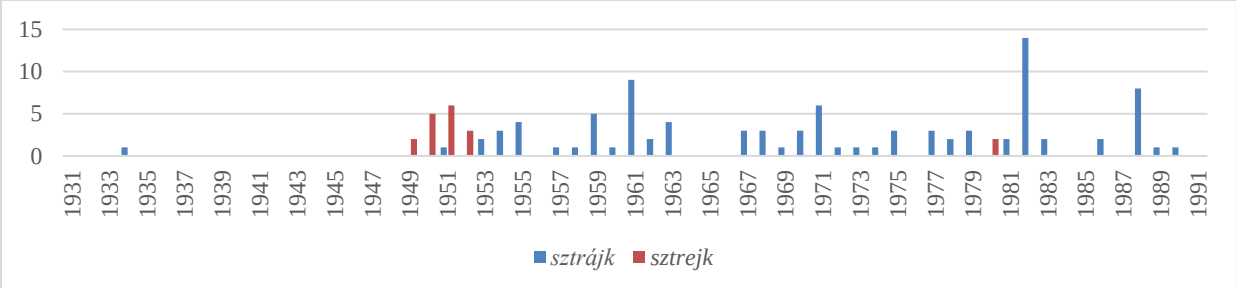
A három célszópár közül a *Svájc/Svejc* fordult elő a leggyakrabban a filmhíradókban, 225 alkalommal. A *Svájc* ejtésváltozattal 1933 és 1991 között 144 alkalommal, a *Svejc* változattal az 1942 és 1990 között rögzített filmhíradókban 81 alkalommal találkozunk (1. ábra). Vagyis az időintervallum nagyjából azonos, és 64 vs. 36%-os arány látszik a *Svájc* alak javára A *Svejc* alak egy kivétellel férfi beszélőktől hangzik el, ami feltételezhetően elsősorban annak köszönhető, hogy mindösszesen 197 esetben fordul elő a célszó valamelyik változata férfi beszélőnél, és csak 28-szor nőknél. Az első női beszélőtől származó adat a célszópár valamely tagjára 1954-ből való, korábban csak férfiak szerepelnek (addig összesen 42 adattal). Mindkét alakváltozat 90%-a bemondóktól és 10%-a más megszólalóktól (színművészek, interjúalanyok)

adatolható, ugyanakkor a *Svejc* változat az 1970-as évektől már alig fordul elő a bemondók ejtésében.
Összességében a *Svájc* ejtés tehát jelentősen gyakoribb volt a teljes korpuszban, de az 1950-es években a *Svejc* realizáció dominált: ebben az évtizedben 38 előfordulást számlált 11 Svájc ejtésváltozattal szemben. A gyakorisági adatok értelemszerűen a beszélők (többségükben bemondók) egyéni ejtésmódja szerint alakulnak, többnyire következetesen, de találunk példákat a beszélőn belüli variabilitásra is. Például egy 1987-es filmhíradóban megszólaló vállalatigazgató 5 alkalommal ejtette ki az ország nevét valamilyen kontextusban, az első alkalommal *svejci*-t, majd kétszer *Svájcban*-t és végül kétszer *svájci*-t mondott. Bán György bemondó egy filmhíradóban a *Svájc, Svájcban* alakot mindig *á*-val, de a melléknévképzős *svejci* alakot következetesen *e*-vel ejtette. Szepesi György sportkommentátorként igen gyakran szerepel a korpuszban, szinte minden esetben az *ej*-változattal, de nála szintén előfordul *áj*-változat is. A toldalékolást tekintve a magánhangzó-harmónia a *Svejc e*-jét minden esetben neutrálisnak mutatja *(Svejcban, Svejcból, svejciak)*, bár ezek előfordulása igen ritka.



1. ábra *A* Svájc *és* Svejc *alakváltozatok előfordulási gyakorisága a filmhíradókban*

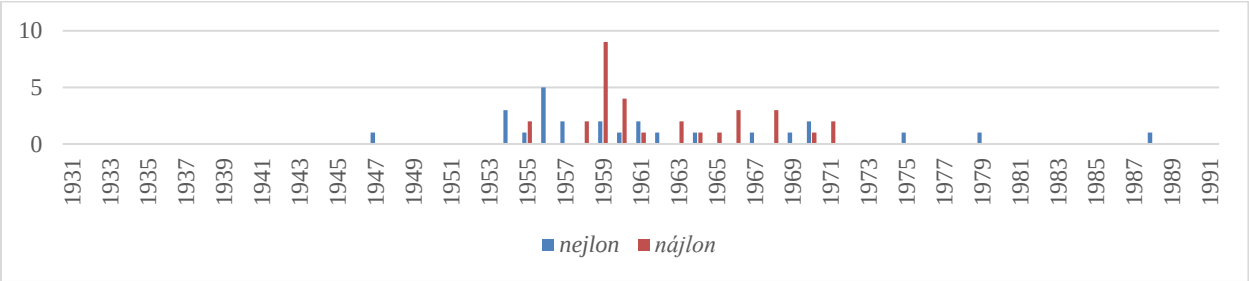
112 előfordulást találtunk a *sztrájk/sztrejk* szópárból, ezek közül 94 *sztrájk* és 18 *sztrejk*, vagyis az arányok itt 84 és 16%. Ezek között 96 férfi és 16 női beszélőktől származó adat van, de a *sztrejk* alakváltozat kizárólag férfi beszélők esetében fordult elő. Az összes előfordulás 80%-a bemondóktól származik, és hasonló az arány az egyes alakváltozatokat tekintve is. A *sztrájk* első előfordulása 1934-ből dokumentálható, majd hosszú szünet után az adatok többsége az 1940-es évek végétől jelentkezik (2. ábra). A *sztrejk* megjelenése a bemondók beszédében 1949 és 1952 közé tehető, az 1980-ból származó két adat egy politikus ejtéséből származik. A toldalékolást tekintve a magánhangzó-harmónia a *sztrejk e*-jét szintén minden esetben neutrálisnak mutatja *(sztrejkba, sztrejkok, sztrejkkal)*, és ezek előfordulása ugyancsak ritka.



2. ábra *A* sztrájk *és* sztrejk *alakváltozatok előfordulási gyakorisága a filmhíradókban*

A *nejlon/nájlon* szóalakpár összesen 57 előfordulást számlál 1947 és 1988 között (3. ábra), minden esetben bemondó a beszélő. A *nejlon* 26 adattal a minta 46%-át teszi ki, míg a 31 *nájlon* 54%-ot. A nemek közötti arány kiegyenlített, 27 férfi és 30 női beszédből származó adat van, ezen belül a *nejlon* a férfiak esetében gyakoribb ejtésmód (15 vs. 11), a *nájlon* pedig a nők esetében (19 vs. 12). A *nájlon* változat az 1970-as évektől már nem fordul elő, de a *nejlon* is

igen ritkán. Az egyénen belüli variabilitás itt is megfigyelhető: egy 1955-ös filmhíradóban ugyanaz a női beszélő mindkét alakváltozatot használja.



3. ábra *A* nejlon *és* nájlon *alakváltozatok előfordulási gyakorisága a filmhíradókban*

Következtetések

A három vizsgált szóalakpár nemcsak az abszolút gyakoriságát és az időbeli eloszlását tekintve tér el, hanem a variabilitás mintázataiban is. A filmhíradókban nem adatolható a G. Varga által 1968-ban a „választékos köznyelvben" megfigyelt kiegyenlített mintázat, a *Svájc* és a *sztrájk* esetében is mindvégig jelentősen gyakoribb az *áj*-t tartalmazó alak. Ezzel szemben a *nejlon/nájlon* alakpárnak az 1950-60-as évekre jellemző előfordulása kiegyenlített mintázatot mutat, ami nem vetíti előre a mai sztenderd forma előretörését.

Forró (2013) írott adatainak megoszlásával egyezően a toldalékolás a *Svejc* esetében neutrális. Habár Forró a *sztrejk* alakváltozat hiányát állapította meg, a filmhíradós korpusz alapján létező és egyes beszélők esetén gyakori előfordulású változatról van szó, amely szintén neutrálisan toldalékolódik.

A *sztrejk* az 1950-es, a *Svejc* pedig az 1970-es évektől oly mértékben visszahúzódik a bemondók ejtésében, ami alapján feltehető, hogy ez nem pusztán az egyéni ejtési mintázatok sajátosságaiból adódik, hanem valamilyen, a bemondókra vonatkozó előírásnak tudható be. Habár ezekkel a szóalakokkal kapcsolatban erre utaló forrást nem találtunk, ismeretes, hogy a filmhíradók bemondóinak a körében elvárás volt, hogy az idegen szavakat (közneveket) a magyar köznyelv hangkészletével ejtsék (vö. pl. Csók 1967), így például az illabiális *ȧ* ejtése kerülendő volt. Nem lehetetlen, hogy a *Svejc* és a *sztrejk* alakok visszaszorulása a filmhíradókban hasonló irányelvekkel függ össze.

Köszönetnyilvánítás

Köszönöm Kurutz Mártonnak, a Magyar Nemzeti Filmarchívum munkatársának, a bemondók azonosításában nyújtott segítségét.

Irodalom

Csók P. 1967. Az idegen szavak és nevek ejtése. In: Grétsy L. – Szathmári I. (szerk.): *Helyes kiejtés, szép magyar beszéd. Az egri kiejtési konferencia anyaga.* Budapest: Tankönyvkiadó. 210–213.

Forró O. 2013. *Ingadozás a magyar elölségi harmóniában. Szempontok a variabilitás szinkróniájának és diakróniájának feltárásához és értelmezéséhez.* PhD-disszertáció. Piliscsaba: Pázmány Péter Katolikus Egyetem.

G. Varga Gy. 1968. *Alakváltozatok a budapesti köznyelvben.* Nyelvészeti tanulmányok 10. Budapest: Akadémiai Kiadó.

Gocsál, Ákos 2024. Voices of the past: Research into old Hungarian newsreels. In: Bóna, J. (ed.): *HSCR 2024. Proceedings of the Sixth International Workshop on the History of Speech Communication Research. Budapest, September 20-21, 2024.* Dresden: TUDpress. 136–145. https://www.isca-archive.org/hscr_2024/gocsal24_hscr.pdf

TOWARD SPEAKER-INDEPENDENT ACOUSTIC-TO-ARTICULATORY INVERSION GUIDED BY ANATOMICAL GRID

Nhat-Nam Nguyen[1] - Pierre-André Vuissoz[2] - Yves Laprie[1]

[1]Université de Lorraine, CNRS, Inria, F-54000 Nancy, France
[2]Université de Lorraine, Inserm, IADI U1254, F-54000 Nancy, France

nhat-nam.nguyen@loria.fr, pa.vuissoz@chru-nancy.fr, yves.laprie@loria.fr

Introduction

Acoustic-to-articulatory inversion aims to reconstruct speech articulator gestures from the acoustic speech signal. Recent work has shown that complete vocal-tract contours can be reconstructed from real-time MRI data, producing coordinates for several articulators from the glottis to the lips [1]. However, these coordinates remain tied to the anatomy of a specific speaker, which may contribute to inversion errors if the system is used for multi-speaker inversion due to anatomical mismatch, particularly when no speaker-specific anatomical reference in the form of a midsagittal MRI image is available.

Prior work has used cross-speaker transformations to synthesize speaker-adapted vocal-tract MRI data [5] and thin-plate-spline normalization to reduce inter-speaker morphological differences in articulatory space [6, 7]. We address this issue as cross-speaker anatomical normalization, using an anatomically anchored vocal-tract grid as a shared reference for aligning speaker-specific contours.

Methods

For each speaker, contours were extracted from the same phonetic target, the vowel /u/, ensuring that cross-speaker comparisons were made under a matched articulatory condition.

The proposed normalization framework uses a curvilinear vocal-tract grid constructed from reproducible midsagittal anatomical anchors. This grid preserves speaker-specific vocal-tract geometry while providing structured correspondences across speakers.

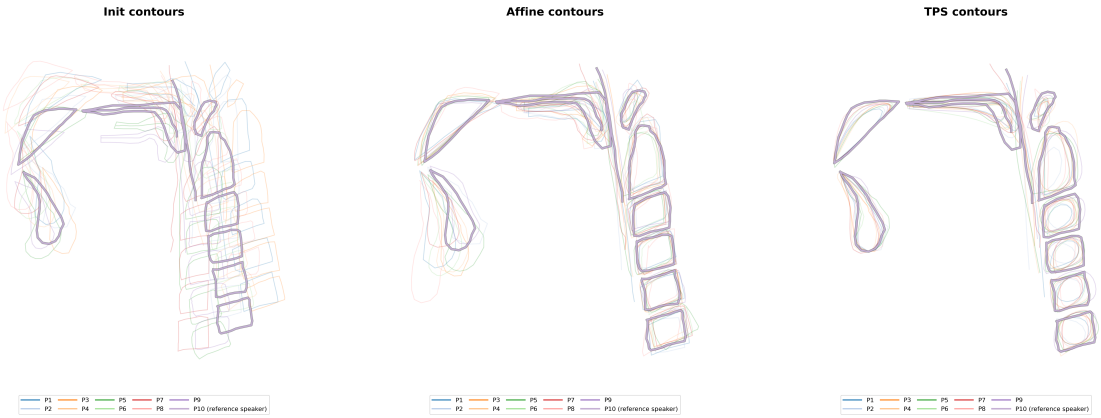
A fixed reference speaker is selected by choosing the candidate that maximizes Dice overlap [4, 3] between closed anatomical masks after affine alignment. After reference selection, all speakers are evaluated in three stages: the initial stage, corresponding to the original contours before normalization; the affine stage, in which contours are mapped into the selected reference space using the affine transform; and the affine-plus-TPS stage, in which TPS[2] refines the affine-normalized contours. We define Dice distance as $d_{\mathrm{Dice}}(A, B) = 1 - \mathrm{Dice}(A, B)$. For each stage, we compute the mean of d^2_{Dice} over all speaker pairs.

Results

Using the selected reference speaker, the mean squared pairwise Dice distance in the all-speaker cohort decreased from 0.7547 for the unnormalized initial contours to 0.1764 after affine alignment and 0.0975 after affine plus TPS. This corresponds to reductions of approximately 77% after affine alignment and 87% after the full affine-plus-TPS pipeline, relative to the initial stage. Thus, affine alignment removes most global mismatches, while TPS adds local refinement. Figure 1 illustrates the progressive reduction in cross-speaker contour variability after affine and affine-plus-TPS normalization.

Conclusion

Anatomically anchored grids provide an interpretable basis for normalizing vocal-tract shape across speakers. Under the proposed ROI-based Dice-distance criterion, the grid-guided affine-plus-TPS pipeline reduces cross-speaker anatomical variability for the vowel /u/. These results motivate the use of such grids as candidate intermediate representations between speaker-specific contour coordinates and speaker-independent acoustic-to-articulatory models. However, the present evaluation is restricted to a single phonetic target, /u/, and the observed reduction in anatomical variability cannot yet be assumed to generalize to other phonemes or linguistic contexts. Further validation is needed to determine whether this normalization improves phoneme-level articulatory consistency and downstream inversion performance.



(a) *Initial stage* **(b)** *After affine* **(c)** *After affine + TPS*

Figure 1. *All-speaker contour overlays before and after normalization. Quantitatively, the mean squared pairwise Dice distance is reduced by approximately 77% after affine alignment and by 87% after the full affine-plus-TPS pipeline.*

Acknowledgements

The authors thank the anonymous reviewers for their constructive comments.

References

[1] Sofiane Azzouz, Pierre-André Vuissoz, and Yves Laprie. Reconstruction of the complete vocal tract contour through acoustic to articulatory inversion using real-time MRI data. *arXiv preprint arXiv:2510.00914*, 2025.

[2] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[3] William R. Crum, Oscar Camara, and Derek L. G. Hill. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Transactions on Medical Imaging*, 25(11):1451–1461, 2006.

[4] Lee R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[5] Ioannis Douros, Asterios Toutios, Jangwon Kim, Shrikanth Narayanan, and Yves Laprie. Mri vocal tract sagittal slices estimation during speech production of cv. In *Proceedings of the European Signal Processing Conference (EUSIPCO)*, 2020.

[6] Jianguo Wei and Jianwu Dang. Vocal tract normalization in articulatory space using thin-plate spline method. *The Journal of the Acoustical Society of America*, 123(5):3885–3885, 2008.

[7] Jianguo Wei and Jianwu Dang. Morphological normalization of vocal tract shape. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 4186–4189, 2010.

VARIATION IN MID VOWELS PRODUCED BY SPEAKERS WITH T21: AN EXAMINATION OF TWO BRAZILIAN DIALECTS

Marian Oliveira[1] – Lucrécia de Aquino Santos[2] – Deise Santana Maia[3]

[1]Universidade Estadual do Sudoeste da Bahia – Núcleo Saber Down
[2]Universidade Estadual do Sudoeste da Bahia – Núcleo Saber Down
[3]Laboratoire CRIStAL – ULille

marian.oliveira@uesb.edu.br, 2021a0089@uesb.edu.br, deise.santanamaia@univ-lille.fr

Introduction

Camara Jr. (1992) proposes a seven-vowel system for Brazilian Portuguese (BP) in stressed position (/a, i, u, e, o, ɛ, ɔ/), which undergoes successive neutralizations in unstressed contexts, reducing the inventory to five, four, and three vowels. In pretonic position, mid vowels may be realized through maintenance [e, o], lowering [ɛ, ɔ], or raising [i, u]. Previous studies conducted in Bahia (Pacheco, Oliveira & Ribeiro, 2013; Santos, 2020) indicate that lowering is the predominant pattern in Vitória da Conquista, consistent with Nascentes' (1953) observation that speakers from Brazil's North and Northeast favor [+open] realizations.

Given that variation in mid-vowel height functions as a salient dialectal marker in Bahia, members of this community — including individuals with disabilities — are expected to preserve such features. The present study therefore examines, through acoustic analysis, the production of pretonic mid vowels by speakers with Trisomy 21 (T21) in Vitória da Conquista and Porto Seguro. T21, or Down syndrome, is a genetic condition characterized by an extra chromosome and is often associated with phonoarticulatory alterations (Oliveira, 2011). Despite these alterations, we hypothesize that speakers with T21 maintain both the phonological distinctions and the dialectal traits of their communities.

