NVFW
Nederlandse Vereniging voor Fonetische Wetenschappen

Dag van de Fonetiek

ABSTRACTS

31 oktober 2025

Utrecht
Janskerkhof 15A, zaal 1.01

Dag van de Fonetiek 2025

vrijdag, 31 oktober 2025

Janskerkhof 15A, Utrecht, zaal 1.01

09:30 - 10:00	**Inloop**	
10:00 – 10:45	**Keynote**	The road towards inclusive speech technology and why phonetic knowledge and sciences are essential *Odette Scharenborg, TU Delft*
10:45 - 11:45	**Postersessie 1 (met koffie en thee) – zaal 1.01 en 1.05**	
11:45 - 12:05	Breaking the ice: apparent-time change in Frisian vowel breaking *Cesko Voeten, Martijn Kingma*	
12:05 - 12:25	A computer simulation of the reducing effect *Anthe Sevenants, Dirk Speelman, Freek Van de Velde, Dirk Pijpops*	
12:25 - 12:45	The documentation and description of labial-velar and implosive consonants in the Ubangi River basin: a field report - *Lorenzo Maselli*	
12:45 - 13:00	**Algemene Ledenvergadering**	
13:00 - 14:15	**Lunchpauze**	
14:15 - 14:35	Before words: How innate knowledge shapes preverbal infants' use of prosody to express communicative functions - *Elanie van Niekerk, Caroline Junge, Iris-Corinna Schwarz, Lisa Gustavsson, Ellen Marklund, Aoju Chen*	
14:35 - 14:55	Word stress in self-supervised speech models: A cross-linguistic comparison *Martijn Bentum, Louis ten Bosch, Tom Lentz*	
14:55 - 15:15	Yes/no question prosody in Meʃegbe (Tɔŋugbe) Ewe *Man Yan Priscilla Lam, Yiya Chen*	
15:15 - 16:15	**Postersessie 2 (met koffie en thee) – zaal 1.01 en 1.05**	
16:15 - 16:35	Conservatively radical: Cross-modal language variation in entral Eastern Norwegian dialect speech - *Eva van Kampen, Remco Knooihuizen*	
16:35 - 16:55	Early prosodic boundary perception: Innate biases in preterm newborns *Jorik Geutjes, Caroline Junge, Maria-Luisa Tataranno, Manon Benders & Aoju Chen*	
16:55 - 17:15	Helpt klokje of klȯkje 't maist tegen de kol? De samenval van hok en bȯk in het Noord-Nederlands - *Maureen Prins, Remco Knooihuizen*	
17:15 - 18:00	**Borrel**	

Z.o.z. voor postertitels en auteurs

NVFW
Nederlandse Vereniging voor Fonetische Wetenschappen

Postersessie 1 (10:45 - 11:45)
Velar or Uvular? The Standard German Ach-Laut Revisited *Zoe Tholen*
Accommodation to a tongue height obstruction: Acoustic outcomes and aftereffects *Xinyu Zhang, Rob Schoonen, Esther Janse*
The evolution of uptalk in Standard Southern British *Alanna Tibbs, Jiseung Kim, Amalia Arvaniti*
The pitch patterns of trisyllabic Japanese loanwords in Standard Chinese by native speakers *Jueyu Hou, Tim Laméris*
The effect of neural immaturity after gestational diabetes on rapid auditory processing abilities in newborns *Eline de Groot, Rachida Ganga, Maria-Luisa Tataranno, Frank Wijnen, Aoju Chen*
It's not what you say, it's how you say it: Creating a Prosodic Proficiency Test Battery *Ronny Bujok, Lieke van Maastricht*
Flexibility of vowel categorisation in newly learned words *Stephanie Cooper, Brechtje Post*
Multimodal Prosodic Phrasing in Infant-Directed Speech: Testing the Cumulative-Cue Hypothesis with Gesture Restriction *Roos Ledeboer, Victoria Reshetnikova, Roy Hessels, Aoju Chen*
Turn-taking and final boundary tones in Dutch: a corpus study in replication of Caspers (2003) *Ariëlle Reitsema*
Auditory Kernels - Learning Acoustic Building Blocks from Speech *Dimme de Groot, Odette Scharenborg, Jorge Martinez*

Postersessie 2 (15:15 – 16:15)
Higher-Order Phonological Processing in Pre-Readers with and without Familial Risk of Dyslexia *Maaike Smit, Ana Carbajal Chavez, Marlies Gillis, Maaike Vandermosten, Pol Ghesquière, Jan Wouters*
Linking variability in prosody production and perception *Constantijn van der Burght, Ture Berg*
Attitudes towards accent variation in word-final nasal vowels in Polish *Weronika Polakowska*
The Role of Phonetic Entrainment in Second Language Vowel Learning *Jing Tang, Laura Smorenburg, Hugo Quené, Aoju Chen*
Development of Phonetic Distinctiveness in infants and (pre)readers at risk for dyslexia *Klara Spooren, Elise Lefèvre, Jolijn Vanderauwera, Stéphane Dupont, Jan Wouters, Pol Ghesquière, Maaike Vandermosten*
Weighing auditory, visual, and semantic cues in lexical stress perception in Spanish *Floris Cos, Lieke van Maastricht, Hans Rutger Bosker, Matteo Maran, Esther Janse*
Vowel systems of Standard and Kudar varieties of Iron Ossetic *Varvara Petrova, Timofei Mets*
Musicians vs. non-musicians: who produces sentence stress better? Analysis of fundamental frequency and syllable duration in a negative sentence among French-speaking learners of Dutch *Thomas De Wispelaere, Pauline Degrave*
Prosody of Mandarin sentence-final particles in spontaneous conversational corpus *Katrina Kechun Li, Yiya Chen*
Plosive bursts in Seoul Korean *Michaela Watkins*

KEYNOTE

The road towards inclusive speech technology and why phonetic knowledge and sciences are essential

Odette Scharenborg
TU Delft

Automatic speech recognition (ASR) is increasingly used, e.g., in emergency response centers, domestic voice assistants, and search engines. Because of the paramount relevance spoken language plays in our lives, it is critical that ASR systems are able to deal with the variability in the way people speak (e.g., due to speaker differences, demographics, different speaking styles, and differently abled users). ASR systems promise to deliver objective interpretation of human speech. Practice and recent evidence however suggest that the state-of-the-art ASRs struggle with the large variation in speech due to e.g., gender, age, speech impairment, race, and accents. The overarching goal in our research is to develop inclusive speech technology, i.e., speech technology that works for everyone irrespective of their voice, language, or the way they speak. In this talk, I will present systematic experiments aimed at quantifying, identifying the origin of, and mitigating bias in state-of-the-art ASRs on speech from different "diverse", typically low-resource, groups of speakers, and I will argue why phonetic knowledge and the phonetic sciences are essential in developing inclusive speech technology.

Breaking the ice: apparent-time change in Frisian vowel breaking

Cesko Voeten[12], Martijn Kingma[21]
[1]University of Amsterdam, Amsterdam Center for Language and Communication, [2]Fryske Akademy

Frisian vowel breaking is an opaque synchronic process whereby the ingliding diphthongs [iə,yə,uə,eə,oə] alternate with the 'broken' vowels [jɪ,jø,wo,jɛ,wa] (Tiersma 1978ff). Cross-linguistically, synchronically opaque alternations tend to regularize diachronically (e.g. Sneller 2018), as has been explicitly predicted for Frisian vowel breaking (Arndt-Lappe & Ernestus 2020). Synchronically, such ongoing sound change presents as variation, and indeed, variation in Frisian vowel breaking has been observed anecdotally (e.g. Stefan 2022). However, phonetic measurements, including of change over time, are lacking; the most recent ones still date from 1985 (de Graaf).

We, hence, use the recently-completed Boarnsterhim corpus (Kingma et al submitted) to conduct a novel phonetic investigation of Frisian vowel breaking. The corpus contains 112 speakers from the Boarnsterhim municipality born between 1897 and 2001 recorded in 1980 and/or 2010. Both spontaneous and read speech are available. We present two analyses of this rich phonetic dataset. The first analysis used GAMs (Wood 2017) to model apparent-time changes in ingliding and breaking vowels' F1–F2 trajectories, taking account of potentially-nonlinear formant dynamics (Voeten, Heeringa, & Van de Velde 2022). We discuss evidence of [jɪ] and [wa] merging with [iː,uː], and evidence suggesting a reconfiguration of breaking along the front–back dimension.

The second analysis, using read speech only, used mixed-effects logistic regression to determine if words have changed their participation in the breaking allomorphy. We discuss both general factors (specifically: vowel class, part of speech, morphological derivation, ambisyllabicity of the following consonant; these three predictors show significant effects, in addition to other predictors that were not significant) as well as individual differences among words (based on the by-word random effects; cf. Voeten 2021). We discuss which words appear to be changing particularly strongly, with particular attention to morphophonological predictors of that change (cf. Bergsma et al in progress).

References

Arndt-Lappe, S., & Ernestus, M. (2020). Morpho-phonological alternations: The role of lexical storage. In V. Pirrelli, I. Plag, & W. U. Dressler (Eds.), *Trends in Linguistics: Studies and Monographs: Vol. 337. Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon* (pp. 191-227). Mouton de Gruyter.

Bergsma, F., Fingerhut, K., Kingma, M. & Van 't Veer, M. (in progress). Breaking ground: Frisian breaking in adjectives reanalysed.

De Graaf, T. (1985). Phonetic aspects of the Frisian vowel system. *North-Western European language evolution, 5*, 23-40.

Kingma, M., Pinget, A-F., Heeringa, W. & Van de Velde, H. (submitted). The Boarnsterhim Corpus: A Frisian-Dutch bilingual speech corpus in apparent- and real-time. *Language Resources and Evaluation*.

Sneller, B. (2018). *Mechanisms of phonological change*. PhD dissertation, University of Pennsylvania.

Stefan, M. H. (2022). *Spoken Frisian: Language contact, variation and change*. PhD dissertation, Rijksuniversiteit Groningen.

Tiersma, P. M. (1978). Bidirectional leveling as evidence for relational rules. *Lingua, 45*, 65-77.

Tiersma, P. M. (1979). Breaking in West Frisian: a historical and synchronic approach. *Utrecht Working papers in Linguistics, 8*, 1-41.

Tiersma, P. M. (1979). Aspects of the phonology of Frisian based on the language of Grou. *Meidielingen fan de stúdzjerjochting Frysk oan de Frije Universiteit yn Amsterdam, 4*.

Tiersma, P. M. (1980). *The lexicon in phonological theory*. PhD dissertation, Indiana University.

Tiersma, P. M. (1982). Local and general markedness. *Language, 58*, 832-849.

Tiersma, P. M. (1983). The nature of phonological representation: evidence from breaking in Frisian. *Journal of Linguistics, 19*, 59-78.

Voeten, C. C. (2021). Individual differences in the adoption of sound change. *Language and Speech, 64*(3), 705-741.

Voeten, C. C., Heeringa, W., & Van de Velde, H. (2022). Normalization of nonlinearly time-dynamic vowels. *The Journal of the Acoustical Society of America, 152*(5), 2692-2710.

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd edn.). Chapman & Hall.

A computer simulation of the reducing effect

Anthe Sevenants[1], Dirk Speelman[1], Freek van de Velde[1], Dirk Pijpops[2]
[1]KU Leuven, [2]Universiteit Antwerpen

The reducing effect (Bybee, 2003) is a common mechanism of phonetic change in language, causing frequent constructions to become phonetically reduced over time (e.g. *I don't know* becomes *dunno*). High-frequency constructions especially are said to reduce faster and more strongly due to "neuromotor automation" (Bybee, 2006, p. 5). Corpus studies show reduction empirically, but cannot explain how communication remains successful despite the phenomenon. We do not know, for example, what requirements keep language users from reducing "too far", avoiding communicative chaos.

To find these requirements, we built a computer simulation with virtual speakers ('agents'). Each agent has a memory of constructions represented as vectors (cfr. Baevski et al., 2020). During communication, these are compared on the basis of phonetic distance to determine what construction was 'heard'. To simulate acoustic reduction, speakers can reduce an exemplar's vector, leading to sparser representations over time.
We show that two requirements are necessary for successful reduction (i.e. reduction that is strongest, but never overly strong, in high-frequency constructions). First, the frequency distribution of constructions must be Zipfian, else the acoustic space will fail to be distributed efficiently among constructions. Second, when applying reduction, speakers should check if they are able to understand their reduced utterance themselves ("re-entrance," Steels, 2003). Without this check, speakers reduce too far, with mass confusion as a consequence.

Our simulation shows that there might be more to the reduction principle than just a link between usage and sparsity, as certain properties inherent to language (Zipfian distribution, inner voice) are indispensable in our model world. An experimental spin-off could help confirm this.

References

Baevski, A., Zhou, H., Mohamed, A., & Auli, M. (2020). *wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations*. arXiv. https://doi.org/10.48550/arXiv.2006.11477

Bybee, J. (2003). *Phonology and Language Use*. Cambridge University Press.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, *82*(4), 711–733. https://doi.org/10.1353/lan.2006.0186

Steels, L. (2003). Language re-entrance and the 'inner voice'. *Journal of Consciousness Studies*, *10*(4-5), 173–185.