The phonological interpretation follows Autosegmental Phonology, particularly Wetzels' (1992) model for mid vowels in Portuguese. The working hypothesis is that, regardless of oromyofunctional alterations, participants reproduce the dialectal features characteristic of their speech community. Thus, if the local variety tends to exhibit open mid vowels in pretonic position, speakers with T21 are expected to follow the same pattern.

Methods

The study examined speech data from eight participants from two cities in Bahia — Vitória da Conquista and Porto Seguro — equally distributed by sex (male/female) and condition (with/without T21). A corpus of words containing open and close-mid vowels, both front and back, in pretonic, tonic, and word-final unstressed positions was elicited through a PowerPoint slideshow, with each item repeated five times. Participants were affiliated with the Research and Study Center on Down Syndrome (Saber Down – UESB), and the study was approved by the Brazilian Ethics Committee.

Data collection took place in an acoustic booth, and instrumental analysis focused on the extraction of F1 and F2 formants using Praat (Boersma & Weenink, 2026). The dataset was organized in Excel and subjected to statistical analysis in RStudio. The dependent variables comprised the formants of pretonic mid vowels [e, o] and the seven stressed oral vowels [i, e, ɛ, a, ɔ, o, u]. The independent variables included syllabic environment (pretonic vs. tonic), vowel type, geographical location (Vitória da Conquista vs. Porto Seguro), sex/gender (male vs. female), and speaker condition (typical vs. atypical).

Results

The results indicate that, phonetically, women tend to produce more open vowels, as reflected in higher F1 values, whereas men exhibit lower F1 values, indicating reduced openness. From a dialectal and phonological perspective, speakers with T21 show patterns consistent with regional expectations: in pretonic position, the mid vowels [e, o] frequently acquire the feature [+low] and are realized as [ɛ, ɔ] in most occurrences. Contextual analysis also reveals assimilation to the height of the tonic vowel. Although lowering predominates — as expected for the dialect — maintenance [e, o] also occurs in some contexts. Overall, the findings confirm the initial hypothesis: phonoarticulatory alterations associated with T21 do not prevent speakers from maintaining features characteristic of their local speech community.

The social variable of locality emerges as a significant factor in vowel quality. The estimated intercept indicates that vowels produced in Porto Seguro exhibit lower formant frequencies and are therefore higher than those produced in Vitória da Conquista, with a difference of 52.49 Hz. Moreover, the analysis revealed differences in vowel height between male and female speakers, with higher F1 values among women, corresponding to a difference of 103.46 Hz. From a physiological perspective, the smaller size of the female vocal tract results in higher frequency values. From a linguistic perspective, however, such differences may carry distinct social meanings. Sociolinguistic research has suggested that women often display more conservative linguistic behavior than men, favoring socially prestigious forms.

The differences between the two localities are evident (Figure 1; Table 2), with vowels in Vitória da Conquista being lower. Regarding speaker condition, atypical participants (T21) align with the linguistic patterns of their respective communities. This finding reinforces a central tenet of sociolinguistics: linguistic variation is systematically constrained rather than random. In other words, even in the presence of significant oromyofunctional alterations, speakers with T21 are able to adjust their linguistic trajectories so as to preserve salient features of the language in general and of their dialect in particular.

Conclusion

The results demonstrate that mid vowels in pretonic position may either be preserved or undergo lowering, a process observed among speakers both with and without T21. These findings suggest that, despite the oromyofunctional alterations associated with the syndrome, individuals with T21 are able to maintain phonetic and phonological features characteristic of their linguistic communities. This evidence strengthens the hypothesis that sociolinguistic factors play a decisive role in the development and maintenance of speech patterns, irrespective of speakers' genetic condition.

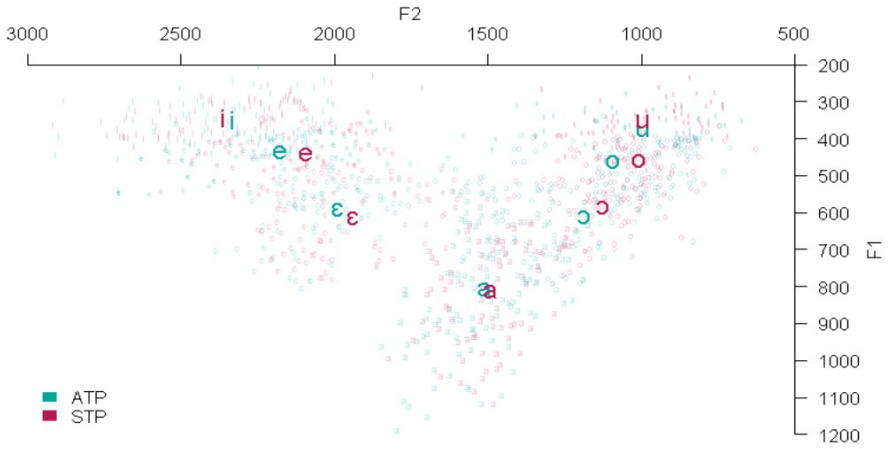


Figure 1. Vowel space of oral vowels produced by speakers with and without T21, from Porto Seguro (PSG) and Vitória da Conquista (CON)

	[i]	**[e]**	**[ɛ]**	**[ɔ]**	**[o]**	**[u]**
Vitória da Conquista (CON)	▲	■	●	●	■	▲
Porto Seguro (PSG)	▲	■	●	●	■	▲

Note: ● = highest frequency; ■ = medium frequency; ▲ = lowest frequency.

Table 2. Distribution of Variants by Locality and Number of Occurrences

Acknowledgements

References

Boersma, Paul – Weenink, David 2026. Praat: doing phonetics by computer [Computer program]. Version 6.4.62, retrieved 13 March 2026 from https://praat.org.

Camara Jr., J. M. 1992. Estrutura da língua portuguesa. 21st ed. Vozes.

Kent, R. D. – Read, C. 1992. Acoustic analyses of speech. Singular Publishing Group.

Nascentes, A. 1953. O linguajar carioca. Organização Simões.

Oliveira, M. 2011. Sobre a produção vocálica na síndrome de Down: Descrição acústica e inferências articulatórias. Doctoral thesis, Universidade Estadual de Campinas.

Pacheco, V. – Oliveira, M. – Ribeiro, P. J. 2013. Em busca da melodia nordestina: As vogais médias pretônicas de um dialeto baiano. Linguística 29. 165–187.

Santos, L. A. 2020. Análise acústica de vogais orais produzidas por sujeitos com T21: Um contraste entre falantes conquistenses e porto-segurenses. Master's thesis, Universidade Estadual do Sudoeste da Bahia.

Wetzels, L. 1992. Mid vowel neutralization in Brazilian Portuguese. Cadernos de Estudos Linguísticos 23. 19–55.

MONDHATJUK-E, HOGY „A MAGYARBAN” MEGKÜLÖNBÖZTETÜNK HOSSZÚ ÉS RÖVID FELSŐ NYELVÁLLÁSÚ MAGÁNHANGZÓKAT?

Oszkó Beatrix[1,2] – Vargha Fruzsina Sára[1] – Várnai Zsuzsa[1]

[1]ELTE Nyelvtudományi Kutatóközpont

[2]Újvidéki Egyetem

oszko.beatrix@nytud.elte.hu, oszkobea@gmail.com vargha.fruzsina@nytud.elte.hu, varnai.zsuzsa@nytud.elte.hu

1. Bevezetés

A 2025 októberében rendezett Beszédkutatás konferencián a Deme Andrea és kollégái által a magánhangzók hosszúságáról tartott előadást követő vitában szóba került, hogy a felső nyelvállású magánhangzóknál még a nyelvész kollégák körében is állandó nehézséget okoz a hosszúság, illetve rövidség írásban való jelölése. Vagyis – nyelvérzékük alapján – nem mindig tudják biztonsággal eldönteni, hogy egy adott szó adott alakját hosszú vagy rövid ékezettel helyes-e leírni. A hosszú *í, ú, ű* használatában való bizonytalanság (az *ly* mellett) a legáltalánosabb forrása az iskolai dolgozatok helyesírási hibáinak is.

Felmerül a kérdés, hogy itt a magyar nyelvre általánosan érvényes jelenségről van-e szó. Vannak olyan nyelvi sajátosságok a magyarban, amelyek valamikor az egész nyelvterületre jellemzők voltak, de mára csak bizonyos nyelvváltozatokban vannak jelen. Ilyen (az írásban általában jelöletlen) zárt *ë* fonéma, amely a magyar nyelvterület nagy részén a mai napig megjelenik a beszélők nyelvhasználatában, és ilyen az *ly* is, amely a kiejtésben ma már csak nyomokban bukkan föl palóc vagy moldvai beszélőknél. A hosszú felső nyelvállású magánhangzók esetében feltehető, hogy egy olyan hangváltozásról van szó, amely a nyelvterület különböző részeit különböző mértékben érintette, illetve bizonyos esetekben a terület egy jelentős részén egyáltalán nem jelent meg (Abaffy 2003: 331), hasonlóképpen az *ö*-zéshez (*Szöged, vödör, mögvötte*) vagy az *í*-zéshez (*szíp, legíny, nígy*), amelyek szintén nem jelentkeznek (legalábbis nem azonos mértékben) az egész nyelvterületen.

2. Módszer

A felső nyelvállású magánhangzók ejtését különböző helyzetekben A magyar nyelvjárások atlasza és A romániai magyar nyelvjárások atlasza integrált térképlapjai alapján vizsgáljuk a teljes nyelvterületen. A térbeli változatosság feltárása mellett vizsgáljuk a helyesírási szabályozás ezen magánhangzókra vonatkozó részeit, a szabályozást alakító általános alapelveket, nyelvészeti ideológiákat, illetve a szabályozás változását az 1900-as évek elejétől kezdődően.

3. Eredmények

A nyelvföldrajzi térképek tanúsága alapján a hosszú magánhangzók használatának egyaránt van lexikai, fontokatikai és földrajzi meghatározottsága. Az 1. ábra a *búza* első szótagi magánhangzójának időtartamát szemlélteti az atlaszadatok fonetikus lejegyzése alapján. A Dunántúl nagy részén, az északkeleti kutatópontokon, és sok helyütt a mezőségi, székely és moldvai nyelvjárásokban a rövid ejtés a jellemző. A második ábra alapján, ahol az *odú* szó *ú*-jának ejtésbeli variabilitását látjuk, már csak a nyelvterület középső részén jellemző a hosszú magánhangzó használata. Feltételezhető, hogy szóvégi helyzetben a rövid ejtésmód általánosabban jelentkezik.

Az 1950 előtti helyesírási szabályozásban is megjelent ugyan a rövid-hosszú szembenállás a felső nyelvállású magánhangzóknál, azonban egyfelől a szabályzatok megalkotói nem törekedtek egyöntetűségre már a szószedetben megadott példákban sem (a *mindig* és a *mindíg*

pl. egyaránt helyes), illetve nem írták elő a megkülönböztetést kötelező érvénnyel, hanem ajánlásként fogalmazták meg. Ebben megtartották azt a hozzáállást, amit még Simonyi Zsigmond fogalmazott meg 1903-as Az új helyesírás című, iskolai felhasználásra szánt helyesírási útmutatójában, ahol amellett érvel, hogy különösen szóvégi helyzetben, illetve olyan szavakban, ahol ingadozás van a nyelvterület különböző részei között (pl.: *nyújt, sújt, gyújt*) nem szükséges a hosszú magánhangzó írásbeli használata (Simonyi 1903: 30). Külön megemlíti azt is, hogy a dunántúli iskolásoktól egyáltalán nem várható el a felső nyelvállású magánhangzóknál a hosszú alakok jelölése, mivel az az ő nyelvjárásuktól idegen.

Az 1950-es években gyökeres szemléletváltás következik be a nyelvi változatossággal kapcsolatosan. Ahogyan Deme László fogalmaz: „A nemzeti nyelv, Sztálin elvtárs megfogalmazása szerint, egységes az egész nemzet számára, s közös annak minden tagja számára" (Deme 1953a: 28). Az egységességet Deme és munkatársai úgy képzelik el, hogy mindenhol egyváltozatúságra törekednek, kötelezően előírva a használandó szóalakokat. Így ahol a hosszú ékezetes írásmódot írják elő, ott kötelező lesz mindenkinek úgy írni. Sőt, Deme még azt is szorgalmazza, hogy az írásban jelölt időtartamot várják el a kiejtésben is. A hosszú felső nyelvállású magánhangzók megkülönböztetésének fontosságát Deme azzal indokolja, hogy a megkülönböztetés és annak írásban való kötelező jelölése hiányában „három önálló hangeszköz" kiveszne a magyar nyelvből (Deme 1953b: 213).

4. Következtetések

A felső nyelvállású magánhangzók helyesírásában megmutatkozó bizonytalanság hátterében minden bizonnyal az áll, hogy ebben a tekintetben a magyar nyelvet nagyfokú térbeli változatosság jellemzi. A helyesírási szabályzat azonban – az 1950-es 9. kiadástól kezdődően – egyváltozatúságra törekszik, és kötelezően előírja, hogy mely szavakban kell a felső nyelvállású magánhangzókat hosszú ékezettel írni. Előadásunkban amellett érvelünk, hogy társadalmi haszonnal járna a helyesírási szabályozásban átgondolni a felső nyelvállású magánhangzók kötelező és minden szóalak esetében csak egyféle változatra korlátozódó jelölését. Támogatást jelenthetne ehhez, ha a nyelvészeti kutatásokban a jelenleginél nagyobb teret kapna a nyelvi változatosság felismerése, és ezzel összefüggésben a nyelvi tények alaposabb feltárása. A Budapesti Szociolingvisztikai Interjúhoz kapcsolódó kutatásokból például kiderül, hogy nem beszélhetünk egységes nyelvváltozatról Budapest esetében sem (Borbély–Kontra: 2021), így nem meglepő, hogy a magukat köznyelvi beszélőként definiáló budapesti értelmiségiek nyelvhasználatát is befolyásolhatja a (sokszor nem dokumentált, irrelevánsnak tartott) nyelvjárási háttér, például a felső nyelvállású magánhangzók időtartama tekintetében.

Irodalom

E. Abaffy Erzsébet 2003. Hangtörténet. In: Kiss Jenő – Pusztai Ferenc (szerk.): *Magyar nyelvtörténet.* Budapest: Osiris Kiadó. 301–351.

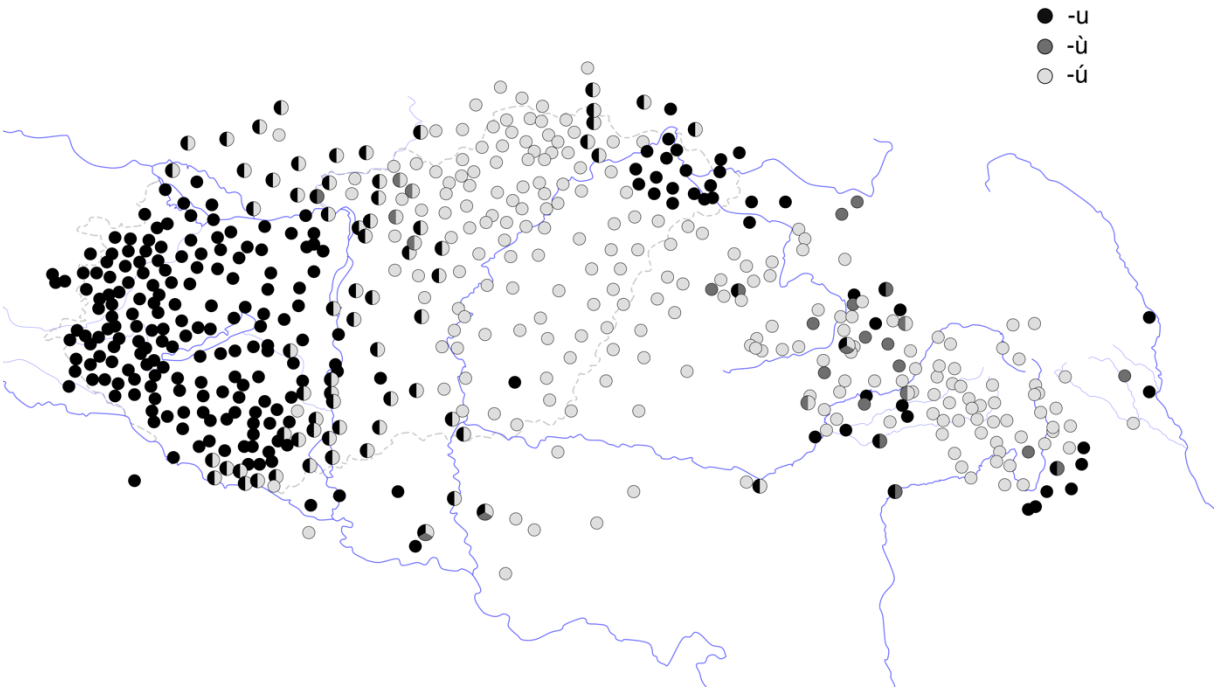
Deme Andrea – Szánthó Zsuzsa – Gráczi Tekla Etelka 2025. A magánhangzók hosszúsági szembenállása a nyelvi környezet függvényében. Előadás a Beszédkutatás – Speech Research konferencián, 2025. október 10-én.

Deme László 1953a. A nemzeti nyelv rétegei a helyesség szemszögéből. In: Lőrincze Lajos (szerk.): *Nyelvművelésünk főbb kérdései.* Budapest: Akadémiai Kiadó. 15–48.

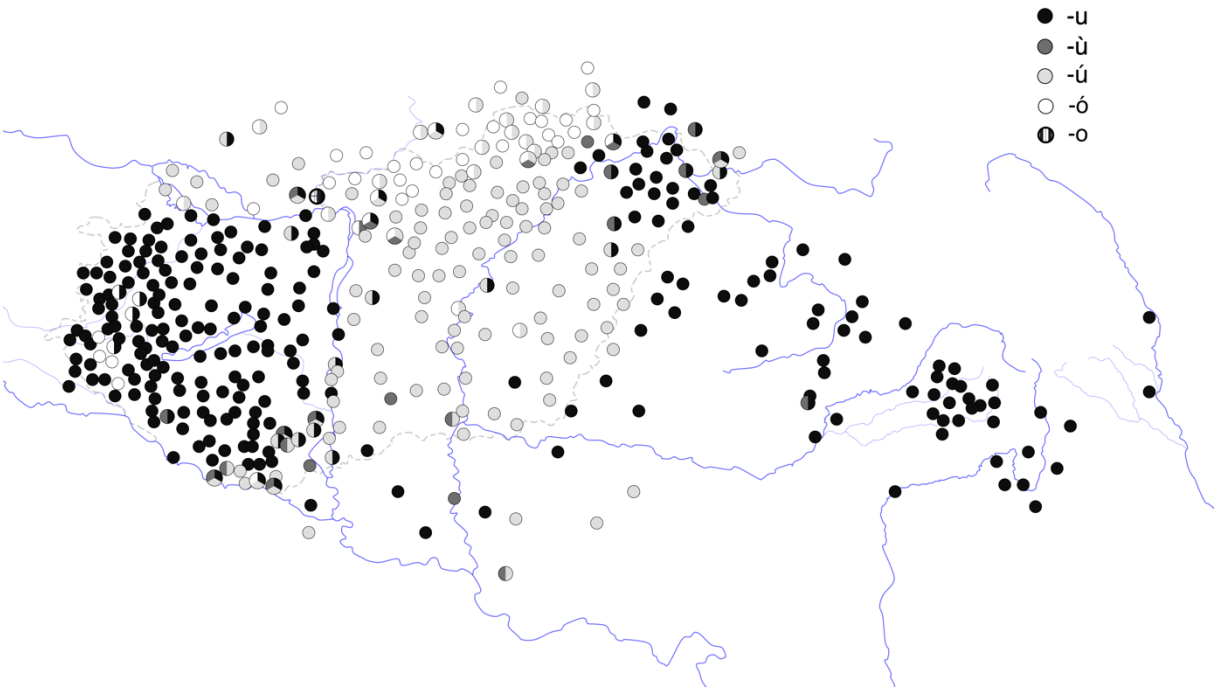
Deme László 1953b. A helyes magyar kiejtés kérdése. In: Lőrincze Lajos (szerk.): *Nyelvművelésünk főbb kérdései*. Budapest: Akadémiai Kiadó. 199–239.

Kontra Miklós – Borbély Anna (szerk.) 2021. Tanulmányok a budapesti beszédről a Budapesti Szociolingvisztikai Interjú alapján. Budapest: Gondolat.

Simonyi Zsigmond 1903. Az új helyesírás. A Vallás- és Oktatásügyi m. kir. miniszter 1903. március 14-én 862. szám alatt kelt rendeletével kiadott magyar iskolai helyesírás. Nyelvészeti Füzetek 5.



1. ábra *A* búza *első szótagi magánhangzójának időtartama*



2. ábra *Az* odú *második szótagi magánhangzójának időtartama*

INTRASPEAKER VARIABILITY OF NASALANCE SCORES: DATA FROM GREEK

Ioannis Papakyritsis[1] -Emilia Michou[1]

[1] Department of Speech and Language Therapy, University of Patras, Greece

papakyritsis@upatras.gr, emiliamichou@upatras.gr

Introduction

Hypernasality, excessive emission of sound through the nose during speech, is a common symptom of velopharyngeal dysfunction caused by neurological, structural or functional disorders, that can affect the overall naturalness and intelligibility of speech output (Kummer, 2008). Nasalance is an acoustic measure of the relative contribution of the sound pressure level of the nasal speech signal to the total sound pressure level, that is used to quantify oral-nasal balance and related disorders (see Bressman, 2021 for a review). Treatment protocols targeting the reduction of hypernasality incorporate principles of motor learning (Maas et al., 2008). Quantification and real-time visualization of nasalance via the nasometer and other alternative devices, allows the systematic use of augmented, clinician-guided feedback, which is a key motor learning principle.

This study investigates the clinical utility of a portable version of the nasometer device that uses a handheld plate for the separation of the oral and nasal microphones. While the device has been marketed as a biofeedback instrument, the manufacturers advise against its use for diagnostic purposes due to possible acoustic spillover between oral and nasal signals. The study aim is to measure intra-speaker variation of the nasalance scores generated by the device in order to assess its effectiveness as a biofeedback device.

Methods

Participants were 60 young Greek speakers aged 20-30 years old, and 20 older adults over 70 years old. Participants were drawn from a larger study aimed to derive nasalance norms for Greek adults speakers. Participants read aloud a corpus of oral and nasal linguistic material of varying complexity, syllables, sentences and texts, adapted from Okalidou, Karathanasi & Grigoraki (2011), using a portable nasometer device. The material was adjusted to involve mulitiple repetitions of each syllable, sentence and text to allow the estimation of intraspeaker variability. Additionally, in order to identify specific sources of speech variation, a subpart of the sentence material was produced in three speaking rate conditions, habitual, slow and fast; three attempts at each target rate were carried out. Nasalance score variability measures were calculated.

Results

Overall, the Intraclass Correlation Coefficient (ICC) measurements indicated high intraspeaker reliability of the nasality scores, acceptable for clinical use. Table 1 presents preliminary findings from the female young speakers group (N=40). ICC was close to 0.9 for most material; only bilabial sentences obtained ICC <0.7. Similarly, ongoing analysis shows relatively small effects of speaking rate on nasalance score variation. The analysis of the data from the older speaker group is ongoing.

NASALITY SCORES	SYLLABLES				
	/pa/	/ta/	/ka/	/sa/	/ma/
MEAN	18.24%	18.50%	18.76%	19.57%	54.71%
SD	8.64%	8.90%	8.40%	10.03%	11.30%
ICC	0.861	0.845	0.871	0.897	0.919
	SENTENCES				
	Bilabial	Alveolar	Velar	Sibilant	Nasal
MEAN	13.66%	17.72%	19.27%	14.87%	43.14%
SD	3.77%	5.1%	5.21%	6.33%	8.58%
ICC	0.567-0.711	0.862-0.906	0.823-0.900	0.813-0.943	0.906-0.946
	TEXTS				
	Oral	Oronasal	Nasal		
MEAN	17.02%	26.38%	43.17%		
SD	5.99%	8.15%	7.32%		
ICC	0.772-0.947	0.898-0.963	0.834-0.947		

Table 1. *Nasality score metrics (mean, SD, ICC) for the young female group*

Conclusions

Conclusively, analysis shows acceptable levels of intraspeaker variability of nasalance scores generated by a portable instrument used for measuring oral-nasal balance. Thus, the handheld nasometer device is a potentially appropriate clinical tool that can be effective as biofeedback for patients with nasalance problems.

References

Kummer, A.W. (2008) Cleft Palate and Craniofacial Anomalies: Effects on Speech and Resonance (2nd ed.). Clifton Park, NY: Delmar.

Maas, E., Robin, D.A., Hula, S.N.A., Freedman, S.E., Wulf, G., Ballard, K.J., & Schmidt, R.A. (2008). Principles of Motor Learning in Treatment of Motor Speech Disorders. *American Journal of Speech - Language Pathology*, *17* (3): 277-298.

Okalidou, A., Karathanasi, A., & Grigoraki, E. (2011). Nasalance norms in Greek adults. Clinical Linguistics and Phonetics, 25(8), 671–688.

Bressmann, T. (2021). Nasometry. In M. J. Ball (Ed.), *Manual of Clinical Phonetics* (1st ed., pp. 385–399).

VARIATION AND PRIMING IN THE HUNGARIAN 2SG SUBJUNCTIVE

Henrietta Pokk[1] – Péter Rácz[1]

[1] Budapest University of Technology and Economics

henriettapokk@edu.bme.hu, racz.peter.marton@ttk.bme.hu

The Hungarian 2SG subjunctive varies between long and short forms (e.g. indefinite *várjál/várj, kérjél/kérj;* definite *várjad/várd*, *kérjed/kérd*). Within the definite set, vowel harmony creates an asymmetry in surface similarity across the paradigm: in back stem verbs, the long 2SG.SBJV definite (*várjad*) closely resembles the 3SG indicative definite (*várja*), while in front stem verbs this similarity does not hold (*kérjed* vs. *kéri*). This asymmetry is much weaker for the indefinite forms (Rebrus 2025), consequently this study focuses on definite 2SG.SBJV forms only (see Table 1). This allows us to ask whether this variation is primarily driven by the frequency of definite / indefinite forms, or whether phonological similarity also plays a role.

	3SG.INDEF	**3SG.DEF**	**2SG.SBJV.DEF**	2SG.SBJV.INDEF
back (*vár*)	**vár**	**várja**	**várd/várjad**	várj / várjál
front (*kér*)	**kér**	**kéri**	**kérd/kérjed**	kérj / kérjél

Table 1: *Forms of* vár *'wait' and* kér *'ask', boldface forms were included in the design*

Webcorpus data (Nemeskey 2020) show evidence for frequency effects, but not phonological similarity effects, on 2SG.SBJV variation. We wanted to test this pattern in speaker variation without the confounds of style, register, and communicative pressures that shape the corpus data. We designed a visual word priming experiment, run online on Pavlovia, in which a prime word appeared and disappeared before each target; participants made a lexical decision on the target, pressing "j" if it was a real, familiar Hungarian word form and "f" if it was not. We compare two effects: morphological similarity, which operates both in front and back vowel stems, and phonological similarity, which is only relevant for back vowel stems. The four conditions are the following: same-cell primes (2sg subjunctive short, e.g. *várjad→várd*) were contrasted with phonologically similar cross-cell primes (3SG indicative, e.g. *várjad→várja*), crossed with vowel harmony class. Front stem pairs (*kérjed→kérd* vs. *kérjed→kéri*), equidistant in phonological form, served as a control to delineate competition from surface similarity effects. The stimulus set comprised 48 verb lemmas, counterbalanced across vowel harmony classes (back vs. front) and stem structure (monosyllabic vs. polysyllabic). To ensure attention and valid responding, we included attention check trials (prime = 2SG.SBJV long, target = scrambled nonword). The design has been preregistered (https://aspredicted.org/92c82r.pdf).

Under the null hypothesis (H0), reaction times should not differ across conditions. We had two main hypotheses. Under the morphological blocking hypothesis (H1), same-cell primes should produce slower reaction times compared to cross-cell primes indicating morphological competition within the paradigm cell. Under the phonological facilitation hypothesis (H2), the back form-similar condition should produce faster reaction times than all other conditions suggesting that phonological similarity facilitates access to the target form even across paradigm cells.

Using generalized linear mixed effects modeling (Brooks et al. 2017) with crossed random intercepts for participants and items, we found that back cross-cell primes (*várjad/várja*) elicited significantly slower reaction times compared to back same-cell primes (*várjad/várd*; β = 0.090, SE = 0.025). Front cross-cell primes (*kérjed/kéri*) also elicited significantly slower reaction times relative to the back same-cell baseline (β = 0.069, SE = 0.031). Front same-cell primes (*kérjed/kérd*) showed no significant difference from the baseline (β = 0.010, SE = 0.031).

These findings are particularly interesting for two reasons. First, we expected that same-cell primes would elicit slower reaction times due to competition within the paradigm, yet we observe the opposite pattern: cross-cell primes elicited significantly slower reaction times than same-cell primes, for both back-vowel (*várjad/várja* vs. *várjad/várd*) and front-vowel (*kérjed/kéri* vs. *kérjed/kérd*) items (see Figure 1 below). This suggests that shared paradigm cell membership, rather than phonological similarity, is what drives facilitation. Second, contrary to our expectations, the phonologically similar back cross-cell condition did not stand apart from the other conditions: front cross-cell primes, which do not share the back stems' phonological overlap with their targets, showed a comparable slowdown. This shows that cross-cell primes lead to slower responses regardless of whether they resemble the target phonologically.

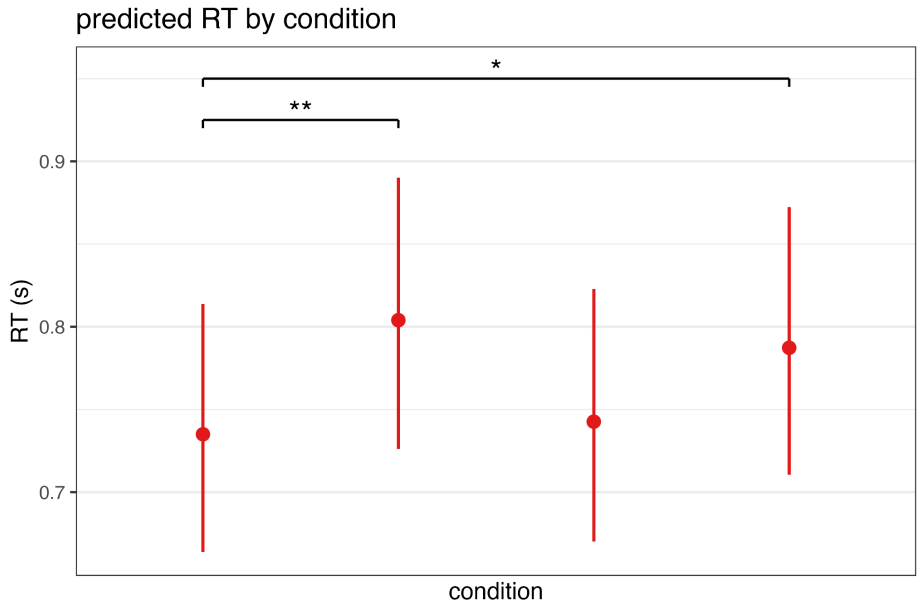


Figure 1. *Predicted reaction times by condition (in order: back same-cell, back cross-cell, front same-cell, front cross-cell)*

References

Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Maechler M, Bolker BM (2017). "glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling." *The R Journal*, **9**(2), 378–400. doi:10.32614/RJ-2017-066.

Nemeskey, Dávid Márk (2020). *Natural Language Processing methods for Language Modeling*. PhD thesis. Eötvös Loránd University.

Rebrus, Péter (2025). *A fonológiai mintázatok paradigmatikus alapjai*. Akadémiai nagydoktori mű, Eötvös Loránd Tudományegyetem.

EFFECTS OF SPEECH RATE MANIPULATION IN DUTCH WORD-INITIAL CLUSTERS: AGE-RELATED AND INDIVIDUAL VARIABILITY

Teja Rebernik[1,2] – Katharina Polsterer[3,4] – Valentine Lucquiault[3] – Valerie Querner[3] – Martijn Wieling[2,3] – Anne Hermes[1]

[1]Laboratoire de Phonétique et Phonologie (CNRS & Université Sorbonne Nouvelle)
[2]Brussels Centre for Language Studies (Vrije Universiteit Brussel)
[3]Center for Language and Cognition Groningen (Rijksuniversiteit Groningen)
[4]Acoustics Research Institute (Austrian Academy of Sciences)

teja.rebernik@vub.be, k.m.polsterer@rug.nl, v.lucquiault@rug.nl, v.f.j.querner@rug.nl, m.b.wieling@rug.nl, anne.hermes@sorbonne-nouvelle.fr

1. Introduction

The timing and coordination of word-initial clusters can reveal the underlying syllable structure, with complex onsets displaying tighter and more stable coordination, and non-complex onsets displaying looser and more variable coordination (e.g., American English: Marin & Pouplier, 2010; Italian: Hermes et al., 2013; French: Rebernik et al., 2025a). In the current study, we examine the stability of these word-initial clusters under speech rate manipulations, specifically focusing on age-related effects. While no existing study has looked at age-related articulatory timing adaptations to speech rate manipulations, one prior study has shown that older speakers have more stable and consistent coordination patterns in German complex onsets /pl, kl/ compared to younger speakers (Mücke et al., 2020).

2. Methods

We collected data of 66 Dutch native speakers (32 male, 34 female), between 20 and 76 years. They were recruited at the Noorderzon performing arts festival in Groningen (the Netherlands) and tested inside the sound-dampened booth of SPRAAKLAB, the mobile laboratory of the Faculty of Arts, University of Groningen (Rebernik et al., 2025b). Kinematic data was collected using the NDI VOX EMA, with a minimal setup including reference sensors and two sensors placed just behind the tongue tip (TT) and on the vermillion border of the lower lip (LL). Acoustic data was collected concurrently using a headset microphone. Speech rate was manipulated using a visual cue, namely a red box moving across a screen, spanning ten steps from 0.5 to 2 seconds (similar to Hermes et al., 2020). For each rate step, speakers produced the carrier phrase "Zij heeft [target word] gezegd" three times, each time with one of the three target words "plakken" (*transl: to stick*), "smakken" (*transl: to smack (lips)*) (both real Dutch words) and "fnakken" (a nonce word). For the acoustic analysis, we annotated the target words in Praat; for the ongoing kinematic analysis, we used Mark Tiede's *mview* to annotate the lowest positions of the LL during the vowel /e/ preceding the target word and the vowel /a/ of the target word, and extracted the full TT and LL trajectories. This allowed us to visualise the change in articulatory coordination over time.

3. Results

Preliminary acoustic analysis was conducted on twelve speakers, including six younger speakers (aged 20-25), and six older speakers (aged 59-73), with an equal number of male and female speakers. Overall, the speech rate manipulation task was successful, with most speakers decreasing their word duration with increasing speed (exceptions include speaker NDZ41, as seen in Figure 1). However, speakers reacted to speech rate manipulations in different ways, in line with prior studies showing speaker-specific speech rate effects on articulatory timing: some (e.g., NDZ06) are fairly linear, with each consecutive production faster than the previous, while others (e.g., NDZ16) are non-linear with large trial-to-trial variability. This speaker sample shows no age-related effects in duration, variability or the ability to adapt to the speech rate manipulations. However, we do see an age-related differentiation between the different

word-initial clusters, with younger speakers maintaining similar timing for the three clusters /pl, sm, fn/, while the older speakers group /sm, fn/ and distinguish them from /pl/. This may reflect syllable structure-dependent adaptation. Further in-depth articulatory data analysis, shedding light on inter-articulatory timing and coordination, will be presented at the workshop.

4. Conclusion

This study investigates the effects of speech rate manipulation on word-initial cluster coordination in Dutch speakers across different age groups. Preliminary results show no age-related effects on duration or variability in adapting to speech rate. However, differences seem to emerge in the grouping of word-initial clusters, with younger speakers maintaining similar durations across all clusters (/pl/, /sm/, /fn/), while the older speakers group /sm/ and /fn/ separately from /pl/. These findings suggest that aging may influence the coordination in a syllable structure-dependent manner: in Dutch, /pl/ is uncontroversially considered as true complex onset, while /s/-clusters are not considered complex onsets (Fikkert & Freitas, 2004; van der Hulst, 1984). Future analyses of kinematic data will provide deeper insights into the inter-articulatory timing and gestural coordination underlying these patterns.