The documentation and description of labial-velar and implosive consonants in the Ubangi River basin: A field report

Lorenzo Maselli[1,2,3]
[1]Universiteit Gent, [2]Université de Mons, [3]Tokyo University of Foreign Studies

This contribution is intended as a field report on recent phonetic and phonological documentation conducted by the presenter at Université de Bangui (Central African Republic) between February and March 2025. The aim of the mission was to collect new empirical evidence on the production of labial-velar and implosive consonants in the Bantu, Central Sudanic, and Ubangi languages spoken in the Ubangi River basin. The broader objective is to develop the first comprehensive acoustic and articulatory analysis of these articulations in Central Africa, integrating electroglottographic, aerodynamic, and acoustic data.

The mission yielded over 100 GB of speech material from 125 speakers representing 20-30 linguistic varieties; this constitutes the largest phonetic corpus yet assembled in Central Africa. Recordings include both local lects and the Central African lingua franca, Sango, making this simultaneously the most extensive corpus ever collected for any individual language in the region.

In this field report, I will discuss methodological aspects of data collection, including elicitation protocols, recording conditions, and the use of specialised equipment in the field, all information that may be of practical use to others planning similar missions. I will also cover aspects relative to ongoing data analysis, including measurement selection and Praat processing of the acoustic and articulatory traces. The report will close with reflections on the execution of instrumental phonetic fieldwork, particularly regarding the implementation of laboratory methods in under-documented linguistic contexts.

Before Words: How innate knowledge shapes preverbal infants' use of prosody to express communicative functions

Elanie van Niekerk[1], Caroline Junge[1], Iris-Corinna Schwarz[2], Lisa Gustavsson[2], Ellen Marklund[2], Aoju Chen[1]
[1]Utrecht University, the Netherlands; [2]Stockholm University, Sweden

Preverbal infants systematically vary prosody to express different communicative functions (Esteve-Gibert & Prieto, 2012), but the mechanisms underlying this ability remain unclear. We asked how infants begin to acquire prosodic form–meaning mappings, focusing on pitch. We hypothesised that infants first use pitch following innate biases (H1), and gradually rely less on these biases as they gain language-specific knowledge (H2).

We examined the Frequency Code (henceforth, FC; Ohala, 1983), which outlines that smaller larynxes produce higher pitch than larger ones; consequently, speakers use higher pitch or a rising pitch pattern to sound 'small' (e.g., uncertain, as in questions/requests) and lower pitch or a falling pitch pattern to sound 'big' (e.g., confident, as in statements/comments). We compared Dutch- and Stockholm Swedish-exposed infants because Dutch typically uses rising questions and falling statements (Haan, 2001), following FC (Ohala, 1983), whereas Swedish uses falling contours for both (House, 2004).

Monolingual infants (13 Dutch-exposed, 12 Swedish-exposed) participated in 15-minute home play sessions (see Image 1) twice per month at 3, 5, and 7 months. Sessions included four play conditions designed to elicit requests and comments. Audio was segmented (see Image 2) in Praat (Boersma, 2024). Using videos in ELAN (Max Planck Institute for Psycholinguistics, 2024), speech-like vocalisations were coded as requests or comments (analysis-1), and requests were classified as initial or follow-up bids (analysis-2). Mean pitch per vocalisation was extracted using ProsodyPro (Xu, 2013), and the effects of communicative function, language, and age were analysed using linear mixed-effects models (van Niekerk et al., 2024).

Analysis-1 (n=2689 comments, n=1454 requests) revealed a main effect of function ($p<0.001$) and function×age interaction ($p<0.001$) with requests having higher pitch than comments across ages, but with larger differences at 7 than at 3 months. Analysis-2 (n=98 initial requests, 980 follow-up requests) revealed function×age interaction ($p<0.01$) with 3-month-olds marking follow-up requests with higher pitch than initial requests, but 7-month-olds marking both request types with similarly high pitch. No language-related effects emerged in either analysis.

Infants systematically used pitch cross-linguistically, with patterns evolving between 3–7 months yet remaining consistent with the FC (Ohala, 1983). Results support H1 but not H2, suggesting biologically motivated biases underpin prosodic form–meaning mappings throughout the early preverbal phase.

References

Esteve-Gibert, N., & Prieto, P. (2012). Prosody signals the emergence of intentional communication in the first year of life: Evidence from Catalan-babbling infants. *Journal of Child Language*, *40*(4), 919–944. https://doi.org/10.1017/S0305000912000359

Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica, 40*(1), 1–18. https://doi.org/10.1159/000261678

Haan, J. (2001). *Speaking of questions: An exploration of Dutch question intonation* (Doctoral dissertation, Netherlands Graduate School of Linguistics [LOT]).

House, D. (2004). Final rises and Swedish question intonation. In *Proceedings of Fonetik 2004*.

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 6.4) [Computer software]. https://www.praat.org/

Max Planck Institute for Psycholinguistics, The Language Archive. (2024). *ELAN* (Version 6.9) [Computer software]. Retrieved from https://archive.mpi.nl/tla/elan

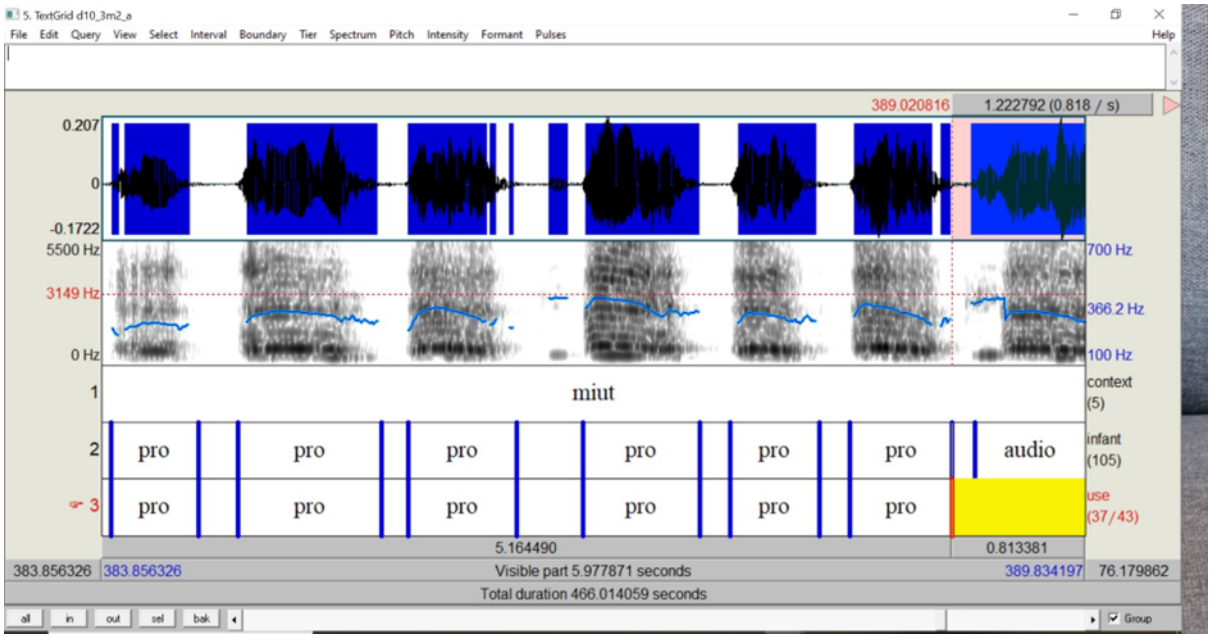
Xu, Y. (2013). ProsodyPro: A tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, 7–10. Aix-en-Provence, France.

van Niekerk, E., Junge, C., & Chen, A. (2024). Role of innate mechanisms in the acquisition of prosodic form-meaning mappings. *AsPredicted*. https://aspredicted.org/pe4hc.pdf

Images and Figures

Image 1: Observational set-up in home setting with participant-led positioning

Image 2: Vocalisation segmentation in Praat using audio recordings



Word stress in self-supervised speech models: A cross-linguistic comparison

Martijn Bentum[1], Louis ten Bosch[1], Tomas O. Lentz[2]
[1]Centre for Language Studies, Radboud University
[2]Department of Communication and Cognition, Tilburg University

Self-supervised speech models (S3Ms) learn general-purpose representations of spoken language that can be fine-tuned for tasks such as speech recognition, speaker identification, and emotion detection. Yet the end-to-end nature of S3Ms makes them difficult to interpret. A common approach is diagnostic classification, where simple classifiers probe model layers for linguistic information. Prior work has shown that S3Ms encode phonetic, semantic, syntactic, and prosodic cues (e.g., Pasad, Chou & Livescu, 2021; Bentum, ten Bosch & Lentz, 2024). This study extends that approach to investigate how word stress is represented in S3Ms across languages. Specifically, we investigate the S3M representations of word stress for five different languages: Three languages with variable or lexical stress (Dutch, English and German) and two languages with fixed or demarcative stress (Hungarian and Polish).

Word stress refers to the relative prominence of syllables within words (Gussenhoven, 2004), realized acoustically through correlates such as duration, intensity, pitch, spectral tilt, and formant peripherality (e.g., van Heuven, 2018). These cues vary in reliability across languages: in fixed-stress languages (e.g., Hungarian, Polish) stress occurs in predictable positions, while in variable-stress languages (e.g., Dutch, English, German) stress is lexically distinctive and less predictable. Human listeners show corresponding sensitivity, with “stress deafness” observed in fixed-stress language speakers (PeperKamp, Vendelin & Dupoux, 2010).

We used the multilingual Wav2vec 2.0 XLS-R model (Babu et al., 2021), trained on 128 languages, and examined bisyllabic words in read-aloud sentences from Common Voice. Stress labels were assigned using CELEX for variable-stress languages and rule-based methods for fixed-stress languages. Classifiers were trained on both acoustic features and model embeddings extracted from different model layers.

Results show that stress can be reliably decoded from S3M embeddings in all five languages, with peak performance around transformer layer 17. Unlike acoustic correlates, model representations consistently revealed language-specific clustering, separating fixed- from variable-stress languages. These findings suggest that S3Ms encode abstract, language-specific stress representations beyond acoustic correlates, offering new insights into how prosody is captured in multilingual models

References

Babu, A., Wang, C., Tjandra, A., Lakhotia, K., Xu, Q., Goyal, N., Singh, K., von Platen, P., Saraf, Y., Pino, J., Baevski, A., Conneau, A., & Auli, M. *XLS-R: Self-supervised cross-lingual speech representation learning at scale*. arXiv preprint arXiv:2111.09296, 2021.

Bentum, M., ten Bosch, L., & Lentz, T. (2024). The processing of stress in end-to-end automatic speech recognition models. In *Proceedings of Interspeech 2024* (pp. 2350-2354).

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press.

Pasad, A. Chou, J. C.& Livescu, K. (2021). Layer-wise analysis of a self-supervised speech representation model. In *2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*. (pp. 914–921).

Peperkamp, S., Vendelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics*, 38(3), 422-430.

van Heuven, V. J. (2018). Acoustic correlates and perceptual cues of word and sentence stress: towards a cross-linguistic perspective. In R. Goedemans, J. Heinz and HulstH. van der (Eds.), *The Study of word Stress and accent: Theories, Methods and data.* (15–59). England: Cambridge University Press.

Yes/no question prosody in Mɛfɛgbɛ (Tɔŋugbe) Ewe

Man Yan Priscilla Lam[1,2], Yiya Chen[1,2]
[1]Leiden University Centre for Linguistics, [2]Leiden Institute for Brain and Cognition

In question prosody research, high-pitched, rising intonation has long been considered a strong cross-linguistic universal (e.g., Bolinger 1978; Ohala 1983, 1984). This pattern has been attributed to the frequency code (Gussenhoven 2004) and is corroborated by evidence from Chinese tone languages (see review in Chen 2022). However, this view is challenged by lax question prosody, a regional feature observed in many African languages, where a different set of features are used to encode questions: falling pitch, vocalic lengthening, breathy termination and sentence-final open vowel (Rialland 2009; Downing & Rialland 2017). While Rialland (2007) distinguishes lax prosody from tense prosody (high-pitched markers), a handful of African languages exhibit features from both (Cahill 2012, 2013; Salffner 2017).

This study examines yes/no questions in Mɛfɛgbɛ (Tɔŋugbe), a variety of Ewe (Kwa, Niger-Congo) spoken in Ghana. Using a semi-controlled experimental design, we investigate if, and how, Mɛfɛgbɛ utilizes lax prosody. Acoustic and statistical analyses of the speech data collected through an interactive game (26 speakers; 1,324 yes/no questions, 1,329 statements) reveals that in the final rhyme of the utterance, questions are distinguished from statements by various features: rising f0 contour, larger f0 range, higher f0 mean, higher intensity, and increased breathiness. While some of these features align with the predictions of lax prosody, others do not. Mɛfɛgbɛ therefore does not present a clear case of lax prosody. These findings highlight interesting issues for discussion. First, the rising pattern is observed across all tested final lexical tone conditions (H, M, L). This echoes findings from Cantonese and Mandarin (Yuan 2004; Ma et al. 2011; Xu & Mok 2011; Chen 2022), showing that Mɛfɛgbɛ shares similarities with Chinese tone languages in how yes/no questions are prosodically marked. Furthermore, the mixed set of features employed by Ewe bears resemblance to Ikaan, which, as discussed by Salffner (2017), invites reconsideration of lax prosody analysis, in relation to the frequency code.