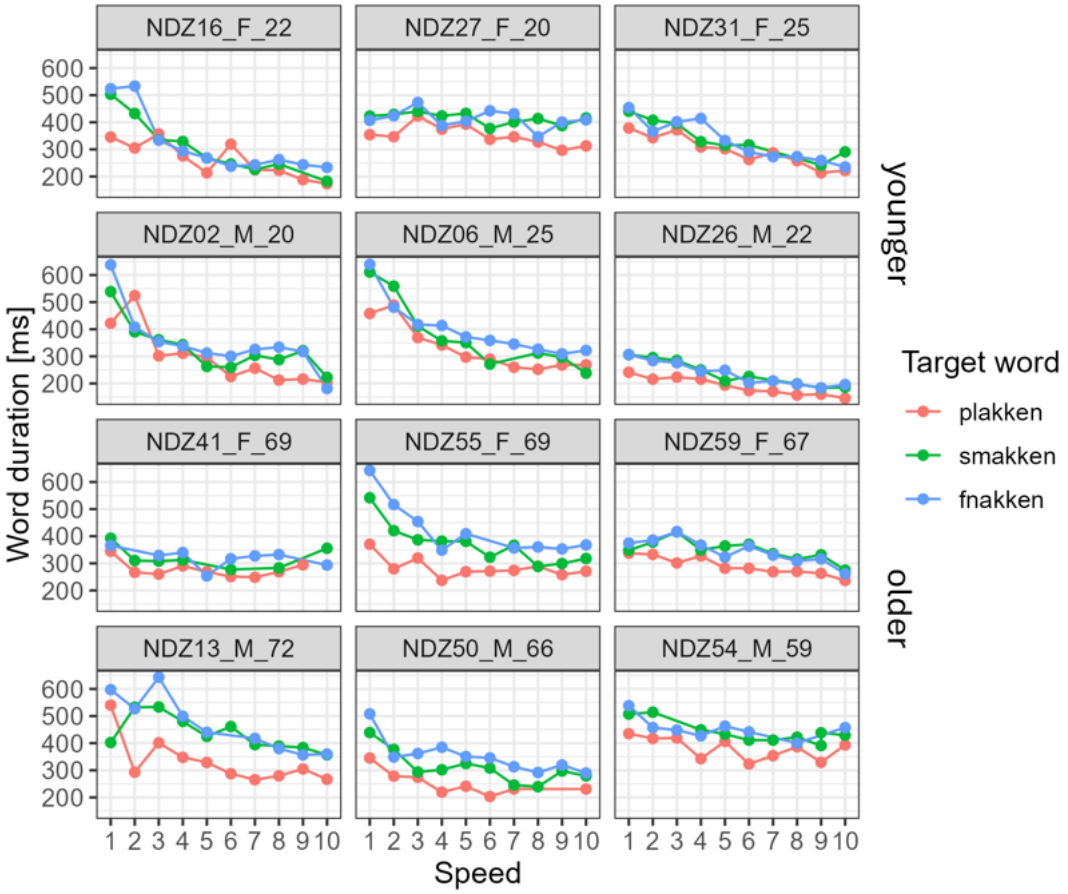


Figure 1. *Effects of speech manipulation on (acoustic) target word duration in milliseconds. The upper two panels present younger speakers, while the lower two panels present older speakers. F: female, M: male; speakers' age is indicated after speaker sex.*

Acknowledgments

We would like to thank the festival visitors for participating in our study as well as members of Speech Lab Groningen for helping with recruitment. We would also like to acknowledge funding from the ANR JCJC INSPECT project (ANR-23-CE28-0017), which made this work possible.

References

Fikkert, J.P.M. – Freitas, M.J. 2004. The role of language-specific phonotactics in the acquisition of onset clusters. *Linguistics in the Netherlands* 21. 58–68.

Hermes, A. – Mücke, D. – Grice, M. 2013. Gestural coordination of Italian word-initial clusters: the case of 'impure s'. *Phonology* 30/1. 1–25.

Hermes, A. – Tilsen, S. – Ridouane, R. 2020. Cross-linguistic timing contrast in geminates: A rate-independent perspective. *12th International Seminar on Speech Production* (December 2020, New Haven, United States)

Marin, S. – Pouplier, M. 2010. Temporal Organization of Complex Onsets and Codas in American English: Testing the Predictions of a Gestural Coupling Model. *Motor Control* 14/3. 380–407.

Mücke, D. – Hermes, A. – Tilsen, S. 2020. Incongruencies between phonological theory and phonetic measurement. *Phonology* 37/1. 133–170.

Rebernik, T. – Rialland, A. – Buech, P. – Hermes, A. 2025a. Tight and loose coordination as a marker of complex onsets and extrasyllabicity in French. *22èmes Rencontres du Réseau Francais de Phonologie* (1-3 July 2025, Agadir, Morocco).

Rebernik, T. – Jacobi, J. – Buurke, R. – Tienkamp, T. B. – Abur, D. – Wieling, M. 2025b. SPRAAKLAB - mobile laboratory for speech recorded acoustically and kinematically. *Behavior Research Methods* 57/205.

Van der Hulst, H. 1984. *Syllable Structure and Stress in Dutch*. Leiden: Foris Publications.

VARIÁCIÓ A SZERBHORVÁT *E*-KONJUGÁCIÓS IGÉK JELEN IDEJŰ ALAKJAIBAN

Rebrus Péter[1] – Tóth Szilárd[2]

[1]ELTE Nyelvtudományi Kutatóközpont
[2]ELTE Bölcsészettudományi Kar

rebrus@nytud.hu, sztothilard@student.elte.hu

A szerbhorvátban az *e*-konjugációs osztályba tartozó tövek, amelyek veláris mássalhangzóra végződnek, jelen idejű alakjai három paradigmatípusba sorolhatók aszerint, hogy a tematikus magánhangzók kiváltanak-e tőben morfofonológiai váltakozást, esetünkben az ún. velárisok palatalizációját, vagyis a velárisok posztalveoláris mutációját. A három mintát az 1. táblázat mutatja be: az 1SG és 3PL viselkedése eltér a többi cellától. A *moći* 'possum' ige egyedi esetében [1] mintában az 1SG és 3PL alakokban *-u-* tematikus magánhangzó jelenik meg, míg a többi cellában elülső *-e-* mellett van palatalizáció. A [2] mintában azon infinitívuszban *~ći* végű igék tartoznak (pl. *seći* 'vágni', *leći* 'feküdni', *peći* 'sütni' stb.), melyekben 3PL kivételével minden cellában van palatalizáció. A [3] mintát az infinitívuszban *-a-*, de a konjugációban *-e-* tematikus magánhangzót tartalmazó igék alkotják (pl. *skak-**a**-ti* 'ugrálni', *jah-**a**-ti* 'lovagolni', *lag-**a**-ti* 'hazudni' stb.), ahol palatalizáció az egész paradigmában megjelenik, még ott is, ahol 3PL-ben nem-elülső *-u-* áll. A szubsztenderd nyelvváltozatokban elterjedt, hogy mindkét kritikus cellában megjelenhet eltérés: a [1] típusban %***mož-em*** 1SG-ben *mog-u* helyett (pl. a dráva-menti szlavóniai dialektusban (Mance 2014), a bunyevác nyelvváltozatban (Bošnjaković–Sikimić 2013)), illetve %*mož-u* 3PL-ben *mog-u* helyett (pl. bunyevác, bánáti nyelvváltozat (Ivić et al. 1997)), a [2] típusban pedig %*seč-u* 3PL-ben *sek-u* helyett (pl. a délkelet-szerbiai torlaki dialektusban (Sobolev et al. 2023), a szlavóniai dialektusban). Az alábbi paradigmatípusok sorrendjét az igék típusfrekvenciája szerint adtuk meg: [3] > [2] > [1].

	[3]	[2]	[1]	
INF	*ska**k**-ati*	*se**ć**-i*	*mo**ć**-i*	
TŐ	*skak-*	*sek-*	*mog-*	
1SG	*ska**č**-em*	*se**č**-em*	%*mož-em*	*mo**g**-u*
3PL	*ska**č**-u*	%*se**č**-u* *se**k**-u*	%*mož-u*	*mo**g**-u*
2SG	*ska**č**-eš*	*se**č**-eš*	*mož-eš*	
3SG	*ska**č**-e*	*se**č**-e*	*mož-e*	
1PL	*ska**č**-emo*	*se**č**-emo*	*mož-emo*	
2PL	*ska**č**-ete*	*se**č**-ete*	*mož-ete*	

1. táblázat *az* e*-konjugációs veláris-igék három típusa a szerbhorvátban*

Az elemzés elméleti keretét a paradigmauniformitás fogalma (Steriade 2000, Kenstowicz 2001) adja: a paradigmán belüli és paradigmák közötti analógiás vonzások képesek felülírni a fonológiai szabályosságokat, ha a paradigmatikus általánosítások dominánsnak bizonyulnak. A szubsztenderd alakok megjelenésében két egymást erősítő mechanizmus játszik szerepet. Egyrészt intraparadigmatikus nyomás érvényesül: a [2] paradigmán belüli *seč-*… alakok (*seč-em*, *seč-eš*, *seč-ete* stb.) vonzást gyakorolnak az eltérő *sek-*… 3PL alakra. Másrészt

transzparadigmatikus hatást gyakorol a magas típusfrekvenciájú [3] minta 3PL cellájában rendszeresen megjelenő …***č-u*** szekvencia és analógiás mintaként szolgál a [2] típus szubsztenderd %***seč-u*** alakjaihoz. Az [1] esetében a jelenségekre vonatkozó általánosítások bonyolultabbak: az 1SG és 3PL kanonikus alakjai közötti homonimia nem egyetlen, egységes javítóstratégiát hív elő, hanem több, egymással versengő megoldást, amelyek a dialektális kontinuum különböző pontjain eltérő mértékben realizálódnak (pl. *mog-u* '1SG' ~ *mož-edu* '3PL', *mož-em* '1SG' ~ *mog-u* '3PL', *mor-em* '1SG' ~ *mor-u*/*mor-edu* '3PL'). Az 1SG–3PL kontraszt fenntartásának igénye, az intraparadigmatikus nyomás és a [2], [3] osztályok transzparadigmatikus vonzása egyidejűleg, egymással versengő erőkként érvényesülnek, amelyek a rendszer eltérő pontjain eltérő kimeneteket produkálnak. A 2. táblázat a fenti általánosításokat összegezi: a nyilak az analógiás hatásokat jelölik, a szám- és személyjegyek sorrendje nem számít.

	[3]	[2]	[1]
3SG	*skač-e*	*seč-e*	*mož-e*
3PL	*skač-u*	→ %*seč-u*	→ %*mož-u*
1SG	*skač-em*	*seč-em*	→ %*mož-em*
2PL	*skač-ete*	*seč-ete*	*mož-ete*

2. táblázat *transz- és intrapradigmatikus hatások az* e*-osztályban*

A paradigmák közötti kontrasztok neutralizálása nem csupán ezt az egy osztályt jellemzi. Mivel a legnagyobb variabilitás a rendszerben a 3PL-ben jelentkezik – mind a tőváltakozás, mind a végződések tekintetében –, erős tendencia figyelhető meg a dialektális kontinuumon belül ezen kontrasztok kiegyenlítődésére és részleges eltűnésére. Így a 3PL a paradigmarendszer leggyengébb pontja, ahol a legtöbb közel azonos ellentétes analógiás hatás érvényesül (Rebrus 2012).

Irodalom

Mance, Nina (2014). *Suvremeni podravski govori slavonskoga dijalekta [A szlavóniai dialektus mai dráva menti nyelvjárásai]*. Doktori értekezés. Osijek: Sveučilište Josipa Jurja Strossmayera u Osijeku, Filozofski fakultet.

Bošnjaković, Žarko – Biljana Sikimić (2013). *Bunjevci. Etnodijalektološka istraživanja 2009 [Bunyevácok. Etnodialektológiai kutatás 2009]*. Novi Sad: Matica srpska – Subotica: Nacionalni savet bunjevačke nacionalne manjine.

Ivić, Pavle – Žarko Bošnjaković – Gordana Dragin (1997). Banatski govori Šumadijsko-vojvođanskog dijalekta. Druga knjiga: morfologija, sintaksa, zaključci, tekstovi. [A šumadija-vajdasági dialektus bánáti nyelvjárásai. Második kötet: morfológia, szintaxis, konklúziók, szövegek]. *Srpski dijalektološki zbornik XLIII.* Beograd: Institut za srpski jezik SANU – Srpska akademija nauka i umetnosti.

Kenstowicz, Michael (2001). Paradigmatic uniformity and contrast (manuscript, MIT).

Rebrus Péter (2012). Gyenge pontok a paradigmában: a múlt idő változatai. In: É. Kiss Katalin – Hegedűs Attila (szerk.), *Nyelvelmélet és dialektológia* 2. Piliscsaba: PPKE BTK Elmélet Nyelévszet Tanszék – Magyar Nyelvészeti Tanszék. 174-185.

Sobolev, Andrey N. – Mirjana Mirić – Daria V. Konior – Svetlana Ćirković (2023). Torlak: Areal Embedding and Linguistic Characteristics. *Zeitschrift für Slavische Philologie* 79/1. 23-46.

Steriade, Donca (2000). Paradigm uniformity and the phonetics-phonology boundary. In: Michael Broe – Janet Pierrehumbert (eds.). *Papers in Laboratory Phonology* 5. Cambridge: Cambridge University Press. 313-334.

INTRA-SPEAKER VARIATION ACROSS THE LIFESPAN: GRAMMATICAL CHANGE IN A HIGH-DENSITY CORPUS OF AN AMERICAN TV PERSONALITY

Lena Senger[1] – Peter Uhrig[2]

[1]Friedrich-Alexander-Universität Erlangen-Nürnberg
[2] Friedrich-Alexander-Universität Erlangen-Nürnberg

lena.ls.senger@fau.de, peter.uhrig@fau.de

Research on language change in variationist sociolinguistics has predominantly relied on apparent-time designs and real-time trend studies. While these approaches have yielded important insights into language change at the community level, they provide only indirect evidence of the variability of individual speakers' linguistic behaviour over time. Although an increasing number of researchers have recently emphasised the relevance of panel studies for addressing key questions about lifespan change and the "interaction between individual malleability and community-wide language change" (Buchstaller & Wagner 2019: 2), such longitudinal research remains rare due to practical challenges (Anthonissen & Petré 2019: 3; Bülow & Vergeiner 2021: 1). Most existing panel studies tend to be limited in scale, typically tracking only small samples of speakers and often restricting the analysis to two points in time (Beaman 2021; Bowie 2005; Buchstaller 2016; Rickford & Price 2013; Sankoff & Blondeau 2007). Harrington (2006) applies a more continuous approach by phonologically analysing the annual Christmas broadcasts of Queen Elizabeth II over a 50-year period. In their large-scale longitudinal corpus study on written Early Modern English, Anthonissen & Petré (2019) provide evidence of grammatical change across the lifespan.

In order to contribute to this line of variationist research on intra-speaker variation and change, the present study investigates spoken American English using a corpus-linguistic approach. The data used in this study consists of recorded episodes of the American daily talk show *Live!*. The recordings come with time-aligned subtitle texts, which allow the creation of individual speaker corpora. The data is analysed using the CQPweb corpus interface. This pilot study is designed as a longitudinal case study, focusing on Kelly Ripa and comparing her language use to a larger sample drawn from the community of practice of American TV broadcasters. The approach has several advantages, for example a similar communicative setting and speech style across time, provided by the show's format, while still covering a wide range of topics (and interactions with different interlocutors). In addition, the availability of dense, high-quality broadcast data covering more than fifteen years enables time-series analyses of gradual, continuous intra-speaker variation and change, similar to Harrington's (2006) investigation, but with roughly 200 recordings per year rather than one.

Not only does this study address how the linguistic behaviour of an individual speaker develops over time but also how such individual trajectories relate to community-level patterns of variation and change. The linguistic variables selected for analysis are primarily grammatical in nature and are motivated by well-documented community-level changes in American English, as for instance described by Leech et al. (2009). The investigation may be extended to include phonetic/phonological and pragmatic variables in order to capture intra-speaker variability across different linguistic levels. The findings are expected to reveal more fine-grained patterns of age-grading, lifespan change, and retrograde change, which have been identified in previous real-time research (Bülow & Vergeiner 2021: 3). This case study thus aims to contribute to variationist research by addressing the relative lack of dense longitudinal data and by showing how individual speakers' language use relates to ongoing changes in the wider speech community.

References

Anthonissen, L. – Petré, P. 2019. Grammaticalization and the linguistic individual: New avenues in lifespan research. *Linguistics Vanguard* 5/2. 1–12.

Beaman, K. V. 2021. Identity and mobility in linguistic change across the lifespan: The case of Swabian German. In: Ziegler, A. – Edler, S. – Oberdorfer, G. (eds.): *Studies in Language Variation* 27. 27–60.

Bowie, D. 2005. Language change over the lifespan: A test of the apparent time construct. *Penn Working Papers in Linguistics* 11/2. 45–58.

Bülow, L. – Vergeiner, P. C. 2021. Intra-individual variation across the lifespan: Results from an Austrian panel study. *Linguistics Vanguard* 7/2. 1–11.

Buchstaller, I. 2016. Investigating the effect of socio-cognitive salience and speaker-based factors in morpho-syntactic life-span change. *Journal of English Linguistics* 44/3. 199–229.

Buchstaller, I. – Evans Wagner, S. 2019. Introduction: Using panel data in sociolinguistic study of variation and change. In: Buchstaller, I. – Evans Wagner, S. (eds.): *Panel Studies of Variation and Change*. New York: Routledge. 1–18.

Harrington, J. 2006. An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcast. *Journal of Phonetics* 34/4. 439–457.

Leech, G. – Hundt, M. – Mair, C. – Smith, N. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.

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

Sankoff, G. – Blondeau, H. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83/3. 560–588.

TEMPORAL CHARACTERISTICS OF SPEECH BREATHING IN MULTIPLE SCLEROSIS: A STUDY OF EXPIRATORY PATTERNS AND PHONATION TIME IN HUNGARIAN SPEAKERS

Zsuzsa Szánthó[1,2,3] – Tekla Etelka Gráczi[2,3]

[1]ELTE Eötvös Loránd University
[2]ELTE Research Centre for Linguistics
[3]MTA–ELTE NYTK Lendület "Momentum" Neurophonetics Research Group

szantho.zsuzsa@btk.elte.hu, graczi.tekla.etelka@nytud.hu

1. Introduction

The present study examines the characteristics of speech breathing among speakers with multiple sclerosis (MS). Speech breathing differs from physiological breathing: the duration of exhalation is proportionally longer, and it forms a close unity with phonation, making it fundamental to speech production (Fuchs–Rochet-Capellan 2021). MS is a chronic disease of the central nervous system which leads to the demyelinisation of the nerves, and in addition to motor and cognitive symptoms, can also affect speech, weaken the respiratory muscles (Joy–Johnson 2001). Shorter exhalation and phonation times, as well as shorter speech units were found in MS in English speakers (Nordio et al. 2018). Furthermore, studies on Hungarian speakers confirm distinguishability between SM speakers and neurotypical speakers based on acoustic temporal patterns (Gosztolya et al. 2022). The aim of this study is to explore the relationship between speech and exhalation among native Hungarian speakers with MS. The research question is whether differences can be detected between MS and neurotypical speakers in speech breathing. The hypothesis of the study is that MS speakers produce shorter exhalation phases, shorter speech units, and lower maximum phonation times than neurotypical speakers.

2. Methods

31 female speakers—17 speakers with MS (MS group: 47±7 years) and 14 neurotypical speakers (CO group; 47±8 years)—took part in the present study. None of the speakers was previously at a phoniatrist or speech therapist. Respiration and speech recordings were made in a quiet room using RespTrack Model 16AD (Heldner et al. 2019). Silent breathing, reading aloud of a text of appr. 100 words (*The North wind and the Sun*), as well as the pronunciation of the speech sounds [i], [a], and [u] for their maximum duration, to measure maximum phonation time. 7 vowels had to be eliminated. The respiration was analysed in R (R Core Team 2026) adapting RespInPeace (Włodarczak 2019). Speech signal was annotated in Praat software (Boersma et al. 2026). We identified the speech units in the reading, labeled the sustained speech sounds. Finally, we determined the duration of the speech units, sustained speech sounds, and the respiration times (breath cycles, and inspiration and expiration separately). Linear Mixed Models and t-test (for number of silences in reading) were run (lme4: Bates et al. 2015, and lmerTest packages: Kuznetsova et al. 2017).

3. Results

The sustained vowels' duration was significantly shorter in the MS group (difference of group means [a]: 32, [i]: 22 and [u]: 24%), than in the CO group (Fig. 1). Neither the total duration of the breath cycles, nor the duration of the exhalation or inhalation phases showed significant difference between the two groups in silent or speech breathing. However, the within speaker variation was larger in the inhalation phases of silent breathing in the MS group, and while it was larger during reading in the CO group (Fig. 2). Neither the number of silent pauses was different in the readings of the two groups.

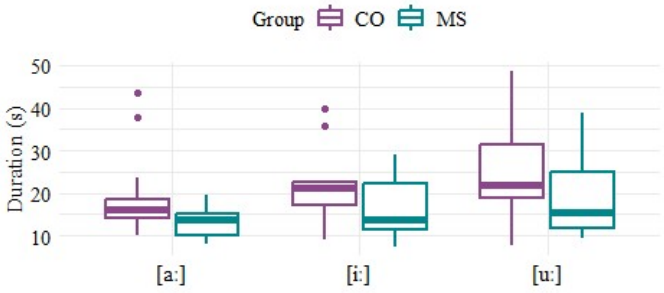


Figure 1. *Maximum vowel durations (s)*

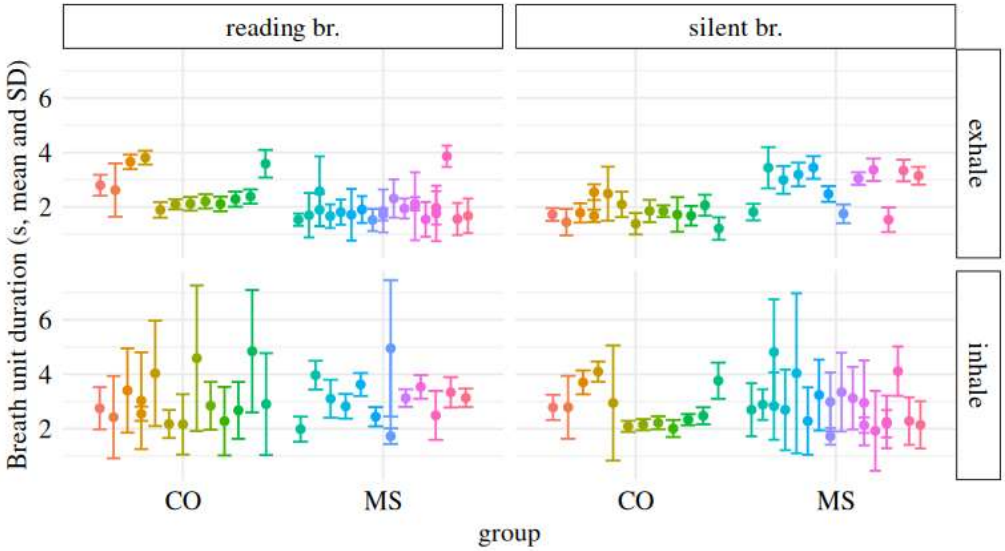


Figure 2. Breath unit duration (s)

4. Conclusion

While the breath data did not support our hypotheses, the maximal phonation time did. The results indicate that the MS speakers in our study may have dysphonia, while at a group level their respiration is intact. However, results showed high within speaker variation in duration of inhalation in both groups, although under different breathing conditions. While in neurotypical speakers the highest within speaker variation emerged during the inspiration phase of speech breathing, which can be explained by the varying demands of speech planning, in the MS speakers it emerged during silent breathing. This may suggest that variation in respiratory timing is higher in MS speakers than in neurotypical speakers during automatic, physiological breathing. Based on these present findings, we will continue the investigation by examining pause types, the temporal characteristics of pauses and speech units, and maximal inspiratory volume.

Acknowledgement

Zs. Sz. was supported by the EKÖP-26 University Research Scholarship Program of the Ministry of Science and Technology from the source of the National Research, Development and Innovation Fund.

References

Bates, D. – Maechler, M. – Bolker, B. – Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. 10.18637/jss.v067.i01

Boersma, P. – Weenink, D. – Shchupak, A. 2026. *Praat: doing phonetics by computer* [Computer program]. Version 6.4.67, retrieved 21 May 2026 from https://praat.org

Fuchs, S. – Rochet-Capellan, A. 2021. The Respiratory Foundations of Spoken Language. *Annual Review of Linguistics*, 7. 13–30.

Gosztolya G. – Tóth L. – Svindt V. – Bóna J. – Hoffmann I. 2022. Sclerosis Multiplex hangalapú felismerése akusztikai alapú beágyazások használatával. *XVIII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged, 2022. január 27–28. 151–160.

Heldner, M. – Włodarczak, M. – Branderud, P. – Stark, J. 2019. *The RespTrack System*. Phonetics Laboratory, Department of Linguistics, Stockholm University.

Joy, J. E. – Johnston Jr., R. B. 2001. *Multiple Sclerosis: Current Status and Strategies for the Future*. Institute of Medicine (US) Committee on Multiple Sclerosis: Current Status and Strategies for the Future. National Academies Press (US).

Kuznetsova, A. – Brockhoff, P.B. – Christensen, R. H. B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software,* 82(13), 1–26.

Nordio, S. – Bernitsas, E. – Meneghello, F. – Palmer, K. – Stabile, M. R. – Dipietro, L. – Di Stadio, A. 2018. Expiratory and phonation times as measures of disease severity in patients with Multiple Sclerosis. A case-control study. *Multiple Sclerosis and Related Disorders*, 23, 27–32.

R Core Team 2026. R: *A Language and Environment for Statistical Computing. R Foundation for Statistical Computing,* Vienna, Austria. https://www.R-project.org/

Włodarczak, M. 2019. RespInPeace: Toolkit for processing respiratory belt data. In Proceedings of Fonetik 2019. 115–118. 10.5281/zenodo.3246019

A BESZÉD, A BESZÉDSZERVI MOZGÁS ÉS A FELSŐ VÉGTAGI MOZGÁS VARIABILITÁSÁNAK SAJÁTOSSÁGAI PERZISZTÁLÓ BESZÉDHANGHIBÁK ESETÉN – EGY PILÓTAKUTATÁS TAPASZTALATAI

Tar Éva – Lénárt Zoltán

ELTE Bárczi Gusztáv Gyógypedagógiai Kar

tar.eva@barczi.elte.hu, lenart.zoltan@barczi.elte.hu

1. Háttér

A perzisztáló beszédhanghiba (PBHH) a fonológiai fejlődés lezárulását követően is fennálló beszédzavar, amely kedvezőtlenül befolyásolja az iskolai teljesítményt és a szociális beilleszkedést (Wren et al., 2021; 2023) A zavar hátterében beszédmotoros és tágabb mozgásszabályozási eltérések egyaránt felfedezhetők, amelyek az általános motoros teljesítményben és a nem beszédcélú motoros hálózatok működésében is megjelenhetnek (Redle et al., 2015; Carrigg et al., 2016).

A PBHH egyik meghatározó tünete a beszédprodukció tartós egyénen belüli variabilitása, amely feldolgozási vagy beszédmotoros zavarra utalhat (McNeill et al., 2022). Klinikai jelentősége ellenére a variabilitás különböző szintjei közötti kapcsolatról kevés ismeret áll rendelkezésre. Korábbi kutatások nem találtak összefüggést sem a szegmentális és szóvariabilitás, sem a percepciós és akusztikai mutatók között. Emellett az egyes akusztikai variabilitásmutatók kapcsolata sem igazolódott (Preston–Koenig, 2011; Macrae–Tyler–Lewis, 2014). Ismereteink szerint nincs olyan kutatás, amely három különböző rendszer variabilitását együttesen vizsgálta volna, illetve azt sem tudjuk, hogy célzott terápia hiányában hogyan változik a produkció variabilitása az iskoláskor első éveiben.

Pilótakutatásunk célja a 6-9 éves, célzott beszédterápiában nem részesülő, beszédhanghibás gyermekek beszédprodukciós, beszéd- és végtagmozgás-variabilitásának feltárása volt. Vizsgálatunk három kérdésre irányult: (1) együtt jár-e a beszédprodukció, a beszédszervi mozgások és a végtagmozgások variabilitása; (2) milyen mintázatok jellemzik az egyéni variabilitásprofilokat; valamint (3) kimutatható-e fél év alatt változás a variabilitás mértékében.

2. Módszer

A vizsgálatban 10 beszédhanghibával diagnosztizált, 8;0 éves átlagéletkorú kisiskolás vett részt, akiktől két időpontban, fél éves különbséggel gyűjtöttünk adatokat. A beszédprodukció variabilitását képmegnevezési és szóismétlési beszédminták percepciós elemzésével, a beszédmotoros időzítés variabilitását a /p/ VOT-értékeiből számított variációs együtthatóval (CV) jellemeztük DDK- és képmegnevezési feladatok során. Az általános motoros működés indikátoraként a felső végtag mozgását körrajzolási feladat során, a berajzolt terület mutatói és azok CV-i alapján elemeztük. Az egyéni profilok összehasonlítására robosztus z-értékeket képeztünk; az összefüggések feltárására Spearman-korrelációt és az eloszlásdiagramok elemzését, az idői változások vizsgálatára Mann-Whitney U-próbát alkalmaztunk.

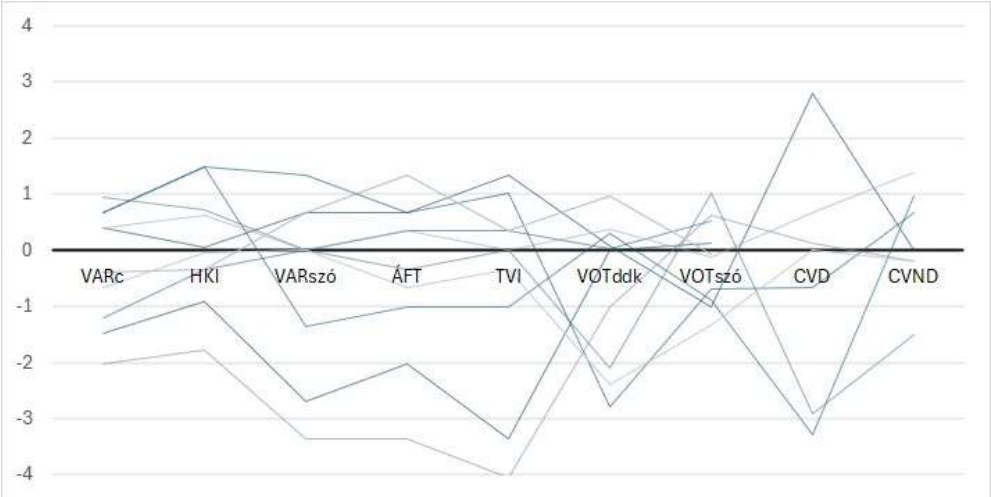
3. Eredmények

Eredményeink arra utalnak, hogy a szegmentumvariabilitás mutatói egymással és a szóvariabilitás mutatók egymással erős pozitív együtt járást mutatnak, míg a beszéd és a beszédmozgás mutatók között nem találtunk kapcsolatot. A domináns végtagmozgás eltérése erős negatív összefüggést mutatott a szóprodukció variabilitásával, míg a nem-domináns végtag variabilitásmutatói egymással pozitívan korreláltak (lásd 1. táblázat).

Az egyéni variabilitásprofilok jól elkülönültek: a gyermekek eltérő területeken mutattak a csoportátlagtól való eltéréseket. A legnagyobb eltérések a szóprodukció és a domináns végtag mutatóiban láthatók (lásd 1. ábra).
A féléves időtávon nem találtunk statisztikailag szignifikáns változást a vizsgált mutatók értékében (minden $p > 0,05$) és a hatásnagyságok is kicsik voltak.

	Beszédprod. Szegmentum		Beszédprodukció Szó			Beszédmozgás (CV)		Végtagmozgás					
	VAR_C	HKI_C	$VAR_{SZÓ}$	ÁFT	TVI	VOT_p DDK	VOT_p SZÓ	CV D	CV ND	TER D	ELT D	TER ND	ELT ND
VAR_C	1,000												
HKI_C	**,881**	1,000											
$VAR_{SZÓ}$	0,359	0,355	1,000										
ÁFT	0,367	0,355	**,954**	1,000									
TVI	0,395	0,309	**,935**	**,982**	1,000								
VOT_p DDK	-0,091	-0,018	-0,408	-0,144	-0,144	1,000							
VOT_p SZÓ	0,218	-0,036	-0,148	-0,018	0,000	0,036	1,000						
CV D	-0,636	-0,649	0,222	0,360	0,360	0,286	0,036	1,000					
CV ND	0,036	0,414	0,074	0,234	0,108	0,607	-0,214	0,143	1,000				
TER D	-0,618	-0,577	-0,593	-0,450	-0,396	0,607	-0,036	0,464	0,143	1,000			
ELT D	-0,385	-0,327	**-,972**	**-,873**	**-,882**	0,559	0,180	-0,108	0,126	0,631	1,000		
TER ND	0,036	0,414	0,259	0,396	0,252	0,429	-0,143	0,214	**,964**	0,000	-0,054	1,000	
ELT ND	0,036	0,414	0,259	0,396	0,252	0,429	-0,143	0,214	**,964**	0,000	-0,054	**1,000**	1,000

1. táblázat. *Variabilitásmutatók közti korreláció.* VAR = Variábilis item száma, HKI = Hiba Konzisztencia Index, ÁFT = Átlagos Fonémikus Távolság, TVI = Token Variabilitás Index, D = Domináns, ND = Nem-domináns. TER = Terület, ELT = Eltérés. Sötét cella = Bonferroni-korrekció után ($p = 0,003$) szignikáns kapcsolat



1. ábra. *A gyermekek sztenderdizált variabilitás értékei a vizsgált mutatók szerint*

4. Következtetés

Az eredmények a vizsgált mintán arra utalnak, hogy a PBBH-ban megfigyelhető variabilitás nem egységes jelenség, hanem a beszédprodukció, a beszédmotoros működés és az általános motoros működés részben független komponenseiből épül fel. A gyermekek egyéni variabilitásprofiljai jelentős heterogenitást mutattak, míg csoportszinten fél év alatt, célzott beszédterápia hiányában nem volt kimutatható változás a vizsgált mutatókban. Ugyanakkor a kis elemszám miatt az eredmények óvatos értelmezést igényelnek. A pilot vizsgálat tapasztalatai hozzájárulnak a mérőeszközök finomításához és további, nagyobb mintán végzett kutatások megalapozásához.

Irodalom

Carrigg, B. – Parry, L. – Baker, E. – Shriberg, L. D. – Ballard, K. J. 2016. Cognitive, linguistic, and motor abilities in a multigenerational family with childhood apraxia of speech. *Archives of Clinical Neuropsychology* 1;31/8. 1006–1025.

Macrae, T. – Tyler, A. A. – Lewis, K. E. 2014. Lexical and phonological variability in preschool children with speech sound disorder. *American Jornal of Speech-Language Pathology* 23/1. 27–35.

McNeill, B. – McIlraith, A. L. – Macrae, T. – Gath, M. – Gillon, G. 2022. Predictors of speech severity and inconsistency over time in children with token-to-token inconsistency. *Journal of Speech, Language, and Hearing Research 65*/7. 2459–2473.

Preston, J. L. – Koenig, L. L. 2011. Phonetic variability in residual speech sound disorders: Exploration of subtypes. *Topics in Language Disorders* 31/2. 168–184.

Redle, E. – Vannest, J. – Maloney, T. – Tsevat, R. K. – Eikenberry, S. – Lewis, B. – Shriberg, L. D. – Tkach, J. – Holland, S. K. 2015. Functional MRI evidence for fine motor praxis dysfunction in children with persistent speech disorders. *Brain Research* 1597. 47–56.

Wren, Y. – Pagnamenta, E. – Peters, T. J. – Emond, A. – Northstone, K. – Miller, L. L. – Roulstone, S. 2021. Educational outcomes associated with persistent speech disorder. *International Journal of Languge Communication Disorders* 56/2. 299–312.