References

Bolinger, D. (1978). Intonation across languages. In: Greenberg, J. (Ed.), Universals of Human Language. Stanford University Press, Stanford, pp. 371–425.

Cahill, M. (2012). Polar question intonation in Kɔnni. In *Selected Proceedings of the 42nd Annual Conference on African Linguistics*: 90–98.

Cahill, M. (2013). Polar question intonation in five Ghanaian languages. In *LSA Annual Meeting Extended Abstracts* (Vol. 4): 10–1.

Chen, Y. (2022). Mind the subtle f0 modifications: The interaction of tone and intonation in Sinitic varieties. *Stellenbosch Papers in Linguistics Plus* 62(2), 113–136. DOI: https://doi.org/10.5842/62-2-904

Downing, L. J. & Rialland, A. (eds.). (2017). *Intonation in African tone languages*. Berlin ; Boston: De Gruyter Mouton.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

Ma, J. K-Y., Ciocca, V. and Whitehill, T. (2011). The perception of intonation questions and statements in Cantonese. *The Journal of the Acoustical Society of America* 129: 1012–1023. https://doi.org/10.1121/1.3531840

Ohala, J. (1983). Cross-language use of pitch: an ethological view. *Phonetica*, 40, 1–18.

Ohala, J. (1984). An ethological perspective on common cross-language utilization of F0 in voice. *Phonetica* 41, 1–16.

Rialland, A. (2007). Question prosody: an African perspective. *Tones and tunes*, 1, 35–64.

Rialland, A. (2009). The African lax question prosody: Its realisation and geographical distribution. *Lingua* 119(6). 928–949. DOI: https://doi.org/10.1016/j.lingua.2007.09.014

Salffner, S. (2017). West African languages enrich the frequency code: Multi-functional pitch and multi-dimensional prosody in Ikaan polar questions. *Laboratory Phonology* 8(1).

Xu, B. & Mok, P. (2011). Final rising and global raising in Cantonese intonation. *Proceedings of the 17th International Congress of Phonetic Sciences*: 2173–2176.

Conservatively Radical: Cross-Modal Language Variation in Central Eastern Norwegian Dialect Speech

Eva van Kampen[1], *Remco Knooihuizen*[2]
[1]University of Groningen, [2]University of Groningen

Norwegian is characterised by variation in both spoken and written form, and between and within varieties (Røyneland, 2009). Bokmål, the focus of this study, is the most widespread of the two written standards and allows for variation in morphology and orthography (Fjeld, 2015; Papazian, 2002). In written form, these variants are labelled as conservative and radical. The variation in writing mirrors morphological and phonetic variation in the central Eastern Norwegian (CEN) dialects, but patterns of variation differ across modalities (Helset, 2024; Kola, 2015; Stjernholm, 2018). Inspired by Lundquist et al.'s (2020) study on Northern Norwegian, this study aimed to further our understanding of the relation between variation in written Bokmål and spoken dialect by analysing how variation is processed and translated into speech by CEN dialect speakers.

The results of this study suggest that written texts are processed and reproduced differently based on the task that the participants carried out – reading or retrieving. When reading Bokmål, the phonology most closely corresponding to the orthography was activated and reproduced almost exclusively. This suggests that even when variation exists in the dialectal repertoire, the written variant maps onto the corresponding form resulting in conservative variants when reading out loud. In the retrieval task, radical forms were more frequent, which is thought to be the result of sociolinguistic meaning being encoded in short-term memory rather than form. As a result, though conservative forms remained dominant, the production of radical variants was facilitated if they matched the individual's idiolect. This hypothesis was further supported by the patterns found based on the formality of the stimulus and individual variables.

References

Helset, S. J. (2024). Tilhøvet mellom fastsette og operative normer i bokmål. *Maal og Minne*, *116*(1), 45-80. https://doi.org/10.52145/mom.v116i1.2273

Kola, K.W. (2015). Kampen for tilværelsen: de radikale variantenes posisjon i nyere bokmålstekster. S*språklig Samling - Årbok 2015*. Landslaget for språklig samling

Røyneland, U. (2009). Dialects in Norway: Catching up with the rest of Europe? *International Journal of the Sociology of Language 2009*(196-197), 7–30. https://doi.org/10.1515/IJSL.2009.015

Stjernholm, K. (2018). Oslo: Une ville, deux dialectes ? (S. Harchaoui, Trans.). *Nordiques*, 35, 117–134. https://doi.org/10.4000/nordiques.1014

Early prosodic boundary perception: innate biases in preterm newborns

Jorik Geutjes[1], Caroline Junge[1], Maria-Luisa Tataranno[2], Manon Benders[2] & Aoju Chen[1]
[1]Utrecht University, [2]University Medical Center - Utrecht

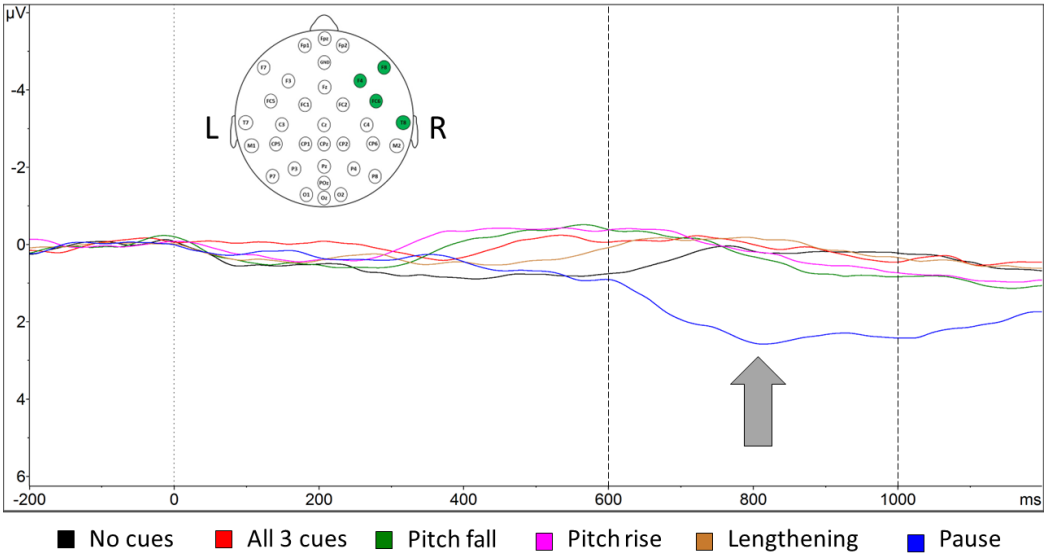
Segmenting continuous speech into meaningful linguistic units is an important first step for newborns acquiring language. The prosodic structure of speech assists in this task. Major speech units, e.g. Intonational Phrases (IPs), are marked by three types of prosodic cues: pitch change, pre-boundary syllable lengthening, and pauses. Although infants can process prosodic structure early on, the underlying mechanisms remain unclear. We hypothesise that infants initially rely on innate perceptual biases to process IP boundaries, namely, physiologically-motivated or cross-species perceptual mechanisms, e.g., the Respiratory Code (RC) and the Iambic-Trochaic-Law (ITL). According to the RC, high pitch is associated with phrase beginnings and low pitch with phrase endings, with pauses between phrases. According to the ITL, lengthened and low-pitched elements occur phrase-finally.

We presented 40 clinically-stable preterm newborns (28-33 weeks of gestation, Dutch-speaking parents) with utterances containing or lacking an IP boundary ([Moni and Lilli and Manu] vs. [Moni and Lilli] [and Manu]), within one week after birth. Boundaries were marked by either one cue or all cues. We measured the EEG component indexing boundary processing, the Closure Positive Shift (CPS), in each condition. We predict that, despite minimal prenatal and postnatal language exposure, preterm newborns can process IP boundaries using the individual cues, based on biologically-motivated principles.

Linear mixed effects modelling shows the CPS was elicited only in the pause condition in the right-frontotemporal region ($p < 0.001$). This implies newborns initially process major prosodic boundaries based on pauses, partially supporting our hypothesis. Associations between prosodic boundaries and other cues may be developed via input-driven learning.

Figure 1

ERP waveforms for frontotemporal electrodes on the right hemisphere, timelocked to the onset of the preboundary syllable.



Note:

Typically, CPS is a positive (downward-facing) deflection observed on frontal electrodes between 500-800ms after detecting a prosodic boundary. This response may be delayed in preterm newborns due to incomplete myelination of the brain (e.g. 600-1000ms, indicated by dashed lines). The gray arrow indicates the CPS, elicited when the boundary in the name sequences is marked by (only) a pause.

Helpt klokje of klȯkje 't maist tegen de kol?
De samenval van hok en bȯk in het Noord-Nederlands

Maureen Prins[1], *Remco Knooihuizenr*[2]
[1]Rijksuniversiteit Groningen, [2]Rijksuniversiteit Groningen

In het standaard Nederlands is het onderscheid tussen de korte o-klanken van hok en bȯk al een lange tijd verdwenen, volgens Van den Toorn (1997, p. 45) een "verandering in het klinkersysteem die ongemerkt tot stand is gekomen". Door de jaren heen is de verschuiving van deze klanken in de literatuur opgemerkt (van Dantzig, 1940; van Loey, 1970; Gilbers & Koster, 2024). In 1999 beschrijft Booij dan ook het Nederlandse klinkersysteem met nog maar één korte o /ɔ/. In sommige noordelijke delen van het land wordt er echter nog steeds onderscheid gemaakt tussen de twee (Grune, 2022).

In dit project wordt onderzocht in hoeverre Groningse en Friese sprekers van het Nederlands nog onderscheid maken tussen de hok- en bȯk-klanken en welke factoren daar een rol in spelen. In ons experiment produceerden Nederlands-Fries/Gronings tweetalige deelnemers (N=95; 52 vrouw; leeftijd M= 54, SD = 20.91) zinnen met daarin woorden die een etymologische hok- of bȯk-klank bevatten. De data is geannoteerd in PRAAT waar ook de F1- en F2-waarden gemeten zijn. Om de overlap tussen de twee klanken te meten is gebruik gemaakt van de Pillai-score en de Bhattacharyya's Affinity. Met een analyse in R is te zien dat er over het algemeen niet veel onderscheid gemaakt wordt tussen de hok en bȯk; de verschuiving is dus in een laat stadium. Een nog lopende analyse zal laten zien welke factoren meespelen bij het wel onderscheid maken tussen de twee klanken.

References

Booij, G. (1999). *The Phonology of Dutch*. Oxford University Press.

Dantzig, B. van. (1940). *De korte o-klanken in het Nederlandsch*. Noordhoff.

Gilbers, D., & Koster, L. (2024). De plaats van /o:/ en /ɔ/ in het Nederlandse klinkersysteem. *Tabu*, 57–74. https://doi.org/10.21827/tabu.2023.41262

Grune, D. (2022). Open en gesloten korte o in een klein deel van Oost-Nederland. Beschikbaar op: http://dickgrune.com/NatLang/Dutch/O_of_O/O_of_O.pdf

Loey, A. van. (1970). *Schönfelds Historische grammatica van het Nederlands : Klankleer, vormleer, woordvorming*. http://ci.nii.ac.jp/ncid/BA91504113

Toorn, M. C. van den (1997). *Geschiedenis van de Nederlandse taal.* Amsterdam University Press.

ABSTRACTS

Postersessie 1

(in alfabetische volgorde)

It's not *what* you say, it's *how* you say it: Creating a prosodic proficiency test battery

Ronny Bujok[1], Lieke van Maastricht[1,2]

[1]Centre for Language Studies, Radboud University, Nijmegen, The Netherlands
[2]Donders Institute for Brain, Cognition, and Behavior, Radboud University, Nijmegen, The Netherlands

Prosody, including rhythm, intonation, and word accent, is crucial to speech communication. Second language learners (L2) who struggle with producing prosody correctly may face difficulties being understood (e.g., Munro et al., 2020; Saito et al., 2016) and social stigma (Derwing, 2003). Hence, mastering prosody improves L2 learners' communication skills, leading to less foreign-accentedness and greater comprehensibility (van Maastricht et al., 2021). Yet, currently, there are no science-based and open-access tools to assess the prosody of L2 learners. Therefore, this project aims to develop a fully automatic online test for L2 learners of English to assess their prosody production. During the test, L2 learners are asked to record themselves reading aloud sentences eliciting various prosodic forms and functions (e.g., narrow focus, types of declarative and interrogative sentences). Our pipeline of automatic speech recognition, forced alignment and prosodic feature extraction quantifies prosodic features such as pitch accent placement, pausing, and speech rhythm. The L2 data are compared against reference data from English natives (L1) to evaluate the prosodic proficiency of the L2 learners. L1 data collection is currently ongoing, so the Phonetics Day will provide us with a perfect opportunity to collect valuable feedback on the planned phonetic analyses for the automated prosody assessment test. For both L2 learners and teachers, this test will provide information on a speaker's current prosodic proficiency, as well as insight into potential areas of improvement, helping learners produce more native-like prosody. In the long run, our test will make prosody training more systematic and accessible, ultimately helping learners to improve communication in their L2.