Wren, Y. – Pagnamenta, E. – Orchard, F. – Peters, T. J – Emond, A. – Northstone, K. – Miller, L. L. – Roulstone, S. 2023. Social, emotional and behavioural difficulties associated with persistent speech disorder in children: A prospective population study. *JCPP Advances* 17;3/1. e12126

AUTOMATIKUS ÉLETKORBECSLÉS A PHONEXIA SPEECH PLATFORM ALKALMAZÁSÁVAL

Tatár Zoltán

NBSZ Szakértői Intézet

tatar.zoltan@nbsz.gov.hu

Egy ismeretlen beszélő életkorának meghatározása egy nyomozás során az elkövetői profil készítése keretében egy szűrő funkciót tölt be, amellyel szűkíteni lehet a nyomozás irányát. Ezt a vizsgálatot elsősorban a nyomozó végzi, ritkábban bevonnak hangtechnikai szakértőt is, minden esetben percepciós elemzést végeznek, és saját szakmai tapasztalatuk alapján vonják le a következtetéseiket. Kutatásaim célja ennek a folyamatnak az elősegítése a legújabb technológiák felhasználásával.

Az életkor behatárolására a kutatásom első szakaszában a Vocalise (Vocalise 2021) rendszer képességeit vizsgáltam, amely egy automatikus beszélőazonosításra kifejlesztett szoftver, és arra voltam kíváncsi, hogy milyen mértékben alkalmazható egy beszélő életkorának megállapítására. A Vocalise x-vektor alapú beszélő-összehasonlítást alkalmaz, amelyből egy *score* értéket számol, amely a két beszélő között fennálló hasonlóságra, illetve a beszélők azonosságára vagy különbözőségére utal. A Vocalise hatékonyságát 530 beszélőnek összesen 1102 hangmintájának felhasználásával teszteltem (BEA (Gósy et al. 2012), Longitudinális korpusz (Gráczy et al. 2020) és NBSZ adatbázis). Kétféle osztályozási megközelítést alkalmaztam, egyrészt korcsoportokba sorolást, másrészt a pontos életkor meghatározását.

A gyakorlati beszélőprofil-készítés jelenleg humán jellegű, az első módszert alkalmazzuk, és a cél az, hogy minél szűkebb intervallumba soroljuk az ismeretlen beszélő életkorát. A végeredmény változó: a legrosszabb esetben megállapítható, hogy bizonyos életkor alatti vagy feletti a beszélő, a legjobb esetben pedig egy 10 éves idősávot tudunk meghatározni, például „a kérdéses beszélő 30–40 év között személy lehet”. A Vocalise eredményei elmaradtak a várakozásaimtól, és a szakirodalomban már nagy múltra visszatekintő humán életkorbecslési kutatási eredményektől is: például Gósy Mária (2001) kutatásában az adatközlők 60% feletti helyes életkorbecslési eredményt értek el, ezzel szemben a Vocalise 10 éves korcsoportba sorolásának hatékonysága mindössze 34% volt, a pontos életkormeghatározása 10 év tévedéssel 44% (Tatár 2026). Mindezek miatt ez a módszer a gyakorlatban sem alkalmazható.

Ezeknek az eredményeknek a javítása reményében tesztelem jelen kutatásomban a Phonexia Speech Platform (Phonexia 2026) rendszerébe tartozó Age Estimation modult. Ezt az alkalmazást gépi tanulással támogatott demográfiai elemzésre fejlesztették ki, és a készítők állítása szerint képes arra, hogy meghatározza a hangfelvételen beszélő személy életkorát. A vizsgálatot ugyanazoknak a felvételeknek a felhasználásával végzem, mint előzőleg (lásd fent), azonban a Phonexia esetében nincs szükség összehasonlító adatbázisra, mint a Vocalise-nál volt, ezért csak a Teszt hanganyagokat használom fel. Ez 304 beszélőtől származó összesen 851 beszédrészletet jelent. Újra megvizsgálom a 10 év elteltével újrarögzített beszélők életkorbecslési eredményeit a Longitudinális adatbázis alapján, valamint a mikrofon-GSM hangjellemzők befolyásoló hatásait is az NBSZ saját rögzítésű spontán beszéd hanganyagainak felhasználásával. Az eljárás eredménye minden esetben a beszélő életkora egyetlen számmal kifejezve, nem egy lehetséges életkori intervallum, ezért a Phonexia hatékonyságát úgy tudom mérni, hogy megvizsgálom, hogy ebben a pontos életkorbecslésben mekkora a tévedési faktor. Ezt párhuzamba tudom állítani a korábbi pontos életkorbecslési eredményeimmel, amelynek a hatékonyságánál (44%) jobb eredményt várok ettől a célszoftvertől, amely csak így adhat alapot egy gyakorlati beszélőprofil-készítési módszertan kidolgozásához.

Irodalom

Gósy, M. 2001. A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr*. 125. 478–488.

Gósy, M. – Gyarmathy, D. – Horváth, V. – Gráczy, T. E. – Beke, A. – Neuberger, T. – Nikléczy, P. 2012. BEA: Beszélt nyelvi adatbázis. In Gósy, M (szerk.) *Beszéd, adatbázis, kutatások.* Akadémiai kiadó, Budapest. 9–25.

Gráczy, T. E. – Huszár, A. – Krepsz, V. – Száraz, B. – Damásdi, N. – Markó, A. 2020. Longitudinális korpusz magyar felnőtt adatközlőkről. In Berend, G. – Gosztolya, G. – Vincze, V. (szerk.) *XVI. Magyar Számítógépes Nyelvészeti Konferencia kötete.*

Phonexia 2026. Phonexia Speech Platform. https://phonexia.com/product/phonexia-speech-platform (Utolsó megtekintés: 2026.05.24.)

Tatár, Z. 2026. Automatikus beszélőprofil-készítés – életkorbecslés. In Főző, E. – Kóré, V. – Markó, A. – Mell, P. – Nyíri, M. (szerk.): *Szakértői kutatások a MILAB égisze alatt.* SZIMfonia IV. Nemzetbiztonsági Szakszolgálat, Budapest. 81–106.

Vocalise 2021. Vocalise | Oxford Wave Research. http://oxfordwaveresearch.com/products/vocalise/ (Utolsó megtekintés: 2025.11.04.)

/s/ and /ʃ/ Variation Across Gender-Diverse Speakers: Implications for Forensic Casework Involving Minority Speaker Populations

Katelyn Taylor – Kirsty McDougall
Department of Theoretical and Applied Linguistics, University of Cambridge
klt56@cam.ac.uk; kem37@cam.ac.uk

Phonetic and sociolinguistic research has extensively explored how sex-based factors shape speech production, yet most existing studies and corpora remain centred on cisgender speakers, leaving gender-diverse voices underrepresented (Zimman, 2018, 2021; Merritt, 2023). This gap poses challenges for forensic speaker comparison where gender-diverse individuals are involved, and it affects broader work on voice perception where assumptions about binary gender categories are often taken for granted. Sibilant consonants such as /s/ and /ʃ/ have long been recognised as salient markers in the negotiation of gender and sexuality. Phonetically oriented studies often attribute gendered differences in /s/ production to sex-based anatomical variation (Fuchs & Toda, 2010), while sociolinguistic research emphasises learned articulatory habits and the role of social practice (Avery & Liss, 1996; Zimman, 2017). These competing perspectives highlight the complex interplay between biological and social factors in speech production.

Building on prior perceptual findings that listeners can somewhat successfully attribute speakers' gender identities from speech cues, the present study investigates whether acoustic variation in /s/ and /ʃ/ aligns more closely with gender identity or biological sex, contributing to broader discussions of how phonetic detail encodes social meaning. Twenty-four speakers participated in this study, representing six gender identity groups, outlined in Table 1. Each speaker was recorded reading the Rainbow Passage twice (Fairbanks, 1960). The recordings were processed using Montreal Forced Alignment and manually checked to ensure accurate boundary alignment and reliable IPA transcription. Acoustic measurements of Centre of Gravity (CoG), intensity, skewness, and kurtosis will be extracted using the Vargo (2025) Praat Mass Analyzer script. These variables will be analysed to determine whether variation in /s/ and /ʃ/ productions aligns more closely with previous phonetic or sociolinguistic findings.

Gender Identity	*Number of speakers*	*Sex assigned at birth*
Cisgender Male	4	Male
Cisgender Female	4	Female
Non-binary	8 (4+4)	Male and Female
Transgender Male	4	Female
Transgender Female	4	Male
Sum:	24	

Table 1. *An outline of the sex assigned at birth and gender identity of the 24 speakers recorded reading the Rainbow Passage for the present study.*

As this experiment is ongoing, the acoustic analysis and statistical modelling are still being completed. Findings and emerging patterns and their implications for forensic casework involving speakers from minority populations will be presented at the conference.

References

Avery, J. D., & Liss, J. M. (1996). Acoustic characteristics of less-masculine-sounding male speech. *The Journal of the Acoustical Society of America*, *99*(6), 3738-3748.

Fairbanks, Grant (1960). Voice and articulation drillbook. New York: Harper & Row.

Fuchs, Susanne & Martine Toda. (2010). Do differences in male versus female /s/ reflect biological or sociophonetic factors? In Susanne Fuchs, Martine Toda & Marzena Zygis (eds.), An interdisciplinary guide to turbulent sounds,281–302. Berlin & New York: Mouton de Gruyter.

Merritt, B. (2023). Speech beyond the binary: Some acoustic-phonetic and auditory-perceptual characteristics of non-binary speakers. *JASA Express Letters*, *3*(3).

Vargo, Julian (2025). Praat Mass Analyzer [Praat Script]

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

Zimman, L. (2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass*, *12*(8), e12284.

Zimman, L. (2021). Gender diversity and the voice. In *The Routledge handbook of language, gender, and sexuality* (pp. 69-90). Routledge.

PROSODIC VARIABILITY AT THE MORPHOLOGICAL BOUNDARY: STRESS PATTERNS IN TATAR-RUSSIAN INTRA-WORD CODE-SWITCHING

Alina Timashova[1]

[1]Department of Language Science and Technology, Saarland University

alti00002@stud.uni-saarland.de

Introduction

Tatar, a Kipchak-branch Turkic language spoken primarily in the Republic of Tatarstan, Russia, exhibits frequent code-switching (CS) with Russian in everyday speech. A notable feature of this contact situation is intra-word CS, in which a Russian stem takes a Tatar inflection within a single word. For example, *karana'virustan* ('from coronavirus') combines the Russian stem *karana'virus-* with the Tatar ablative suffix *-tan*. CS has been studied extensively from a morphosyntactic perspective, but its prosodic characteristics remain under-researched (Stefanich et al., 2019).

The Tatar-Russian pair is a particularly interesting case since the two languages have conflicting stress systems. Russian has flexible lexical stress that can fall on any syllable (Shvedova, 1980), whereas Tatar has regular word-final lexical stress (Comrie, 1997). When a Russian stem is integrated into Tatar morphology, the speaker faces a choice: preserve the original Russian stress placement, shift to Tatar final-syllable stress, or produce an intermediate pattern. Recent work on the comparable Kazakh-Russian pair reports that a hybrid stress pattern emerges in experimental settings (Baidildilova et al., 2026).

Our research questions are:

1. How Tatar-Russian bilinguals resolve the stress placement conflict in ICS words?
2. Is the resolution strategy consistent across speakers, or does each speaker adopt an individual pattern?

Method

We are building a prosodically annotated spoken corpus of Tatar-Russian CS, covering both intra- and inter-word switches. The data come from natural conversational speech in Tatar podcasts on the @azatliq YouTube channel. 25 minutes per speaker are analysed, which keeps annotation tractable while yielding a sufficient number of CS instances. 3 speakers have been annotated so far, with the target of expanding to 10. The speakers are Tatar-Russian bilinguals and can be assumed Russian-dominant, however their exact L1/L2 status is not recoverable from the recordings.

Each token is annotated in Praat across multiple TextGrid tiers: phrase boundaries, inter- and intra-word CS labels, syllable boundaries, primary and secondary stress, the morphological CS boundary (the Russian-to-Tatar transition point), and English glosses (Fig. 1).

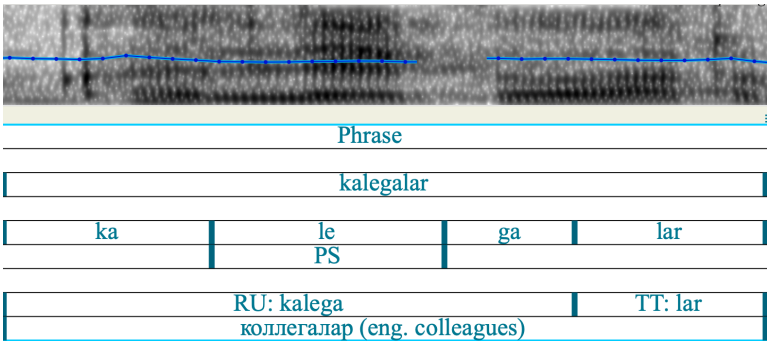


Figure 1. *Annotation example.*

CS tokens are distinguished from borrowings by degree of morphological and phonological integration into Tatar, following Myers-Scotton (2006) and Muysken (2011): established, lexicalised loans are excluded, while Russian stems that retain non-native phonological properties (e.g. not undergo Tatar vowel harmony) are treated as code-switches. As the boundary between CS and borrowings is theoretically contested, ambiguous cases are assigned to a separate Transition tier.

Stress was marked perceptually by a single annotator. To provide an objective check on these categories, syllable-level F0 and duration were extracted with Parselmouth. F0 was converted to normalised semitones (reference 100 Hz for male speakers, 200 Hz for female speakers). Duration was expressed as a ratio of each syllable's duration to the word's mean syllable duration. Intensity was not analysed since it is affected by varying recording conditions. A formal inter-annotator reliability check is planned for the ten-speaker set.

To our knowledge, this is the first spoken-language resource to combine prosodic annotation with morpheme-level language identification for a Cyrillic Turkic-Slavic pair. The existing NMCTT treebank (Taguchi et al., 2022) contains written data only.

Results

Across 3 speakers, 88 intra-word CS tokens were annotated (30, 31, and 27 per speaker). Repetitions of the same word were stress-consistent, so these tokens reduced to 78 unique word types. Each type was categorised by stress placement as Tatar (final) or Russian (non-final): 34 (44%) show the Tatar pattern and 44 (56%) show the Russian pattern. Per speaker, Tatar-pattern types account for 14 of 28, 9 of 26, and 12 of 25, or about 35 to 50% of word types each (Fig. 2).

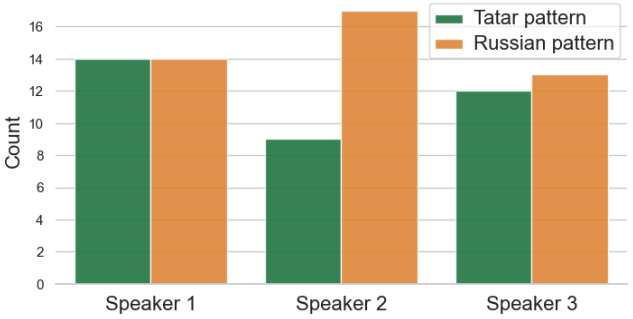


Figure 2. *Stress pattern counts by speaker, over unique word types.*

The two stress patterns differ in how the stressed syllable is marked acoustically. In Russian-pattern words, the stressed syllable shows both a pitch rise and lengthening (F0 -0.06 to +0.09 st; duration ratio 0.94 to 1.16). In Tatar-pattern words, it shows lengthening without a pitch rise (F0 +0.01 to -0.01 st; duration ratio 0.95 to 1.14). Measurements were done over all syllable tokens (Fig. 3).

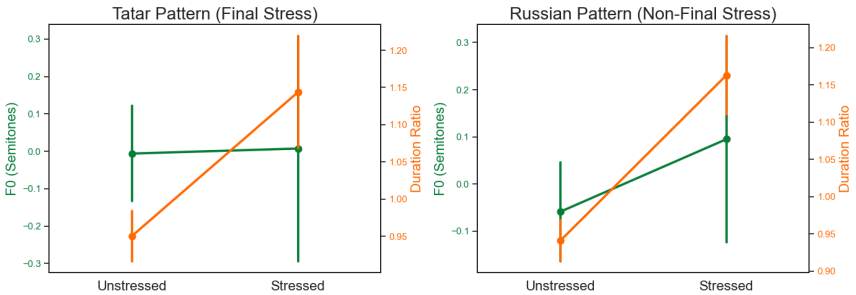


Figure 3. *Normalised semitones (green) and duration ratios (orange) for unstressed vs. stressed syllables, split by stress pattern.*

Conclusion

The roughly balanced split between the two patterns indicates no strong overall preference for either language's stress system at the group level. However, this balance is not identical across all speakers: while all three produce both patterns, Speaker 2 leans more toward the Russian pattern (65%). These between-speaker differences need further analysis on a larger sample.

Resolution varies across the lexicon but is consistent within each word in the repeated examples we observed: a word takes one system's stress placement together with that system's acoustic cues. These preliminary results suggest that speakers keep each morpheme's source-language stress marking intact rather than blending the two systems.

Beyond these findings, the dataset is itself a contribution. The annotation scheme aligns prosodic and morphological information within a single TextGrid structure, and it can serve as a template for other contact situations involving agglutinative languages.

Future work will expand the corpus to ten speakers, which will enable systematic between-speaker comparison. It will also move the acoustic analysis from syllable-level to vowel-level segmentation for greater precision.

References

Baidildinova, M. – Mayer, C. – Scontras, G. 2026. Stress Patterns in Intra-word Code-switching. *Proceedings of the Annual Meetings on Phonology* 2(1).

Comrie, B. 1997. Tatar (Volga Tatar, Kazan Tatar) phonology. In: Kaye, A.S. (ed.): *Phonologies of Asia and Africa* 2. Winona Lake: Eisenbrauns. 899–911.

Muysken, P. 2011. Code-switching. In: Mesthrie, R. (ed.) *The Cambridge Handbook of Sociolinguistics.* Cambridge: Cambridge University Press. 301–314.

Myers-Scotton, C. 2006. *Multiple Voices: An Introduction to Bilingualism.* Malden: Blackwell.

Shvedova, N. Yu. (ed.) 1980. *Russkaja grammatika* [Russian grammar]. Moscow: Nauka.

Stefanich, S. – Cabrelli, J. – Hilderman, D. – Archibald, J. 2019. The morphophonology of intraword codeswitching: representation and processing. *Frontiers in Communication* 4:54.