References

Derwing, T. (2003). What do ESL students say about their accents? Canadian Modern Language Review, 59(4), 547-567. https://doi.org/10.3138/cmlr.59.4.547

Munro, M. J., & Derwing, T.M. (2020). Foreign accent, comprehensibility and intelligibility, redux. Journal of Second Language Pronunciation, 6, 283-309. https://doi.org/10.1075/jslp.20038.mun

Saito, K., Trofimovich, P., Isaacs, T., 2016. Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. Applied Psycholinguistics, 37(2),217–240. https://doi.org/10.1017/S0142716414000502

Van Maastricht, L., Zee, T., Krahmer, E., & Swerts, M. (2021). The interplay of prosodic cues in the L2: How intonation, rhythm, and speech rate in speech by Spanish learners of Dutch contribute to L1 Dutch perceptions of accentedness and comprehensibility. *Speech Communication*, *133*, 81-90. https://doi.org/10.1016/j.specom.2020.04.003

Flexibility of vowel categorisation in newly learned words

Stephanie Cooper, Brechtje Post
University of Cambridge

Listeners can accommodate variation from a range of speakers and contexts to decode the speech signal. Theories differ regarding whether this variation is removed early in processing (e.g. McClelland & Elman, 1986) or coded into lexical representations (e.g. Goldinger, 1998), and whether prior exposure to variation improves future comprehension (e.g. Cooper & Cooper, 2023). These questions are particularly relevant in word learning; new lexical representations may be based on experience from a single context or speaker, thus formed with minimal variation.

This study examined how listeners from monodialectal backgrounds would identify variants of newly learned monosyllabic words that differed in their central vowel. A group of Standard Southern British English (SSBE) speakers and a group of New Zealand English (NZE) speakers were taught new words containing DRESS and LOT vowels from their respective native varieties. Another SSBE-speaking group learned the words in both varieties simultaneously. In the test phase, participants were then asked whether vowel variants of these words were correct pronunciations.

Acceptance and response time results indicate that participants were able to accept substantial variation in vowels, with minimal processing costs, even in words learned with strict vowel uniformity. This supports the argument that vowel perception is uniquely flexible (Cooper & Cooper, 2023; Shaw et al., 2018). However, the group who heard the unfamiliar NZE dialect during training accepted NZE variants at a near-native level. This suggests that speech perception can readily handle some variation, but prior exposure facilitates this further. Hybrid theories of speech perception (e.g. Goldinger, 2007) are therefore supported.

References

Cooper, S., & Cooper, S. (2023). Exposure-independent comprehension of Greek-accented speech: Evidence from New Zealand listeners. *Proceedings of the 20th International Congress of Phonetic Sciences*, 6–10.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.

Goldinger, S. D. (2007). A complementary-systems approach to abstract and episodic speech perception. *Proceedings of the 16th International Congress of Phonetic Sciences*, 49–54.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1–86.

Shaw, J., Best, C. T., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., & Mulak, K. E. (2018). Resilience of English vowel perception across regional accent variation. *Laboratory Phonology, 9*(1), 1–36.

The effect of neural immaturity after gestational diabetes on rapid auditory processing abilities in newborns

De Groot, E.R.[1], Ganga, R.[1*], Tataranno, M.L.[2], Wijnen, F.[1], Chen, A.[1]*
[1] Institute for Language Sciences, Utrecht University, The Netherlands,
[2] Department of Neonatology, UMC Utrecht, The Netherlands.
[*]These authors have contributed equally to this work.

Gestational diabetes mellitus (GDM) affects 5-13% of pregnancies worldwide (Zhu & Zhang, 2016). GDM leads to poorer fetal neural maturation and connectivity, which are associated with poorer cognitive development (Rodolaki et al., 2023) and recognition memory (deRegnier et al., 2000). Moreover, GDM impacts language development (Dionne et al., 2008; Sells et al., 1994). However, why GDM leads to language-related deficits is yet unknown.

As GDM only impacts children prenatally and the fetal neural language network is shaped by prenatal experiences, GDM-exposed children's postnatal language delays may be related to GDM's negative impact on fetal neural maturation. Neural immaturity affects the extent to and speed of sound processing. The ability to distinguish auditory stimuli presented in rapid succession, or rapid auditory processing (RAP) is impacted in infants with a family history of language impairments, explaining much variance in later language outcome (Benasich et al., 2006). We thus hypothesize that GDM affects newborns' RAP.

To test this hypothesis, adopting Benasich et al.'s (2006) EEG oddball paradigm and pure tone stimuli, we assessed RAP in 30 full-term newborns (10 GDM and 20 control; see table 1 for demographics) within 120 hours postnatally. The amplitude of the Mismatch Response (MMR) and the latency and peak amplitude of the Auditory-Evoked Potentials (AEP; N250), were assessed (Benasich et al., 2006).

There were no significant differences between the GDM-group and control group in N250 latency, peak amplitude and MMR amplitude. A clear positive MMR was seen in the right anterior region (see figure 1), similar to Benasich et al. (2006).

Blood glucose levels were well-regulated in the GDM group, which might serve as a protective factor against neural immaturity due to GDM. A larger GDM sample with more variation in blood glucose regulation is needed to validate these findings.

References

Benasich, A. A., Choudhury, N., Friedman, J. T., Realpe-Bonilla, T., Chojnowska, C., & Gou, Z. (2006). The infant as a prelinguistic model for language learning impairments: predicting from event-related potentials to behavior. *Neuropsychologia, 44*(3), 396-411. https://doi.org/10.1016/j.neuropsychologia.2005.06.004

deRegnier, R. A., Nelson, C. A., Thomas, K. M., Wewerka, S., & Georgieff, M. K. (2000). Neurophysiologic evaluation of auditory recognition memory in healthy newborn infants and infants of diabetic mothers. *The Journal of pediatrics, 137*(6), 777-784. https://doi.org/10.1067/mpd.2000.109149

Dionne, G., Boivin, M., Séguin, J. R., Pérusse, D., & Tremblay, R. E. (2008). Gestational diabetes hinders language development in offspring. *Pediatrics, 122*(5), e1073-e1079. https://doi.org/10.1542/peds.2007 -3028

Rodolaki, K., Pergialiotis, V., Iakovidou, N., Boutsikou, T., Iliodromiti, Z., & Kanaka-Gantenbein, C. (2023). The impact of maternal diabetes on the future health and neurodevelopment of the offspring: a review of the evidence. *Frontiers in Endocrinology, 14,* 1125628. https://doi.org/10.3389/fendo.2023.1125628

Sells, C. J., Robinson, N. M., Brown, Z., & Knopp, R. H. (1994). Long-term developmental follow-up of infants of diabetic mothers. *The Journal of pediatrics, 125*(1), S9-S17. https://doi.org/10.1016/S0022-3476(94)70 170-9

Zhu, Y., & Zhang, C. (2016). Prevalence of gestational diabetes and risk of progression to type 2 diabetes: a global perspective. *Current diabetes reports, 16,* 1-11. https://doi.org/10.1007/s11892-015-0699-x

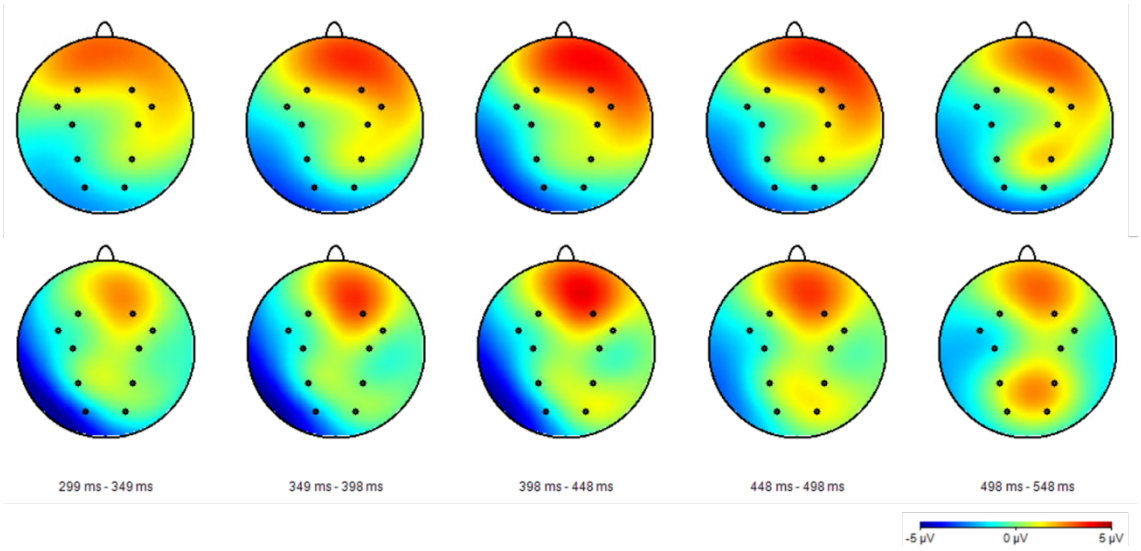
Table 1. Demographics.

	GDM (n=10)	**Control (n=20)**	**P-value**
Gestational age at birth (weeks)	39.00 ± 1.02	39.94 ± 1.20	.044*
Postnatal age at EEG (days)	1.50 ± 1.43	1.05 ± 0.89	.506
Birthweight (grams)	3366.50 ± 380.49	3703.10 ± 424.89	.067
Head circumference (cm)	33.67 ± 1.52	34.57 ± 1.21	.114

Sex (%male)	20%	60%	.044*
Delivery mode *Emergency cesarean section* *Planned cesarean section* *Spontaneous vaginal birth* *Instrumental vaginal birth*	 10% 90%	 15% 30% 50% 5%	
Maternal age (years)	34.50 ± 4.72	32.85 ± 4.40	.366
Maternal pregnancy BMI	32.03 ± 5.07	25.86 ± 4.57	0.004**

The GDM and control group differed with regard to gestational age at birth, which is expected due to the relatively high birthweight of GDM infants, resulting in labor being frequently induced at a slightly lower gestational age. P-values are calculated using a Mann Whitney U test and Chi square test (latter only for sex at birth).

Figure 1. Topography of mismatch response in control (top) and GDM (bottom) groups. The mismatch response is calculated as the difference between the pre-deviant standard and the deviant stimulus.



Auditory kernels – learning acoustic building blocks from speech

Dimme de Groot[1], Odette Scharenborg[1], Jorge Martinez[1]
[1]Technische Universiteit Delft

Unlike linguistic or lexical units such as phonemes or syllables, there are no obvious "information-carrying" acoustic units directly observable in the speech waveform. According to the efficient coding theorem, however, the auditory system should encode incoming sensory information as compactly as possible. Smith and Lewicki (2006) demonstrated that, when learning sparse acoustic building blocks—referred to here as auditory kernels—from speech, the resulting kernels closely resemble reverse-correlation (revcor) filters measured in cat auditory systems. In this work, we extend their analysis to a large cross-linguistic dataset comprising speech of 102 languages. We learn the auditory kernels from each language and examine their statistical properties. We find that the kernels learned on the different languages have a remarkably similar spectral centroid-spread relationship. We also find that, irrespective of language, around 10 kernels are used to represent the content below 500 Hz. These results suggest that this representation might be universal and encourage further research.

The pitch patterns of trisyllabic Japanese loanwords in Standard Chinese by native speakers

Jueyu Hou, Tim Laméris
Leiden University

Recent urban varieties of Mandarin among younger speakers include Japanese phonemic loanwords, e.g., *yàsāxī* 'gentle'. Unlike most Chinese words, typically disyllabic, many Japanese loanwords are polysyllabic and their source forms carry lexically specified pitch patterns, such as L-H-H-H (e.g., *yasashii* 'gentle'). Focusing on trisyllabic loanwords, this study addresses: 1) What pitch patterns do native Mandarin speakers without systematic exposure to Japanese produce? Do they follow Chinese character-based pronunciation (i.e., *yàsāxī* falling-high-high), the Japanese pitch pattern, or something else? 2) What predicts such patterns?

A real-word production task was conducted with 40 native Mandarin speakers (aged 18–35, 20F). After a language-experience questionnaire (LHQ3) [1] and a 2AFC identification task on lexical familiarity with six frequent trisyllabic loanwords, participants read these words aloud in three context conditions: isolated word, in (contextualized and controlled) carrier-sentences (all declarative, with critical words sentence-medial). The Chinese character-based tonal patterns of these words were dipping-falling-high, falling-high-high, and high-rising-high respectively. 1134 tokens were analyzed after data cleaning.

For question 1, hierarchical agglomerative clustering (Ward's linkage, Euclidean distance) identified two optimal clusters via silhouette analysis, featuring two patterns (Fig. 1): L-H-L (61.29%) and L-H-H (38.71%), contrasting the strong-weak-strong pattern typical of Mandarin trisyllabic words [2]. For question 2, context, lexical familiarity, and character tonal patterns were expected to influence the produced pitch patterns. Generalized Additive Mixed Models (GAMMs) showed that beyond individual variation and word-specific contours, context was the strongest predictor of F0 contour variability, while familiarity and character tonal patterns had minor effects.

To conclude, this study observed unique pitch patterns (neither Japanese-like nor Chinese-like, e.g., Fig. 2) of the elicited trisyllabic Japanese loanwords and a strong context effect. Furthermore, individual variation and word-specific pitch patterns while producing such loanwords require further investigation.

Figure 1. *Average F0 contours of the elicited Japanese phonemic loanwords by cluster.*

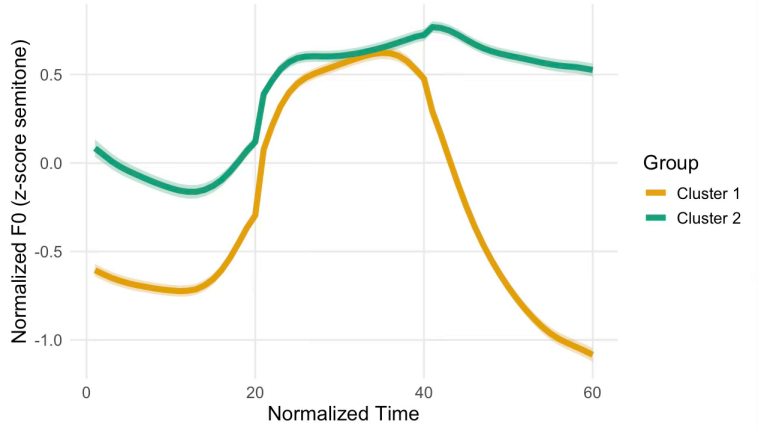
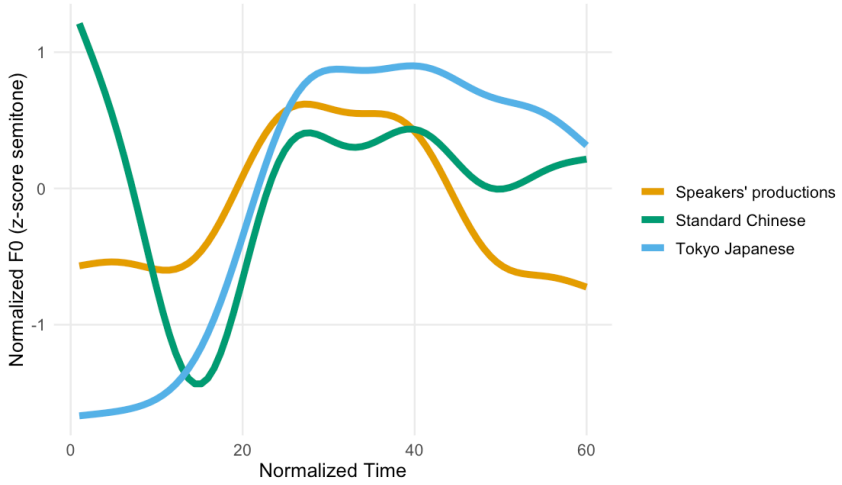


Figure 2. *Average F0 contours of elicited 'yàsāxī' in isolated word condition, compared with the Standard Chinese and Tokyo Japanese pronunciations.*



References

[1] Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition, 23*(5), 938-944. https://doi.org/10.1017/S1366728918001153

[2] Lai, C., Sui, Y., & Yuan, J. (2010). A corpus study of the prosody of polysyllabic words in Mandarin Chinese. *Proceedings of Speech Prosody 2010* (Paper 457). https://doi.org/10.21437/SpeechProsody.2010-33

Multimodal prosodic phrasing in infant-directed speech: testing the cumulative-cue hypothesis with gesture restriction

Roos Ledeboer[1], Victoria Reshetnikova[2], Roy Hessels[3], Aoju Chen[2]

[1]Donders Centre for Cognition, Donders Institute for Brain, Cognition, and Behaviour, Radboud University, [2]Institute for Language Sciences, Utrecht University, [3]Experimental Psychology, Helmholtz Institute, Utrecht University

Human communication is inherently multimodal (McNeill, 1992); the relation between speech and co-speech gestures is well established and facilitates face-to-face interaction (e.g., Church et al., 2017; Cravotta et al., 2019; Esteve-Gibert & Prieto, 2013). However, the underlying cognitive mechanisms remain unclear. Important insights emerge from research on modality disruption and the Cumulative-Cue Hypothesis proposed for prosodic prominence (Ambrazaitis & House, 2023). This hypothesis entails that when facing restrictions in one of the two modalities, speakers redistribute communicative effort within the same modality and only enhance the cues in the other modality in cases of additional effort. By investigating the effects of hand gesture restriction on acoustic and visual prosodic phrasing cues in infant-directed speech, the Hypothesis can be extended, offering a broader framework for understanding the relation between speech and co-speech gestures. Given its engaging and effortful nature, infant-directed speech may be particularly susceptible to modality disruptions, providing a suitable test context. Interactions between three German tutors and 12 Dutch infants (aged 4–5 months and 8–9 months) were recorded in two conditions: Hands Free and Hands Restricted. A total of 732 segmented intonational phrases and 61,151 corresponding video frames were analyzed to compare the tutors' acoustic cues (pitch maximum, pitch minimum, final syllable duration, pause duration) and visual cues (eyebrow movement frequency and intensity) at final intonational phrase boundaries between conditions. Mixed-effects modelling showed that under hand restriction, eyebrows were raised more frequently, but only during interactions with younger infants. Acoustic cues remained unaffected by hand restriction for both age groups. The finding for the younger group supports the Cumulative-Cue Hypothesis and suggests it extends to prosodic phrasing. The age-related differences may be related to changes in communicative intent and speakers' adaptability to infants' developmental needs, yet they complicate the generalization of the findings and warrant further research.

References

Ambrazaitis, G., & House, D. (2023). The multimodal nature of prominence: Some directions for the study of the relation between gestures and pitch accents. In O. Niebuhr & M. Svensson Lundmark (Eds.), *Proceedings of the 13th International Conference of Nordic Prosody* (pp. 262–273). Sciendo. https://doi.org/10.2478/9788366675728

Church, R. B., Alibali, M. W., & Kelly, S. D. (Eds.). (2017). *Why gesture?: How the hands function in speaking, thinking and communicating.* John Benjamins Publishing Company.

Cravotta, A., Bus, . M. G., & Prieto, P. (2019). Effects of encouraging the use of gestures on speech. *Journal of Speech, Language, and Hearing Research*, *62*(9), 3204–3219. https://doi.org/10.1044/2019_JSLHR-S-18-0493

Esteve-Gibert, N., & Prieto, P. (2013). Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research*, *56*(3), 850–864. https://doi.org/10.1044/1092-4388(2012/12-0049)

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* University of Chicago Press.

Turn-taking and final boundary tones in Dutch: a corpus study in replication of Caspers (2003)

Ariëlle Reitsema[1]
[1]Leiden University Centre for Linguistics

We report results of a replication study on the role of utterance-final speech melody in the turn-taking system of Dutch (Caspers 2003), using a larger dataset that had not previously been analysed for this purpose. In natural dialogue, interlocutors typically take turns at speaking and listening. To facilitate smooth and accurately timed speaker transitions, listeners need to gauge whether the speaker intends to continue speaking or to yield the turn at any moment. Among an array of possible end-of-turn cues, the present study examines the patterning of utterance-final boundary tones and syntactic completion in relation to observed instances of turn-keeping and turn-changing in a corpus of Dutch task-oriented spontaneous dialogue. These materials – consisting of eight map-task dialogues (Anderson et al., 1991) – are highly comparable to those analysed by Caspers (2003), allowing for a precise replication and verification of the original findings. The corpus has been divided into inter-pausal units (Koiso et al., 1998) – stretches of single-speaker speech bounded by turn-changes or pauses (>100ms). For each IPU, the turn transition type, final melody (following ToDI), and syntactic completion (Ford & Thompson, 1996) are being annotated. The following hypotheses will be tested: (1) IPUs ending in a level boundary tone (%) will most often be followed by a same-speaker continuation (turn-keeping), while the rising (H%) and low (L%) boundary tones are not expected to pattern consistently with either turn-keeping or changing by themselves; (2) Level boundary tones are expected to pattern with syntactically incomplete IPU-boundaries, and rising and low boundary tones with syntactically complete boundaries. Once verified with this new corpus study, these findings will serve as a stepping stone towards follow-up investigations.

References

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The Hcrc Map Task Corpus. *Language and Speech*, *34*(4), 351–366. https://doi.org/10.1177/002383099103400404

Caspers, J. (2003). Local speech melody as a limiting factor in the turn-taking system in Dutch. *Journal of Phonetics*, *31*(2), 251–276. https://doi.org/10.1016/S0095-4470(03)00007-X

Ford, C. E., & Thompson, S. A. (1996). Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 134–184). Cambridge University Press. https://doi.org/10.1017/CBO9780511620874.003

ToDI second edition (2019, version 2.3) *Transcription of Dutch Intonation, an interactive course.* https://todi.cls.ru.nl/ToDI/home.htm

Koiso, H., Horiuchi, Y., Tutiya, S., Ichikawa, A., & Den, Y. (1998). An Analysis of Turn-Taking and Backchannels Based on Prosodic and Syntactic Features in Japanese Map Task Dialogs. *Language and Speech*, *41*(3–4), 295–321. https://doi.org/10.1177/002383099804100404

Velar or uvular? The Standard German ach-laut revisited.

Zoe Tholen
University of Amsterdam

While the so-called “ach-laut” in Standard German (SG) is commonly transcribed as the velar [x], the current investigation presents a multi-perspective view on the question if a transcription as the uvular [χ] might reflect SG speech more accurately. Two experiments explored the realization and judgment of the ach-laut, specifically examining the above question under the hypothesis that differences might be due to regional variation. During a word-list reading task, all speakers exhibited uvularity in at least some productions of the ach-laut, but also showed both within- and between-subject variation in uvular scrape and center of gravity. While the dialectal background of the speaker did not significantly predict between-subject variation in the phonetic realization, acoustic measures such as the duration of the of the uttered sound emerged as possible predictors. The speech data most likely points to a combination of idiolectal and situational phonetic factors determining the uvularity of the ach-laut, rather than dialect or a sometimes attested complementary allophony of [x] and [χ]. A perception experiment in which participants placed recordings of velar and uvular stimuli on map, depending on where they thought the speaker might be from, reflected a lack of discrimination between [x] and [χ] across all participants. The results of these experiments call for a more specific acoustic description of uvular sounds and, in combination with a review of early and current literature on the ach-laut (which has often confused velar and uvular) they indicate above all that a historical bias against a transcription as uvular might be more responsible for the common [x]-transcription than actual speech patterns.

The evolution of uptalk in Standard Southern British English

Alanna Tibbs, Jiseung Kim, Amalia Arvaniti
Radboud University

Early 21st century studies of uptalk, the use of statement-final pitch rises instead of falls, in Standard Southern British English (SSBE) indicate that at the time uptalk was an innovation: it was infrequent, though increasingly used (Bradford, 1997), especially by women (Barry, 2008) and highly variable in form, with no systematic distinction from other types of rises (Shobbrook & House, 2003). To investigate its current state, we analysed 977 rising utterances from 29 SSBE speakers (19 female) reading scripted dialogues intended to elicit rises indicating requests for confirmation, uncertainty, negotiation, polar questions, listing, and sarcasm (Table 1). The first three functions are associated with uptalk (Arvaniti & Atkins, 2016) and are predicted to have similar shapes, but should differ from non-uptalk rises if uptalk is now an established SSBE feature. Pitch contours were submitted to Functional Principal Component Analysis (Ramsay et al., 2020); the first three principal components (PCs) reflected differences regarding the starting point and extent of the rise (PC1), the scaling of the rise start (PC2), and the rise's overall convex or concave shape (PC3); see Figure 1. LMERs of the PC coefficients of the input curves and a Generalized Additive Mixed Model (GAMM) (van Rij et al., 2022) confirmed the importance of these differences and showed that gender did not affect shape, but pragmatic category did (Figure 1). The uptalk contours were comparable. Sarcasm had a similar shape but scaled lower, polar questions demonstrated a fall-rise pattern absent from uptalk, while listing had a rise-plateau shape. The results support the prediction that uptalk is establishing, since its form is now stable and distinct from rises of competing functions, with no gender differentiation in either frequency or form.