Taguchi, C. – Iwata, S. – Watanabe, T. 2022. Universal Dependencies Treebank for Tatar: Incorporating Intra-Word Code-Switching Information. *Proceedings of the Workshop on Resources and Technologies for Indigenous, Endangered and Lesser-resourced Languages in Eurasia within the 13th Language Resources and Evaluation Conference*. Marseille, France. 95–104.

A PHONEXIA BESZÉDTECHNOLÓGIAI PLATFORM FEJLŐDÉSÉNEK ÖSSZEHASONLÍTÓ ELEMZÉSE (2018–2026)

Varga Zoltán[1] – Tatár Zoltán[1]

[1]NBSZ Szakértői Intézet

varga.zoltan@nbsz.gov.hu, tatar.zoltan@nbsz.gov.hu

A beszédtechnológia fejlődésével újabb és újabb vizsgálati metódusok kerülnek be az igazságügyi szakértői elemzési módszerek sorába. Egy 2018-as tesztsorozat alkalmával feltérképeztük a Phonexia alkalmazás néhány képességét:

Nemmeghatározás (Gender identification–GID)
Életkorbecslés (Age Group Estimation–AGE)
Beszélőszétválasztás (Diarization–DIAR)
Kulcsszókeresés (Keyword spotting–KWS)
Beszélőazonosítás (Speaker identification–SID)

Ezen képességeket a telefonközpontok tudják leginkább kihasználni, mi viszont azt vizsgáltuk meg, hogy milyen segítséget nyújthatnak az NBSZ Szakértői Intézet munkájában. A beszélő nemének és életkorának meghatározása a beszélőprofil készítésében lenne hasznos, a beszélőszétválasztás alapfeltétele a kulcsszókeresésnek és a beszélőazonosításnak, a kulcsszavak kigyűjtése a nyelvész szakértői munkát támogathatja, a beszélőazonosítás modul pedig a hangtechnikai szakértői beszélőazonosításban nyújthat segítséget. Ezeknek a technológiáknak alapvető sajátosságuk, hogy nagy mennyiségű hanganyagokon képesek elvégezni a feladatot, ezzel egy előszűrést is végezhetnek, az utána következő humán feldolgozás számára.
A 2018-as Phonexia képességeit még nagyfokú bizonytalansági faktor jellemezte, amely az alábbi eredményekből is látszik:

Nemmeghatározás: 7,47% hibaarány.
Életkorbecslés: 7 év tévedéssel 53,67% találati arány.
Beszélőszétválasztás: az értékelés humán minősítésen alapult, és nem volt számszerűsíthető, más alkalmazásokkal összevetve (Batvox4.1, Montana, Belux) a második legjobb eredményt érte el.
Kulcsszókeresés: 72% helyes találat.
Beszélőazonosítás: a Batvox-szal összehasonlítva kevésbé pontos, mert a Batvox likelihood ratio-t számol, amely alapján megállapítja, hogy hányszor nagyobb a valószínűsége annak, hogy a két beszélő azonos vagy különböző; a Phonexia SID modulja pedig csak score értéket számol, amely alapján megállapítja, hogy melyik összehasonlító minta hasonlít a legjobban a kérdéses beszélőhöz.

A 2018-as teszt célja a gyakorlati alkalmazhatóság vizsgálata volt, nem az eredmények tudományos pontosságú meghatározása, de ezek az értékek a Phonexia legújabb Speech Platform (Phonexia 2026) nevű rendszere képességeinek vizsgálatához jó összehasonlítási alapot nyújtanak. A technikai fejlődés az igazságügyi szakértői területeken is jelentős volt a világban, és az automatikus vizsgálati metódus mára több módszertan, például az ENFSI beszélő összehasonlítási módszertani ajánlásának (2022) is meghatározó pillérévé vált.
Jelen kutatás célja az, hogy a 2018-as vizsgálatokat újra elvégezzük ugyanazokkal a hanganyagokkal – 266 férfi, 252 nő, GSM és mikrofonos minták –, így meghatározhatjuk a

fejlődés fokát, illetve újra lehet gondolni azt, hogy az igazságügyi szakértői módszertan fejlesztésében milyen mértékben játszhatnak szerepet ezek a technológiák.

Irodalom

ENFSI 2022. Best Practice Manual for the Methodology of Forensic Speaker Comparison. https://enfsi.eu/about-enfsi/structure/working-groups/documents-page/documents/best-practice-manuals/ (Utolsó megtekintés: 2026. 05. 24.)

Phonexia 2026. Phonexia Speech Platform. https://phonexia.com/product/phonexia-speech-platform (Utolsó megtekintés: 2026.05.24.)

INTRASPEAKER AND INTERSPEAKER VARIATION IN SPEECH-RELATED REGISTERS: INTENSIFIERS IN MIDDLE HUNGARIAN

Varga, Mónika[1]
Hungarian Research Centre for Linguistics
varga.monika@nytud.hu

1. Introduction. Intensifiers – those elements that amplify the meaning of the items they modify, in other words, their targets – are characterized by stability, innovation, and competition across different historical periods (Claridge–Jonsson–Kytö 2024). Their use highlights emotional stance and expresses stronger conviction; thus, on the one hand, they are associated with involved, speech-related language use, while, on the other hand, they also occur in more formal registers exhibiting specific patterns in certain historical periods (Culpeper–Kytö 2010, Hiltunen 2021). The present study analyses and compares historical sources from the perspective of intensifiers, focusing on speech-related texts, as this allows for the identification of variations characteristic of individuals and smaller groups as well as for the fine tuning of previous findings. Based on Middle Hungarian historical corpus data and sources outside the corpora, a stable variation can be observed among intensifiers meaning 'very' (Varga 2024, Varga 2026). In fact, the dominant variant is *igen* which is stylistically less neutral and less informal in recent language use, while *nagyon*, which is the default variant in standard Modern Hungarian, appears to have been rare and associated with only certain authors (in Bible translations as well as in private letters and memoirs; cf. Varga 2024a, 2026). The other variants, such as *felette ~ felettébb* or *nagy* tend to appear somewhat more frequently during sub-periods or in specific registers. Furthermore, the frequency of the dominant variant is extremely varied according to authors (based on 10,000-token samples, frequencies range from as few as 2 occurrences to 100 or more), and the registers studied so far allow wide-ranging space for individual preferences.

2. Methods. In the present study, both intraspeaker and interspeaker variation are examined through three case studies. The first concerns variation within ego-documents: János Kemény's memoir (1658) is compared with his letters, which are divided into private correspondence (addressed to his wife) and more formal ones (addressed to the count or to a community). As a supplementary analysis, letters to him written by his wife, Anna Lónyai are also examined. The research question is whether János Kemény uses the same intensifiers with similar frequencies regardless of the register features – or whether the observed patterns vary depending on the situational factors at play.

The second case study allows for a generational comparison within the register of memoirs by examining the diaries of György Vass of Cege (written between 1680 and 1705) and his son, László Vass (written between 1714 and 1738). The question is whether their usage fits the trends characteristic of their respective periods (the second half of the seventeenth century and the first half of the eighteenth century), or whether a pattern characteristic of the family can be observed in their use of intensifiers. An excerpt from the diary of László Vass's wife, Kata Kemény, is also included in the analysis.

The third case study straddles the traditional boundaries of the Middle Hungarian period, and speech-related nature of the sources also requires explanation: two disciplinary works of József Fábián are compared from the late eighteenth century (1799) and the early nineteenth century (1803). The former is written for children and describes the world of nature through imitated dialogues, while the latter is intended for "the common people." The question is to what extent the audience influences the choice of intensifiers, including both their frequency and variety. Regarding the characteristics of intensifiers, the distribution of morphologically unmarked versus marked adverbial variants and the type of the targets are relevant in all three cases (cf.

Varga 2024a, 2025). For memoirs and letters, the study queries for several intensifiers using AntConc (Anthony 2023). Instances that do not serve an intensifier are manually filtered out, resulting in a substantially smaller set for direct comparison. For one of the disciplinary writings, however, the data are collected fully manually from samples of texts. The data sets are tagged according to the syntactic-semantic category of both the intensifier and its targets, as well as by sub-periods, author, and, where necessary, register. This is followed by the cross-tabulation of the individual factors.

4. Results. The preliminary findings are as follows. Despite some variation across his ego-documents, János Kemény proved to be a consistent writer. However, in his memoirs, on a per-10,000-word basis, he uses the default intensifier *igen* considerably more frequently than in his official letters or in those addressed to his wife. The other intensifiers occur no more than five times per 10,000 words (see Figure 1).

The intergenerational analysis revealed similar frequency patterns only in some cases: in terms of overall occurrence, the default *igen* and *annyira* (the latter being a much rarer variant) appear in nearly similar normalized frequencies in the writings of György Vass and his son, László Vass. Certain intensifiers, however, appear only in György Vass's diary, although they are rare, that is 1 occurrence per 10,000 words: the combination of *igen* with *felette*; the intensifiers *ilyen* and *olyan* meaning 'so very'; and the negative emotive *rettenetes* (an unmarked adverb meaning 'terribly'). The diary entries of László Vass's wife exhibit different patterns. The frequency of *igen* is notably higher than that used by either the father or the son, similarly to *mely*. Moreover, *nagyon* (also combined with *igen*) appears exclusively in her entries among the three writers (see Figure 2).

In the third case, *igen* occurs with similar frequency in disciplinary writings by the same author. However, in texts intended for children, other intensifiers are prominent as well, such as *oly* 'so very', *nagyon* 'very', and the morphologically unmarked *iszonyú* 'terribly'. In many instances, the polarity of *iszonyú* is shifted, which means that negative emotive connotation of the amplifier is reinterpreted to express extraordinary intensity of a non-negative feature (e.g., *iszonyú sebesen*,'terribly fast', *iszonyú apró*, 'terribly tiny', *iszonyú erős* 'terribly strong'). Similar elements also appear in the work written for the common people, though in small numbers compared to the overall volume (see Figure 3).

Morphologically unmarked adverbials, such as *nagy* 'very' and *iszonyú* 'terribly', can be observed across all three source groups. These function as modifiers of adjectival and adverbial targets while morphologically marked variants tend to modify verbal targets.

Conclusions. The three case studies on speech-related language use in Middle Hungarian revealed both intraspeaker and interspeaker characteristics regarding the variation of intensifiers. For intraspeaker variety, the author examined in the first case study exhibits a fairly consistent pattern across his ego-documents. However, in the third case study, the author adapts his use of intensifiers to different audiences, including negative emotive words as amplifiers and other usually less frequent variants. For interspeaker variety, in the second case study, the same register reveals varying patterns according to the different writers.

References

AntConc = Anthony, L. 2023. AntConc (Version 4.2.4) [Computer Software]. Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software.

Claridge, C. – Jonsson, E. – Kytö, M. 2024. Intensifiers in Late Modern English. Cambridge: Cambridge University Press.

Culpeper, J. – Kytö, M. 2010. Early Modern English Dialogues: Spoken Interaction as Writing. Cambridge: Cambridge University Press.

Hiltunen, T. 2021. Intensification in Eighteenth Century Medical Writing. *Journal of English Linguistics* 49/1. 90–113. https://doi.org/10.1177/0075424220982649

Varga, M. 2024. Felette igen nagy jó – Intenzitásjelölők variációja 16–18. századi regiszterekben. [Variation of intensifiers in 16th – 18th centuries registers.] *Argumentum* 20. 194–223.
Varga, M. 2025. Rettenetes friss, iszonyú sok, szörnyű vígság: negatív emotív intenzitásjelölők variációjáról 16–18. századi szövegtípusokban. [Variation of negative emotive intensifiers in 16th – 18th centuries text types] *Argumentum* 21. 41–66.
Varga, M. 2026. Variáció és regisztersajátosságok az intenzitásjelölők 16–18. századi használatában [Variation and register perspective in the use of intensifiers in 16th – 18th centuries.] In Gugán, K. – Varga, M. (eds.): *Regiszterfüggő nyelvi változók a középmagyarban* [Variation in Middle Hungarian: a register perspective]. *(To appear.)*

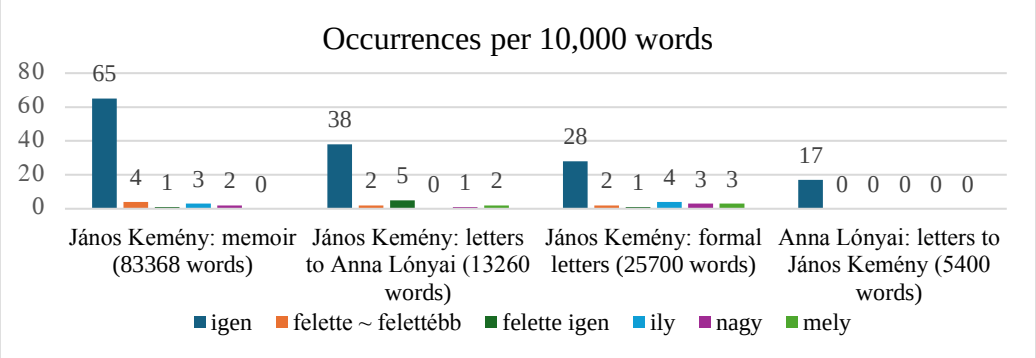


Figure 1. *Variation of intensifiers meaning '(so) very' in the ego-documents of János Kemény (and the letters of his wife to him, in total: 845 instances)*

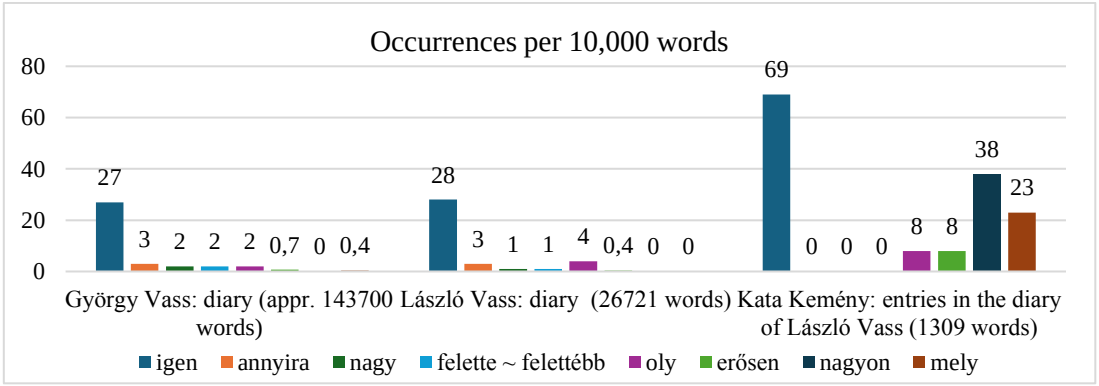


Figure 2. *Variation of intensifiers meaning '(so) very' in memoirs of family members (710 instances)*

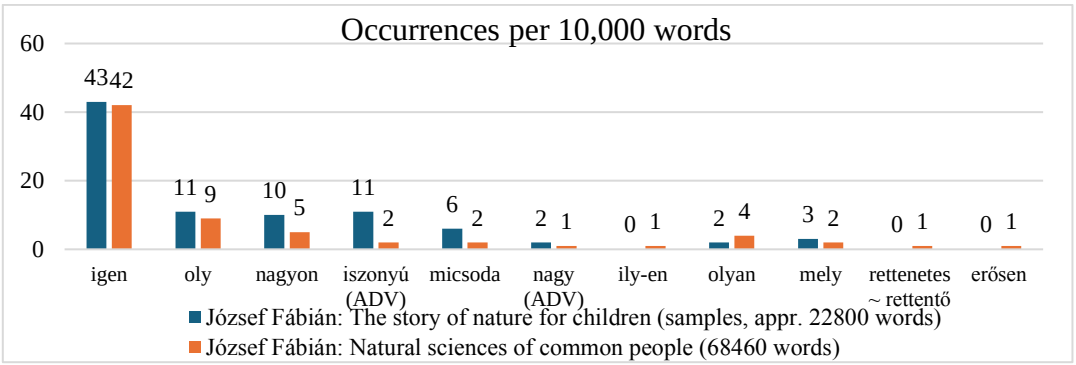


Figure 3. *Variation of intensifiers meaning '(so) very' in disciplinary texts (753 instances)*

BEYOND THE YES-BIAS? MULTILINGUAL SENSITIVITY TO SCHWA OPTIONALITY ACROSS REGISTERS: AN ACCEPTABILITY JUDGEMENT STUDY ON VARIATION AND NORM ADHERENCE

Marie-Theres Weißgerber[1] – Oliver Bunk[1] – Georg Lohfink[1]

[1]Humboldt-Universität zu Berlin

marie-theres.weissgerber@hu-berlin.de, oliver.bunk@hu-berlin.de, lohfinkg@hu-berlin.de

In German spontaneous speech, schwa in first-person singular suffixes is considered a phenomenon of (seemingly) free variation (Mooshammer et al. 2024). Whether or not schwa is realised does not change the meaning of a first-person singular verb in the present tense, e.g. [ʔıç haːbə] vs. [ʔıç haːp] ('I have'). Research has repeatedly demonstrated interactions between situational factors and phonetic parameters (see, e.g., Lancien & Côté 2021, Oschkinat et al. 2025, Ernestus et al. 2015). With respect to schwa optionality, production studies indicate that schwa is realised more often in infrequent verbs (Lange et al. 2024), and in formal compared to informal settings (Mooshammer et al. 2024). This situational pattern holds across monolingual and multilingual speaker groups (Weißgerber 2024). Yet, the perceptual side of schwa optionality remains entirely unexplored, and it is unknown whether language repertoire modulates register-sensitive acceptability judgements for this variable feature. The present study addresses this gap.

Apart from the interlocutor, the formality of a situation and the speakers' language repertoire have been identified as factors that affect the choice of optional phonetic features. In a corpus study, Bunk (2024) observed that non-morphemic [t] in words like *nich(t)* ('not') or *is(t)* ('is') was deleted more frequently in informal settings than in formal ones. Strikingly, multilingual speakers tended to delete [t] in formal settings less frequently than monolinguals, and the difference between conditions was particularly pronounced among Russian-German heritage speakers. Similar results could be identified at the lexical, discourse pragmatic (Bunk 2024) and syntactic level (Bunk et al. 2025). A widely discussed phenomenon in this respect is the concept of the *yes-bias* (see Polinsky 2018), according to which heritage speakers tend to judge ungrammatical constructions in their heritage language as correct more often than monolinguals. To our knowledge, this phenomenon has not been investigated in multilinguals' majority language. Multilingual speakers may tend to mark formal settings more explicitly, possibly due to sociolinguistic performance pressure to align with the monolingual majority society and avoid linguistic othering. For example, multilingual speakers might be stricter when it comes to evaluating deviations from the variety usually associated with formal communicative settings, i.e. a variety that is more closely aligned with the written standard.