Table 1. *Sample dialogues; the target words are bold and underlined.*

Pragmatic Categories	Context	Response
Confirmation Request	*What name is your appointment under?*	***Alanna****?*
Uncertainty	*What time did you get back last night?*	***Eleven****?*
Negotiation	*What should we make for dinner tomorrow?*	***Lasagna****?*
Polar Question	*I visited England recently.*	*Was it* ***rainy****?*
Listing	*Who did you go to the pub with?*	*I went with* ***Melinda****, Malena, and Yolanda.*
Sarcasm	*Why is David Beckham supporting Qatar?*	*Because* ***money****?*

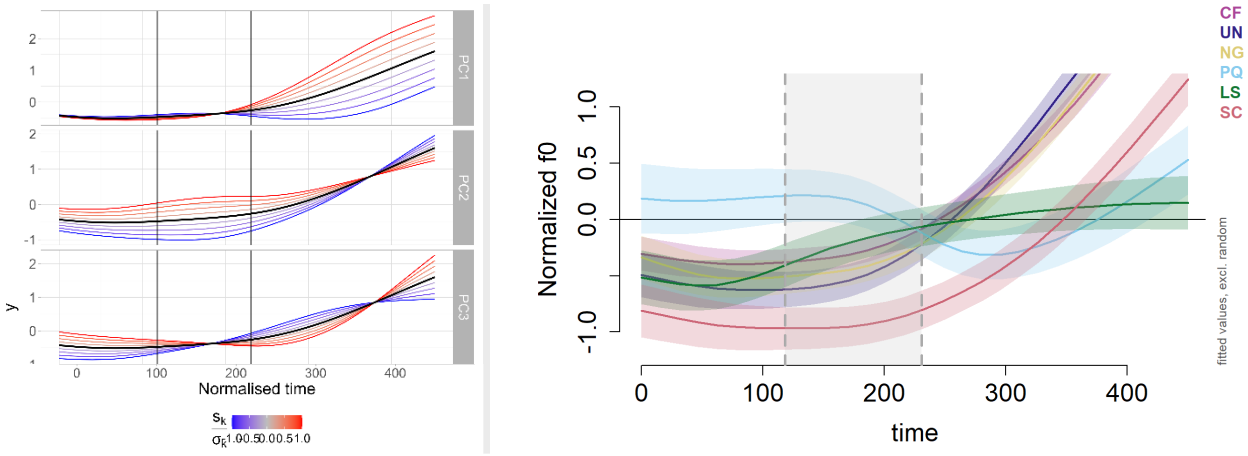


Figure 1. (a) The first three PCs; the lines are PC curves with higher (red) and lower (blue) than the mean (black) PC coefficients. (b) GAMM of the smoothed curves for each pragmatic category; the shaded interval denotes the accented vowel.

References

Arvaniti, A., & Atkins, M. (2016). Uptalk in Southern British English. In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), *Proceedings of Speech Prosody.* Boston: Boston University, pp.153- 157. https://doi.org/10.21437/SpeechProsody.2016-32

Barry, A. S. (2008). *The form, function and distribution of high rising intonation: A Comparative Study of HRT in Southern Californian and Southern British English*. VDM Verlag.

Bradford, B. (1997). Upspeak in British English. English Today, 13(3), 29-36. https://doi.org/10.1017/S0266078400009810

Ramsay, J. O., Graves, S., & Hooker, G. (2020). *fda: Functional Data Analysis*. R package version 5.1.5.1.

Van Rij, J., M. Wieling, R. Baayen & H. van Rijn. (2022). *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*. R package version 2.4.1.

Shobbrook, K., & House, J. (2003). High rising tones in southern British English. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *15th International Congress of Phonetic Sciences.* Barcelona: Universitat Autonoma de Barcelona, pp.1273-127.

Accommodation to a tongue height obstruction: Acoustic outcomes and aftereffects

Xinyu Zhang, Rob Schoonen, Esther Janse
Radboud University

Language acquisition involves learning the mappings between articulatory configurations and their somatosensory and auditory outcomes, creating a close interaction between speech production and perception. Previous work has shown that these processes can influence one another both in infancy and in adulthood.

This study investigates how articulation manipulation and access to one's own auditory feedback affect speech production. In earlier work (Zhang et al, 2023), we found that speaking with a tongue-height obstruction did not alter sound categorization behavior along an /ɪ/–/ɛ/ continuum. Here, we examine the corresponding production data. We compared /ɪ/ and /ɛ/ productions under normal and obstructed conditions, and further assessed the role of auditory feedback by contrasting speakers with access to their own obstructed speech against those whose feedback was masked by speech-shaped noise.

Results demonstrate that the obstruction altered both F1 and F2, albeit in different directions and magnitudes, and that auditory feedback does play a role in adaptation to the altered articulatory configuration. Crucially, no aftereffects emerged: once the obstruction was removed, speakers reverted to their baseline productions. Together with the lack of perceptual change, these findings point to the stability of stored speech sound representations, suggesting that short-term articulatory perturbations do not readily reshape underlying phonological categories.

References

Zhang, X., Schoonen, R., Janse, E. (2023) Sound categorization after speaking with a bite block [pdf] In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th International Congress of the Phonetic Sciences (ICPhS) (pp. 157-161)*.

ABSTRACTS

Postersessie 2

(in alfabetische volgorde)

Weighing auditory, visual and semantic cues in lexical stress perception in Spanish

Floris Cos[1,2], Lieke van Maastricht[1,2], Hans Rutger Bosker[2], Matteo Maran[2] & Esther Janse[1,2]
[1]Centre for Language Studies, Radboud University, Nijmegen, [2]Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen

Simple up-and-down hand movements known as beat gestures influence lexical stress perception: with the same ambiguous acoustic item, listeners may perceive either *CONtent* or *conTENT*, depending on which syllable has a beat gesture aligned to it. This "Manual McGurk effect" has been demonstrated using isolated words in Dutch (Bujok et al., 2025) and notably in Spanish (Rohrer et al., 2025), where lexical stress distinguishes between 1st person present (*repreSENto*, "I represent"), and 3rd person past tense verb forms (*represenTÓ*, "(s)he represented"). The current study investigates how listeners weigh auditory and visual cues to lexical stress in a full, tense-biasing sentence context:

(1) *Anteriormente, *repreSENto / represenTÓ a la ciudad de Málaga.*
"Before, *I represent / (s)he represented the city of Málaga."

In our stimuli, we combined 3 biasing contexts (past tense vs. present tense vs. no bias), 5 auditory versions of the verb form, following an acoustic lexical stress continuum, and 2 visual beat conditions, timed to one of the last two syllables of the verb form. Ninety Spanish natives (42F, 46M, 2 Other; M_{Age} = 28.69, range = 21-40) indicated for each stimulus which pronoun (*yo* "I" or *ella* "she") they thought fit best with the sentence.

The results indicated that the proportion of *yo*-responses decreased as the audio became more *represenTÓ*-like. Compared to the neutral sentence context, the proportion of *yo*-responses increased in present-bias contexts, and decreased in past-bias contexts (Figure 1). Although beat alignment did not influence categorization responses, reaction-time analyses showed that participants responded faster whenever the beat gesture was congruent with the semantic bias (Figure 2). The results thus suggest that in full sentences, beat gestures do not change the ultimate interpretation of the utterance, but congruent (vs. incongruent) gestural timing does facilitate word recognition.

Figure 1.

Proportions of "yo"-responses as a function of Audio Step, Sentence Context and Beat Alignment.

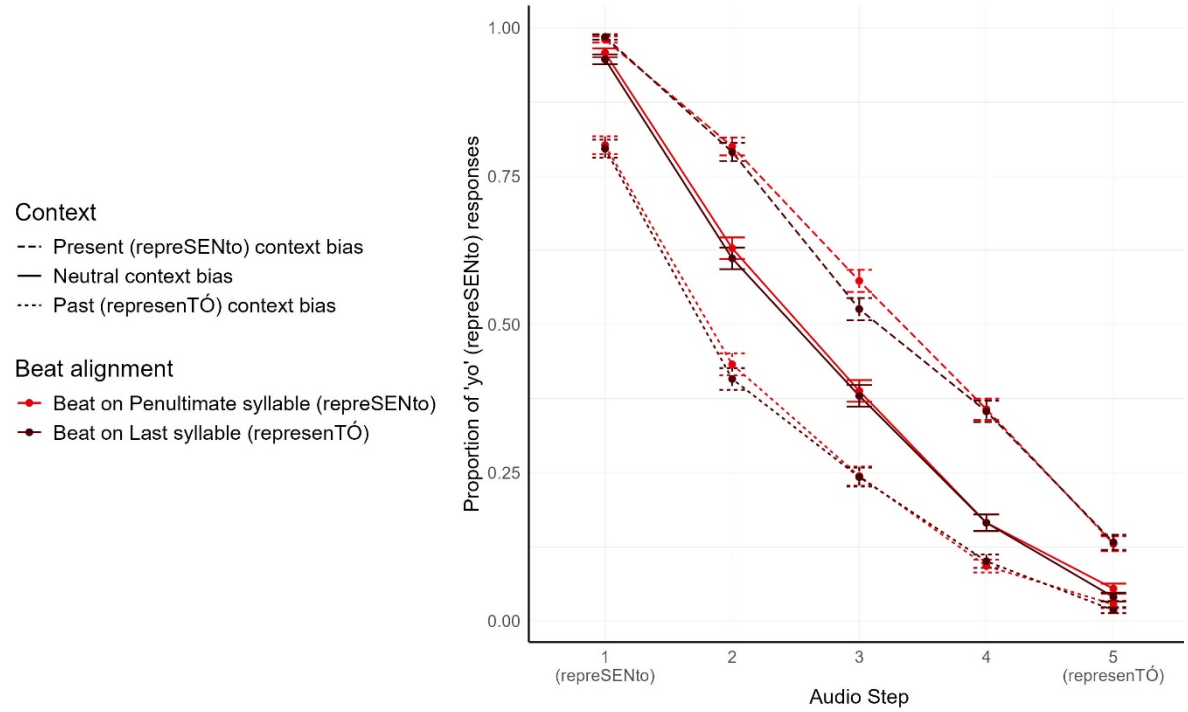
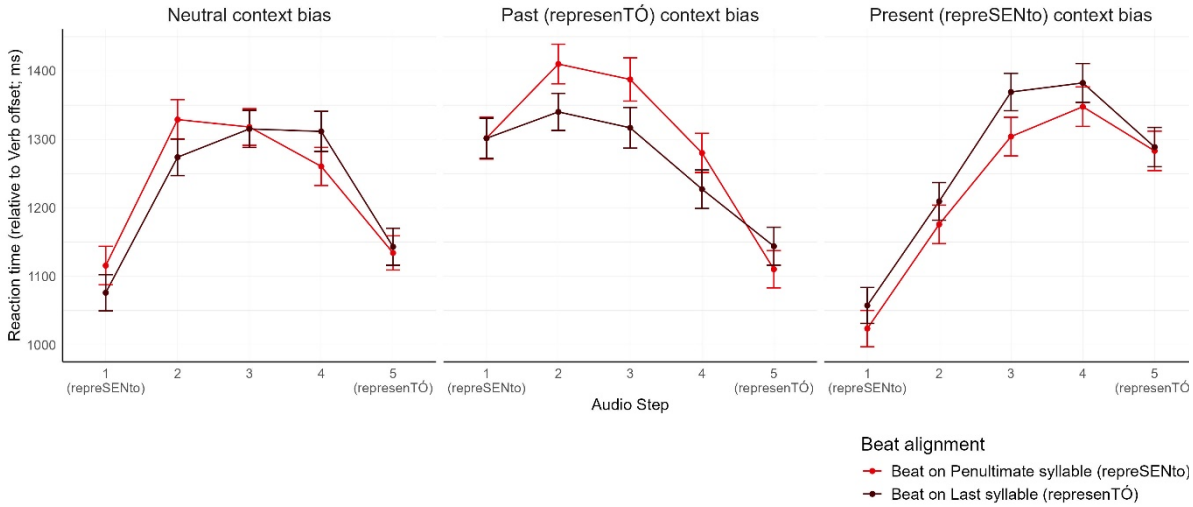


Figure 2.

Reaction times (RTs) for each step of the acoustic continuum and Beat Alignment, separated for each Sentence Context. RTs were measured from the offset of the verb form.