In our study, we explore the acceptability of realisations with and without schwa in Russian-German bilingual (n=60) and monolingual German speakers (n=60). Participants are presented with audio recordings of sentences containing first-person singular verbs with and without the schwa suffix. Verbs differ in frequency and are divided into two groups (low vs. high frequency). In addition to schwa, we included filler sentences as well as pseudo-filler sentences. Pseudo-filler sentences containing *nicht* ('not') and *ist* ('is'), with and without [t], were included to enable future comparison with existing production data (Bunk 2024). Recordings were assessed for authenticity and verb intervals were cross-spliced with Praat (Boersma & Weenink 2026). All items were pseudorandomised and embedded into situational contexts (formal vs. informal). Formality was operationalized with occasion (a casual party vs.

a company party) and the interlocutors' relationship (a close friend vs. a manager at a new job). We will discuss whether multilingual speakers show stricter norm adherence than monolinguals when evaluating phonetic realisations in formal and informal contexts, and what this implies for our understanding of register sensitivity in perception.

Acknowledgements
This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1412, 416591334. The authors would like to thank the participants of the Research Colloquium *Multilingualism, Language Contact & Variation* at Humboldt-Universität zu Berlin for their valuable feedback on the study design. We also thank Christine Mooshammer, Anke Lüdeling, Malte Belz, and Aliona Petrenco for their helpful comments on the final version of the online study and Amelie Ellerich for designing the recruitment flyers.

References

Boersma, Paul and David Weenink. 2026. Praat: doing phonetics by computer [computer program]. http://www.praat.org/.

Bunk, Oliver. 2024. What does linguistic structure tell us about language ideologies? The case of majority language anxiety in Germany. *European Journal of Applied Linguistics*, 12(1): 91–116.

Bunk, Oliver, Shanley E. M. Allen, Sabine Zerbian, Tatiana Pashkova, Yulia Zuban, and Erica Conti. 2025. Information packaging and word order dynamics in language contact. *Linguistic Dynamics in Heritage Speakers: Insights from the RUEG Group*: 379–412.

Ernestus, Mirjam, Iris Hanique, and Erik Verboom. 2015. The effect of speech situation on the occurrence of reduced word pronunciation variants. *Journal of Phonetics*, 48: 60–75.

Lancien, Mélanie and Marie-Hélène Côté. 2021. Hypoarticulation as a tool for assessing social distance: An acoustic study of speech addressed to different types of interlocutors. *STUDII DE LINGVISTICĂ*, 11: 55–84.

Lange, Robert, Bianca Sell, Megumi Terada, Malte Belz, Christine Mooshammer, and Anke Lüdeling. 2024. Schwa realisation in verbal inflection in two dialogue registers of German spontaneous speech. *Zeitschrift für Sprachwissenschaft*, 43(2): 237–266.

Mooshammer, Christine, Marie-Theres Weißgerber, and Robert Lange. 2024. Schwa optionality in spontaneous speech in German: a meta-study of verbal inflection in three corpora. *Proceedings of the Nineteenth Australasian International Conference on Speech Science and Technology*: 62–66.

Oschkinat, Miriam, Melanie Weirich, Daniel Duran, and Stefanie Jannedy. 2025. Register effects of imagined addressees on f0 across generations. *JASA Express Letters*, 5(11): 115203.

Polinsky, Maria. 2018. *Heritage languages and their speakers*. Vol. 159. Cambridge University Press.

Weißgerber, Marie-Theres. 2024. Schwa optionality in verbal inflection in German: the effects of stress and phonetic context. In: *Proc. ISSP 2024*: 59–62.

MELK OR *MELEK?* A MAXIMUM ENTROPY GRAMMAR ACCOUNT FOR SCHWA EPENTHESIS IN DUTCH

Giorgia Zantei[1], Furkan Akkurt[2]

[1]University of Edinburgh, [1]Boğaziçi University

g.zantei@sms.ed.ac.uk, furkan.akkurt@bogazici.edu.tr

Introduction

When Dutch speakers say *melk* 'milk', they subconsciously face a choice: ['mɛlk] or ['mɛlək]. In Dutch, consonant clusters in coda position permit 'free' insertion of schwa. This phenomenon occurs in most varieties of Dutch, and it is considered a highly variable and optional process. It is apparently not conditioned by register or speech rate, and it is no longer considered sociolinguistically marked (Kuijpers et al. 1996; Kuijpeers & van Dondselaar 1998; Collins & Mes 1999; Warner et al. 2001). Examples of the phenomenon are displayed in (1):

(1) Schwa insertion in Dutch (Booij 1995)

a. *melk*	'milk'	['mɛlk]~['mɛlək]
b. *werk*	'work'	['wɛrk]~['wɛrək]
c. *herfst*	'autumn'	['hɛrfst]~['hɛrəfst]

In traditional rule-based models, one input corresponds to one output, and free variation, i.e., interchangeable realizations of the same item, is often overlooked. However, in order to count as free, the two (or more) competing variants should alternate at a frequency of 1:1, but this is rarely the case. This raises the research question: which factors condition the choice between the two competing forms in Dutch, and to what extent can this variation still be considered 'free'? We hypothesize that, beyond the segmental properties of C1 and C2, phrasal factors – specifically word-final voicing triggered by the following word – play a significant, previously unexamined role in conditioning schwa epenthesis.

Methodology

This study employs the case of free variation of epenthetic and non-epenthetic forms in Dutch by analyzing 16 wordforms ending by consonant cluster (formalized as C1C2), balanced per type of C1 (/r/ vs. /l/), and per type of C2 (Fricative vs. Occlusive). The probabilistic analysis was conducted using Maximum Entropy Grammar, which assigns probabilities to multiple outputs, in this case phonetic realizations, of an input. It aims to determine which output, given certain Constraints, is more likely to occur, and is a powerful tool for analyzing competing forms in phonotactics (Hayes & Wilson 2008).

The frequencies of schwa epenthesis in word-final clusters were extracted from *Corpus Gesproken Nederlands* 'Corpus of Spoken Dutch' (CGN), an open-source corpus with phonetic transcriptions. The dataset contains significantly more non-epenthetic than epenthetic forms (about 92% vs. 8%), quite far from the ideal 1:1 proportion hypothesized for genuine free variation.

Results

Schwa insertion has been traditionally seen as the result of avoidance of complex consonant clusters in coda position, without specifically considering how the type of C1 or C2 affects the probability of schwa epenthesis. The study reveals that, although the type of C2 (Fricative vs. Occlusive) in our results has a greater probability to influence the insertion of schwa than the type of C1 (/r/ vs. /l/), the factor with the greatest impact on the occurrences of schwa insertion is word-final voicing, which had never been considered in previous studies. In fact, although

word-final voiced consonants are not allowed utterance-final position in the language, in free speech the initial consonant of the following word, when voiced, triggers regressive voicing of word-final C2. One instance of this is reported in (2).

(2) /ɛn dɑn dun zə dar mɛl**əg** **b**ɛi/
...en dan doen ze daar melk bij
…and then do they there milk at/to
'and then they add milk in it'

Regressive voicing of word-final C2 shows a highly significant correlation with schwa insertion, with a p-value < 0.05, so that, if on one hand 93.67% of voiceless C2s occur as a non-epenthetic form, whereas epenthetic forms occur in 6.33% of cases, and on the other hand, word-final voiced C2s yield 68.87% non-epenthetic forms, while epenthetic forms occur in the 31.13% of the dataset. This means that word-final consonant clusters where the following word starts with a voiced consonant (and triggers regressive voicing of C2) constitutes the environment with the highest probability of schwa epenthesis.

	Voiceless C2 in word-final position	Voiced C2 in word-final position
Epenthetized outputs	68.87%	31.13%
Non-epenthized outputs	93.67%	6.33%

Table 1. *Percentage of epenthetic forms by regressive voicing*

Conclusions

These results confirm our hypothesis that phrase-level phonological context, and not only segmental factors, must be integrated into a comprehensive account of free variation. This also shows how complex free variation can be, especially when the phonological phrase, in addition to the single word, is considered as the domain of variation. This differs from the approach of rule-based models, with one input and one output correspondence. Overall, Maximum Entropy Grammar models proved capable of handling multiple and complex sets of Constraints and to successfully account for free variation in a powerful and insightful way.

References

Booij, G. (1995). *The Phonology of Dutch*. Oxford University Press Inc.

Collins, B. D., & Mees, I. (2003). *The phonetics of English and Dutch*. Brill.

Hayes, B., & Wilson, C. (2008). A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry*, *39*(3), 379–440. https://doi.org/10.1162/ling.2008.39.3.379

Kuijpers, C., & Van Donselaar, W. (1998). The Influence of Rhythmic Context on Schwa Epenthesis and Schwa Deletion in Dutch. *Language and Speech*, *41*(1), 87–108. https://doi.org/10.1177/002383099804100105

Kuijpers, C., Van Donselaar, W., & Cutler, A. (1996). Phonological variation: Epenthesis and deletion of schwa in Dutch. *Proceeding of Fourth Internation Conference on Spoken Language Processing*, *1*, 149–152.

Warner, N., Jongman, A., Cutler, A., & Mücke, D. (2001). The phonological status of Dutch epenthetic schwa. *Phonology*, *18*(3), 387–420. https://doi.org/10.1017/S0952675701004213

Table of contents

Program **5**

Keynotes **9**

Stefano Coretta • BAYESIAN MULTILEVEL MODELLING AND ESTIMATION OF UNCERTAINTY IN SPEAKER VARIATION 11

Anne Hermes • COORDINATIVE STRUCTURES IN SPEECH: LINGUISTIC, AGE- AND DISEASE RELATED FACTORS 12

Presentations **15**

Zsuzsanna Bárkányi, and Zoltán G. Kiss • MULTIPLE PERCEPTUAL PARADIGMS AND LEARNER VARIABILITY IN THE PERCEPTION OF ENGLISH ASPIRATION BY ADVANCED L2 LEARNERS 17

Ákos Blaskovics • VOWEL-TO-VOWEL ISOCHRONY ACROSS DIFFERENT SUPRASEGMENTAL AND INTERVOCALIC CONFIGURATIONS 19

Philipp Buech, Teja Rebernik, Louise Bruzzo, and Anne Hermes • EFFECTS OF AGING ON VOWEL PRODUCTION IN FRENCH: AN ACOUSTIC AND ARTICULATORY STUDY 21

Maria Mendes Cantoni, and Adelino Pinheiro Silva • GLM-BASED DIALECT DECONFOUNDING FOR SPEAKER VERIFICATION ACROSS MULTIPLE BRAZILIAN PORTUGUESE VARIETIES 24

Bernadett Dam • INTRASPEAKER VARIABILITY IN SPEECH CHARACTERISTICS BEFORE AND AFTER STROKE-INDUCED DYSARTHRIA 27

Deme Andrea, Gosztolya Gábor és Tóth László • A MAGYAR MAGÁNHANGZÓS HOSSZÚSÁGI KONTRASZT MEGVALÓSULÁSA KISAGYI STROKE UTÁN JELENTKEZŐ DIZARTRIÁBAN – ESETTANULMÁNY 30

Andrea Deme, Zsuzsa Szánthó, and Tekla Etelka Gráczi • PRODUCTION VARIABILITY IN SEGMENTAL CONTRASTS – LONG AND SHORT VOWELS IN SPEAKERS WITH DOWN SYNDROME 33

Cécile Fougeron • FROM CHAOS TO PATTERNS: UNRAVELING INTER-SPEAKER VARIABILITY IN ARTICULATION RATE ACROSS ADULTHOOD 36

Cécile Fougeron, Nicolas Audibert, and Christine Meunier • DO SPEAKERS DIFFER IN THE WAY THEY VARY? 38

Főző Eszter • A SPONTÁN BESZÉDBEN MEGJELENŐ NYELVHASZNÁLATI SAJÁTOSSÁGOK GENERÁCIÓS KÜLÖNBSÉGEI 40

Tekla Etelka Gráczi, Sarolta Murányi, Katalin Pirsel, Zsuzsa Szánthó, and Ákos Blaskovics • DIADOCHOKINETIC TASKS IN MULTIPLE SCLEROSIS: ACOUSTIC AND ARTICULATORY ASPECTS 43

Reinhold Greisbach, and Anne Hermes • TYPOLOGY OF OVERLAP IN WORD-INITIAL CONSONANT CLUSTERS IN GERMAN 46

Kornélia Juhász, and Katalin Pirsel • FINDING THEIR VOICES: IS IT MORE DIFFICULT FOR SPEAKERS DIAGNOSED WITH DOWN SYNDROME TO INITIATE VOICING? 49

Viktor Kharlamov, Daniel Brenner, and Benjamin V. Tucker • ACOUSTIC PREDICTOR IMPORTANCE AND SPEAKER-SPECIFIC VARIATION IN RANDOM FOREST MODELS OF CONVERSATIONAL ENGLISH FRICATIVES 52
Kohári Anna • JELÖLIK-E ZÖNGEMINŐSÉGGEL A PROMINENCIÁT A MAGYAR ANYANYELVŰ BESZÉLŐK? 54
Kohári Anna, and Csébfalvi Fruzsina • BESZÉLŐK KÖZÖTTI ÉS BESZÉLŐN BELÜLI VARIABILITÁS A FRÁZISVÉGI ZÖNGEMINŐSÉGBEN 57
Xuesong Li • IS NONNATIVE ENGLISH SPEECH MORE VARIABLE? EVIDENCE FROM INTRA- AND INTER-SPEAKER VOWEL VARIABILITY 60
Beier Liu, Huba Bartos, and Katalin Mády • CHOOSE YOUR ACT BY YOUR SOUND: PROSODY-SPEECH ACTS INTERACTION FOR UTTERANCE-FINAL PARTICLES IN MANDARIN CHINESE 63
Markó Alexandra • SZTRÁJK A SVÁJCI NEJLONGYÁRBAN. HAT ÉVTIZED ALAKVÁLTOZATAI A FILMHÍRADÓK KORPUSZÁBAN 66
Nhat-Nam Nguyen, Pierre-André Vuissoz, and Yves Laprie • TOWARD SPEAKER-INDEPENDENT ACOUSTIC-TO-ARTICULATORY INVERSION GUIDED BY ANATOMICAL GRID 69
Marian Oliveira, Lucrécia de Aquino Santos, and Deise Santana Maia • VARIATION IN MID VOWELS PRODUCED BY SPEAKERS WITH T21: AN EXAMINATION OF TWO BRAZILIAN DIALECTS 72
Oszkó Beatrix, Vargha Fruzsina Sára és Várnai Zsuzsa • MONDHATJUK-E, HOGY „A MAGYARBAN" MEGKÜLÖNBÖZTETÜNK HOSSZÚ ÉS RÖVID FELSŐ NYELVÁLLÁSÚ MAGÁNHANGZÓKAT? 75
Ioannis Papakyritsis, and Emilia Michou • INTRASPEAKER VARIABILITY OF NASALANCE SCORES: DATA FROM GREEK 78
Henrietta Pokk, and Péter Rácz • VARIATION AND PRIMING IN THE HUNGARIAN 2SG SUBJUNCTIVE 80
Teja Rebernik, Katharina Polsterer, Valentine Lucquiault, Valerie Querner, Martijn Wieling, and Anne Hermes • EFFECTS OF SPEECH RATE MANIPULATION IN DUTCH WORD-INITIAL CLUSTERS: AGE-RELATED AND INDIVIDUAL VARIABILITY 83
Rebrus Péter és Tóth Szilárd • VARIÁCIÓ A SZERBHORVÁT *E*-KONJUGÁCIÓS IGÉK JELEN IDEJŰ ALAKJAIBAN 86
Lena Senger, and Peter Uhrig • INTRA-SPEAKER VARIATION ACROSS THE LIFESPAN: GRAMMATICAL CHANGE IN A HIGH-DENSITY CORPUS OF AN AMERICAN TV PERSONALITY 89
Zsuzsa Szánthó, and Tekla Etelka Gráczi • TEMPORAL CHARACTERISTICS OF SPEECH BREATHING IN MULTIPLE SCLEROSIS: A STUDY OF EXPIRATORY PATTERNS AND PHONATION TIME IN HUNGARIAN SPEAKERS 91
Tar Éva és Lénárt Zoltán • A BESZÉD, A BESZÉDSZERVI MOZGÁS ÉS A FELSŐ VÉGTAGI MOZGÁS VARIABILITÁSÁNAK SAJÁTOSSÁGAI PERZISZTÁLÓ BESZÉDHANGHIBÁK ESETÉN – EGY PILÓTAKUTATÁS TAPASZTALATAI 94

Tatár Zoltán • AUTOMATIKUS ÉLETKORBECSLÉS A PHONEXIA SPEECH PLATFORM ALKALMAZÁSÁVAL 97
Katelyn Taylor, and Kirsty McDougall • /s/ AND /ʃ/ VARIATION ACROSS GENDER-DIVERSE SPEAKERS: IMPLICATIONS FOR FORENSIC CASEWORK INVOLVING MINORITY SPEAKER POPULATIONS 99
Alina Timashova • PROSODIC VARIABILITY AT THE MORPHOLOGICAL BOUNDARY: STRESS PATTERNS IN TATAR-RUSSIAN INTRA-WORD CODE-SWITCHING 101
Varga Zoltán és Tatár Zoltán • A PHONEXIA BESZÉDTECHNOLÓGIAI PLATFORM FEJLŐDÉSÉNEK ÖSSZEHASONLÍTÓ ELEMZÉSE (2018–2026) 104
Mónika Varga • INTRASPEAKER AND INTERSPEAKER VARIATION IN SPEECH-RELATED REGISTERS: INTENSIFIERS IN MIDDLE HUNGARIAN 106
Marie-Theres Weißgerber, Oliver Bunk, and Georg Lohfink • BEYOND THE YES-BIAS? MULTILINGUAL SENSITIVITY TO SCHWA OPTIONALITY ACROSS REGISTERS: AN ACCEPTABILITY JUDGEMENT STUDY ON VARIATION AND NORM ADHERENCE 109
Giorgia Zantei, and Furkan Akkurt • *MELK* OR *MELEK?* A MAXIMUM ENTROPY GRAMMAR ACCOUNT FOR SCHWA EPENTHESIS IN DUTCH 111