References

Bujok, R., Meyer, A. S., & Bosker, H. R. (2025). Audiovisual perception of lexical stress: Beat gestures and articulatory cues. *Language and Speech*, *68*(1), 181-203. https://doi.org/10.1177/00238309241258162

Rohrer, P. L., Bujok, R., Van Maastricht, L., & Bosker, H. R. (2025). From "I dance" to "she danced" with a flick of the hands: Audiovisual stress perception in Spanish. *Psychonomic Bulletin & Review*, 1-10. https://doi.org/10.3758/s13423-025-02683-9

Musicians vs. non-musicians: who produces sentence stress better? Analysis of fundamental frequency and syllable duration in a negative sentence among French-speaking learners of Dutch

Thomas De Wispelaere[1], Pauline Degrave[1]
[1]Institut Langage & Communication, UCLouvain

French-speaking learners face significant challenges in acquiring Dutch prosody due to fundamental prosodic differences between the two languages. Dutch features, among others, a sentence stress (SS), which serves a contrastive or focus function to emphasize new or contrasted information (Rietveld & Van Heuven, 2016). This SS is produced by an increase of the fundamental frequency in more than 4 semitones, a 10% to 15% longer syllable duration and a greater intensity (Van Heuven, 2018). In contrast, French uses fixed final stress at the end of words and phrases (Di Cristo, 2000) that fulfills a demarcative function. Because of this prosodic difference, French-speaking learners of Dutch often take stress patterns over from their native language. In a negative sentence, they therefore tend to emphasise the negation (at the end of the sentence), whereas the verb form should be stressed (e.g., **Ik weet het niet* instead of *Ik weet het niet*, 'I don't know'; Hiligsmann & Rasier, 2007). Besides, musical training appears to enhance language skills, particularly prosody perception (Degrave, 2022; Jansen, et al., 2023). However, the influence of musical training on prosody production has not yet been investigated on this target group (Rasier & Hiligsmann, 2007). The current study therefore analyses whether French-speaking musicians produce SS acoustically better than non-musicians in the negative sentence *Ik weet het niet* 'I don't know'. 18 musicians and 17 non-musicians pronounced this sentence twice and were recorded as part of a broader study focussing on their SS production (Dejans & Degrave, 2022; Degrave & De Wispelaere, 2025). The results show that musicians do not pronounce acoustically better than non-musicians. The latter make a slightly better use of fundamental frequency, but both groups seem to produce a longer syllable duration on the negation. These results provide more insight into the prosody acquisition of a foreign language and the link with music.

References

Rietveld, T., & Van Heuven, V. (2016). *Algemene fonetiek*. Coutinho.

Van Heuven, V. (2018). Acoustic Correlates and Perceptual Cues of Word and Sentence Stress. Towards a Cross-Linguistic Perspective. In R. Goedemans, J. Heinz, & H. van der Hulst, *The Study of Word Stress and Accent. Theories, Methods and Data*. (pp. 15-59). Cambridge University Press.

Di Cristo, A. (2000). Vers une modélisation de l'accentuation du français [2nd part]. *Journal of French Language Studies*, 10(1), 27-44. https://doi.org/10.1017/S0959269500000120.

Hiligsmann, P., & Rasier, L. (2007). *Uitspraakleer Nederlands voor Franstaligen*. Plantyn.

Degrave, P. (2022). Music training and the use of songs or rhythm: Do they help for lexical stress processing? *International Review of Applied Linguistics in Language Teaching*, 60 (3), 799-824. https://doi.org/10.1515/iral-2019-0081.

Jansen, N., Harding, E. E., Loerts, H., Başkent, D., & Lowie, W. (2023). The relation between musical abilities and speech prosody perception: A meta-analysis. *Journal of Phonetics*, 101, 101278. https://doi.org/10.1016/j.wocn.2023.101278.

Rasier, L., & Hiligsmann, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Cahiers de Linguistique Française*, 28, 41-66.

Dejans, L., & Degrave, P. (2022). Do musicians outperform non-musicians in foreign language prosody production? *SysMys'22*.

Degrave, P., & De Wispelaere, T. (2025). Sentence stress production of French speakers in Dutch: does a musical training help? *EuroSLA34*.

Prosody of Mandarin sentence-final particles in spontaneous conversational corpus

Karina Kechun Li[1,2], Yiya Chen[1,2]
[1]Leiden University Centre for Linguistics, [2]Leiden Institute for Brain and Cognition

Sentence-final particles in Mandarin fulfil versatile communicative functions. Previous research has noted their interaction with sentence intonation (Liu & Xu, 2005), yet the prosody of the particles themselves remains underexplored. Both tonal and intonational factors are relevant: as neutral-tone syllables, their pitch is expected to depend on the preceding lexical tone (Chao, 1968); however, their crucial roles in signalling clause type or speech act suggest that intonation and discourse may also shape their realisations (Hsu & Xu, 2020).

This study investigates the prosody of two Mandarin sentence-final particles, *ma* (吗/嘛) and *ba* (吧), using the MAGICDATA Mandarin Chinese Conversational Speech Corpus (Yang et al., 2022), which comprises about 180 hours of spontaneous conversation from 663 speakers. Recordings were force-aligned with provided transcripts using the Montreal Forced Aligner (McAuliffe et al., 2017), and we focus on the particles that appear at a clause-final position, followed by either a question mark or a period. We tested the effects of preceding tone, clause type (punctuation), and turn position (final or not) on their average pitch. Preliminary results indicate that *ma* has a higher average pitch than *ba*. Preceding tone significantly affects pitch, though the observed patterns deviate from predictions of canonical neutral-tone realisation. By contrast, neither clause type nor turn position showed a robust effect.

Future work will incorporate large language models to refine contextual interpretation, given the inconsistent use of punctuation in the provided transcripts. We also highlight methodological challenges in adapting such large spontaneous corpora for phonetic research, and discuss processing steps that may improve the robustness of prosodic measurements.

References

Chao, Y. R. (1968). *A grammar of spoken Chinese*. University of California Press.

Hsu, Y.-Y., & Xu, A. (2020). Interaction of prosody and syntax-semantics in Mandarin *wh* -indeterminates. *The Journal of the Acoustical Society of America*, *148*(2), EL119–EL124. https://doi.org/10.1121/10.0001676

Liu, F., & Xu, Y. (2005). Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation. *Phonetica*, *62*(2–4), 70–87. https://doi.org/10.1159/000090090

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi. *Interspeech 2017*, 498–502. https://doi.org/10.21437/Interspeech.2017-1386

Yang, Z., Chen, Y., Luo, L., Yang, R., Ye, L., Cheng, G., Xu, J., Jin, Y., Zhang, Q., Zhang, P., & others. (2022). Open source MagicData-RAMC: a rich annotated mandarin conversational (RAMC) speech dataset. *arXiv Preprint arXiv:2203.16844*.

Vowel systems of Standard and Kudar varieties of Iron Ossetic

Varvara Petrova[1], *Timofei Mets*[2]
[1]Higher School of Economics, [2]University of Vienna

Ossetic (< Iranian < IE) is a language spoken in the Caucasian region by circa 576,000 people, the majority of which are native speakers of its Iron dialect (Belyaev, 2021). This dialect is itself subject to regional variation, with Standard Iron having received the most thorough linguistic description (e.g. Dzakhova, 2009). To date, the research concerning the vowel phonemes of another Iron variety, Kudar, has been based exclusively on perception and mainly centered on the peculiarities of their realisation word-initially and in vowel-glide sequences (see e.g. Bekoev, 1985, pp. 178–182) rather than the differences in quality exhibited by the prototypical allophones of Standard and Kudar Iron, despite such differences being noticeable for native speakers.

This study was aimed at investigating the acoustics of vowels of Kudar and Standard Iron using a uniform questionnaire that consisted of monosyllabic words featuring all of the vowels between coronal consonants and disyllabic words with the same vowel phonemes within each word, allowing for quality assessment in regard to stress. The data were collected in Vladikavkaz (North Ossetia, Russia) in 2024 and 2025 from 4 speakers of Standard Iron and 6 speakers of Kudar Iron.

The main differences between the two systems were found to lie in mid vowel subsystems, with /ɜ/ and /a/ being subject to greater variability and /ɜ/ being realized as a higher and more retracted vowel in Standard Iron.

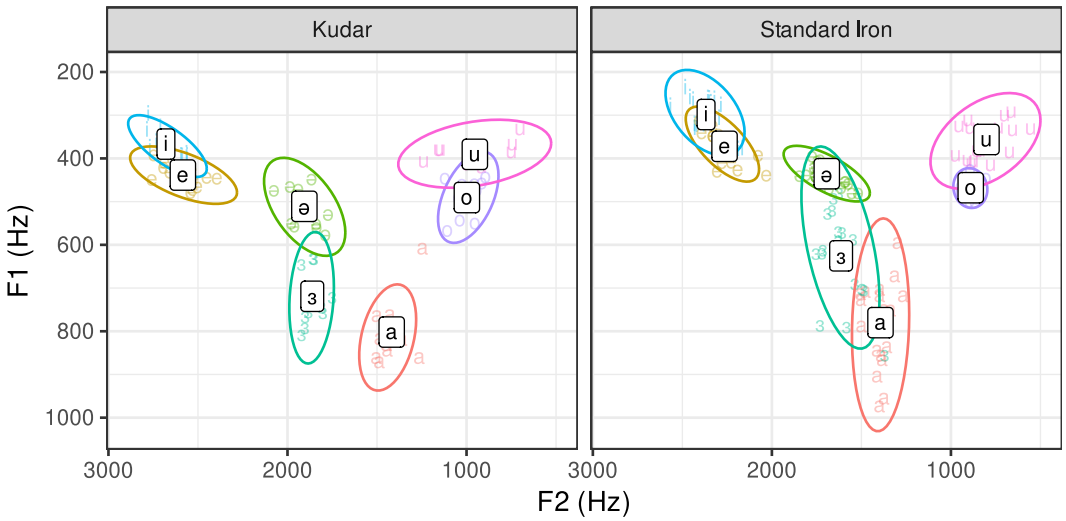


Figure 1. Formants of Kudar and Standard Iron vowels in monosyllabic words.

References

Bekoev, D. G. (1985). *Ironskiy dialekt osetinskogo yazyka [Iron Dialect of the Ossetic Language]*. Iryston.

Belyaev, O. I. (2021, April). *Kontaktno-obuslovlennye yavleniya v osetinskom yazyke [Contact-Indused Phenomena in the Ossetic Language]*. https://iling-ran.ru/workshops/210304_belyaev_handout.pdf

Dzakhova, V. T. (2009). Foneticheskie markery udareniya v osetinskom yazyke (v sopostavlenii s nemetskim) [Phonetic Markers of Stress in Ossetian (in Comparison with German)]. *Slovo i Tekst: Kommunikativnyy, Lingvokul'turnyy i Istoricheskiy Aspekty: Materialy Mezhdunarodnoy Nauchnoy Konferentsii [Word and Text: Communicative, Linguocultural, and Historical Aspects: Proceedings of an International Scientific Conference]*.

Attitudes towards accent variation in word-final nasal vowels in Polish

Weronika Polakowska[1]
[1]University of Amsterdam

Polish is one of the last Slavic languages with nasal vowels (Sussex & Cubberley, 2006). They occur word-medially and word-finally, with the latter exhibiting regional variation. There are three commonly attested variants for the word-final nasal vowels. Nasalised diphthongs [ɔw̃/ɛw̃] are the Standard Polish pronunciation (Gussmann, 2007), considered the most 'correct' in prescriptive literature (Dunaj, 2006). The other variants include non-standard realisations such as denasalisation (i.e. [ɔ/ɛ]) and nasal stopping (i.e. [ɔm/ɛn]). These are viewed as non-standard and thus heavily stigmatised and avoided by speakers (Baranowska & Kaźmierski, 2020; Johnson, 1984). This paper investigates the empirical support for these observations from the literature by means of a speaker-evaluation experiment. It does this by examining attitudes towards accent variation using Polish nasal vowels and comparing the judgements towards standard and non-standard variants.

Eighteen native Polish participants took part in an online matched-guise experiment where they were asked to evaluate speakers based on their suitability to work as a newscaster (cf. Levon & Fox, 2014). The three guises—nasalised diphthongs, denasalisation, and nasal stopping—were each measured along three dimensions: status, solidarity, and pretentiousness (Grondelaers, van Hout & Steegs, 2010; Tamminga, 2017).

Figure 1 summarises the rating averages for the matched-guise task. Linear mixed-effects regression shows that the standard variant (i.e. nasalised diphthong) was judged the most favourably and elicited significantly higher ratings for status and pretentiousness than the non-standard variants. There was no significant difference between denasalisation and nasal stopping. Additionally, although the participants were all from Podlasie, a region characterised by non-standard speech, they notably favoured the standard. The findings highlight the high status and, contrastively, the perceived pretentiousness of the standard language. They also showcase the stigmatisation of non-standard varieties in the face of the standard, in line with previous research.

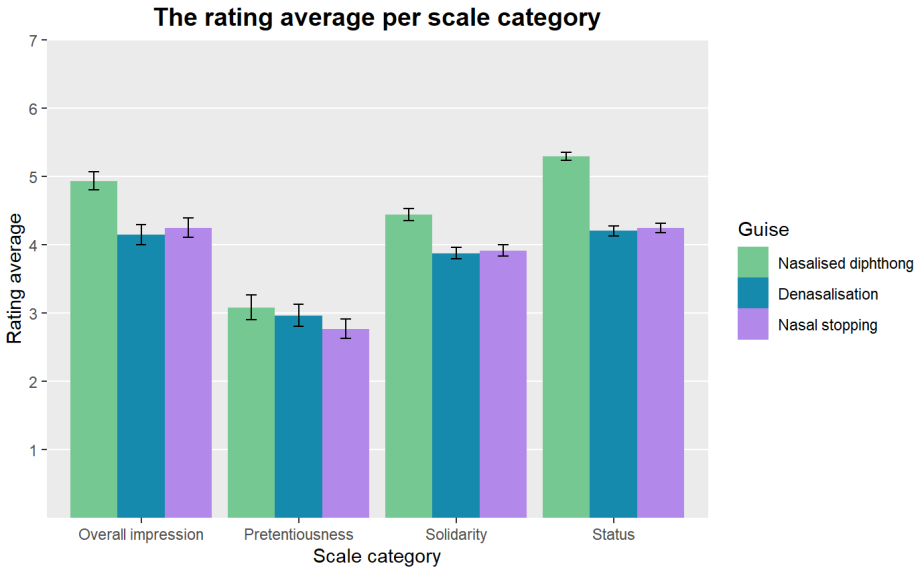


Figure 1. The rating average per scale category. The error bars represent the standard error.

Higher-Order Phonological Processing in Pre-Readers with and without Familial Risk of Dyslexia

Maaike Smit[1], Ana B. Carbajal Chavez[1], Marlies Gillis[1, 2, 4], Maaike Vandermosten[1], Pol Ghesquière[4], Jan Wouters[1]

[1]Experimental Oto-, Rhino-, Laryngology (ExpORL), Department of Neurosciences, KU Leuven; [2]Laboratory of Functional Anatomy (LAF), Université libre de Bruxelles; [3]Laboratoire de Neuroanatomie et Neuroimagerie translationnelles, UNI – ULB Neuroscience Institute, Université libre de Bruxelles; [4]Parenting and Special Education Research Unit, KU Leuven

Developmental dyslexia is characterised by severe and persistent literacy deficits despite adequate instruction and intelligence (Peterson & Pennington, 2015). Its phonological origins and developmental trajectory, particularly before reading acquisition, remain insufficiently understood. Identifying early neural markers of phonological processing is key to improving diagnostics and enabling timely intervention.

According to the Double Deficit Hypothesis (Vukovic & Siegel, 2006; Wolf & Bowers, 1999) people with dyslexia show impairments in phonological processing that extend beyond low-level auditory difficulties to higher-order operations involving linguistic context and lexical access.

In this study, 3-year-old (N = 46) and 5-year-old (N = 32) children with and without a genetic risk of dyslexia listened to naturalistic spoken stories while EEG was recorded. We examined phonological encoding by modelling neural responses to two features central to predictive phonological processing: phoneme surprisal and phoneme entropy. Surprisal reflects how unexpected a phoneme is given its context, while entropy quantifies uncertainty in phoneme prediction (Shannon, 1948), indexing competition among phonological candidates (Gillis et al., 2021). Encoding models were applied to derive temporal response functions (TRFs), allowing us to assess both the latency and the strength of neural responses (Gillis et al., 2022). Analyses are ongoing and results will be presented at the conference.

References

Gillis, M., Van Canneyt, J., Francart, T., & Vanthornhout, J. (2022). Neural tracking as a diagnostic tool to assess the auditory pathway. *Hearing Research*, *426*, 108607. https://doi.org/10.1016/j.heares.2022.108607

Gillis, M., Vanthornhout, J., Simon, J. Z., Francart, T., & Brodbeck, C. (2021). Neural Markers of Speech Comprehension: Measuring EEG Tracking of Linguistic Speech Representations, Controlling the Speech Acoustics. *Journal of Neuroscience*, *41*(50), 10316–10329. https://doi.org/10.1523/JNEUROSCI.0812-21.2021

Peterson, R. L., & Pennington, B. F. (2015). Developmental Dyslexia. *Annual Review of Clinical Psychology*, *11*(1), 283–307. https://doi.org/10.1146/annurev-clinpsy-032814-112842

Shannon, C. E. (1948). A mathematical theory of communication. In *The Bell System Technical Journal* (Issue 3).

Vukovic, R. K., & Siegel, L. S. (2006). The Double-Deficit Hypothesis: A Comprehensive Analysis of the Evidence. *Journal of Learning Disabilities*, *39*(1), 25–47. https://doi.org/10.1177/00222194060390010401

Wolf, M., & Bowers, P. G. (1999). The Double-Deficit Hypothesis for the Developmental Dyslexias. *Journal of Educational Psychology*, *91*(3), 415–438. https://doi.org/10.1037/0022-0663.91.3.415

Development of Phonetic Distinctiveness in infants and (pre)readers at risk for dyslexia

Klara Spooren[1], Elise Lefèvre[1], Jolijn Vanderauwera[2], Stéphane Dupont[3], Jan Wouters[1], Pol Ghesquière[4] and Maaike Vandermosten[1]

[1]Experimental Otorhinolaryngology (ExpORL), Department of Neuroscience, KU Leuven, Leuven, Belgium
[2]Psychological Sciences Research Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
[3]Information, Signal and Artificial Intelligence Lab, University of Mons, Mons, Belgium
[4] Parenting and Special Education Research Unit, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

Purpose. Typically developing infants refine their speech abilities through exposure to their native language, often referred to as perceptual attunement (e.g., Maurer & Werker, 2014). It is hypothesized that infants at risk for dyslexia exhibit reduced attunement (Serniclaes et al., 2004), potentially resulting in less-specified phonetic representations which impacts later phoneme-to-grapheme mapping skills while reading (Boets et al., 2013; Vandermosten et al., 2020).

Methods. This study focuses on a naturalistic setting, conducting day-long audio recordings in the home environment of the child. A self-developed acoustic algorithm will later on analyze all speech sounds. This approach will lead to a more direct and easy-to-implement measure of how phonetic units gradually move towards more narrow phonemic prototypes, thereby reflecting their underlying representation. In a first study, we aim to longitudinally capture the development of phonetic distinctiveness in typically developing infants from 3 - 18 months old (N = 50), and compare that with the phonetic development of infants with a familial risk for dyslexia (N = 50). In a second study, we measure phonetic distinctiveness in pre-reading 5-year-old children with (N = 50) and without (N = 50) a familial risk for dyslexia, as well as preterm born children (N = 50). We will compare phonetic distinctiveness at age 5 between groups, and link the phonetic measure to their reading abilities at age 7.

Significance. We aim to develop an automated, ecologically valid measure of phonetic distinctiveness in early childhood. It will offer critical insights into early language acquisition and dyslexia risk. The findings may inform early diagnosis, leading to timely intervention. Moreover, our shift to naturalistic setting can give insight in potential novel research paradigms.

References

Boets, B., op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., Bulthé, J., Sunaert, S., Wouters, J., & Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science*, *342*(6163), 1251–1254. https://doi.org/10.1126/science.1244333

Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, *56*, 154–178. https://doi.org/10.1002/dev.21177

Serniclaes, W., van Heghe, S., Mousty, P., Carré, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology*, *87*, 336–361. https://doi.org/10.1016/j.jecp.2004.02.001

Vandermosten, M., Correia, J., Vanderauwera, J., Wouters, J., Ghesquière, P., & Bonte, M. (2020). Brain activity patterns of phonemic representations are atypical in beginning readers with family risk for dyslexia. *Developmental Science*, *23*(1). https://doi.org/10.1111/desc.12857

References

Baranowska, K., & Kaźmierski, K. (2020). Polish word-final nasal vowels. *Sociolinguistic Studies, 14*(1-2), 135–162. https://doi.org/10.1558/sols.37918

Dunaj, B. (2006). Zasady poprawnej wymowy polskiej [Rules for correct Polish pronunciation]. *Język polski, 86*(3), 161–172.

Grondelaers, S., van Hout, R., & Steegs, M. (2010). Evaluating Regional Accent Variation in Standard Dutch. *Journal of Language and Social Psychology, 29*(1), 101–116. https://doi.org/10.1177/0261927X09351681

Gussmann, E. (2007). *The phonology of Polish.* Oxford University Press.

Johnson, J. (1984). Variations in Polish nasal /ę/: A contribution to the development of contrastive sociolinguistic methodology. *Papers and Studies in Contrastive Linguistics, 18*, 31–41.

Levon, E., & Fox, S. (2014). Social Salience and the Sociolinguistic Monitor: A Case Study of ING and TH-fronting in Britain. *Journal of English Linguistics, 42*(3), 185-217. https://doi.org/10.1177/0075424214531487

Sussex, R., & Cubberley, P. (2006). Phonology. In T*he Slavic Languages* (pp. 110–191). Cambridge University Press. https://doi.org/10.1017/CBO9780511486807.006

Tamminga, M. (2017). Matched guise effects can be robust to speech style. *The Journal of the Acoustical Society of America, 142*(1), EL18–EL23. https://doi.org/10.1121/1.4990399

The Role of Phonetic Entrainment in Second Language Vowel Learning

Jing Tang, Laura Smorenburg, Hugo Quené, Aoju Chen
Utrecht University

In L2 communication, speakers tend to adapt their phonemes to that of the interlocuter—phonetic entrainment. L2 speakers' phoneme proficiency shows improvement after exposure to a native speech, but the role entrainment plays in facilitating L2 phoneme learning is unclear. This study addresses this issue by examining the relation between an L2 speaker's entrainment tendency on vowels and their speaking proficiency of these vowels, considering L1-L2 vowel distance.

One native speaker of American English and 30 Mandarin-speaking learners of English, all females, participated as the model talker and the L2 speakers. 36 CVC nonwords, consisting of vowels /i, ɪ, æ/ with increasing L1-L2 distance, were tested through a semi-interactive game. F1 and F2 of the vowels were examined. Linear mixed-effects modelling was used to model L2 speakers' entrainment degree, which was then correlated with their vowel proficiency.

Results on F1 showed that all L2 speakers entrained to the native speaker, confirming the prevalence and automaticity of phonetic entrainment; although not significant, the trend of their entrainment degree varying with vowels corroborates the tenet by SLM-r: the larger the L1-L2 difference the easier it will be perceived and thus learned. Yet results on F2 display much variation. We also found no significant correlation between L2 speaker's entrainment degree on vowels and their speaking proficiency of these vowels, not supporting our hypothesis of entrainment tendency as a predictor for the amount of speech learning from interaction, nor as a speech learning ability constrained by existing speaking proficiency.

Linking variability in pitch accent production and perception

Constantijn L. van der Burght[1,2], Ture Berg[2]
[1]Leiden University Centre for Linguistics, Leiden University, Leiden, Netherlands, [2]Max Planck institute for Psycholinguistics, Nijmegen, Netherlands,

Prosody can mark sentence elements occupying parallel roles. In “Mary kissed John, not Peter”, a contrastive accent on *Mary* or *John* cues the implied syntactic role of *Peter*. There is known to be between-listener variability in the perception and interpretation of prosodic phenomena such as contrastive accents. Similarly, there is considerable between-speaker variability in the realisation of prosodic cues. We asked if variability in prosody production and perception are linked. 40 female native speakers of Dutch participated in a production and perception experiment. The experimental sentences (in Dutch) were of the type “The police officer arrested the thief, not the inspector/murderer”. In the production session, participants performed a picture description task in which a lead-in sentence encouraged the realisation of a contrastive accent on the subject or object of the main clause. In the perception session, two weeks later, participants listened to the experimental sentences, recorded by a separate speaker. From these sentences the ellipsis clause was omitted. Participants indicated the focus structure they perceived by selecting (via button press) the semantically appropriate sentence-final noun (presented visually). The perception experiment confirmed individual differences previously reported, with some listeners exhibiting good sensitivity to prosodic information and others relying on a structural bias. The production data were analysed using contour clustering. Again, between-speaker differences were found, with varying degrees to which distinct focus conditions could be categorised based on each speaker’s intonation contours. A forthcoming analysis will test the production-perception link: Can the perceptive ability to discriminate between prosodic categories predict the degree to which two distinct prosodic categories are produced?

Plosive bursts in Seoul Korean

Michaela Watkins[1]

[1]University of Amsterdam

The laryngeal contrast in Korean has been investigated in both perception and production for multiple cues (e.g. Bang et al. 2018, Kang et al. 2022), but rarely bursts**.** In phonology, it has been suggested that fortis and aspirated are marked for laryngeal activity (constricted and spread glottis respectively), whereas lenis remains unmarked (Lombardi 1994). I hypothesize that this relates to the phonetics: fortis and aspirated plosives show consistent and strong bursts, whereas lenis remains inconsistent (i.e. unmarked).

Tokens were taken from a corpus (Yun et al. 2015) and burst duration (ms), burst energy (integrated over time, resulting in ms), and after-burst duration up to vowel onset (ms) was measured. Aspirated stops consistently had a burst, and lenis occasionally lost bursts phrase-internally. Neither of these outcomes are unexpected as lenis exhibits carry-over (gradient) voicing internally (Davidson 2016, Han 2000), therefore losing a burst is possible. However, surprisingly fortis stops lost their burst the most across all positions. After-burst duration was significant for fortis stops at 35.3ms shorter on average compared to the other two stop types ($p = 2e\text{-}16$, c.i. -39.37ms … -31.29ms) and no significant difference was observed between aspirated and lenis, matching prior research. In terms of the two burst measures (duration and normalized energy), both burst models struggled to converge even when random slopes and intercepts were set to (1|Speaker), which is not ideal given expected differences across and within speakers in their (baseline) burst productions, although no significant effect was found in any of the models. The model convergence problems are possibly due to a) insufficient quantity of data to find an effect and/or b) the burst is too varied within and across speakers. Preliminary results with more controlled speech demonstrate burst again being non-significant but distinct patterns of place of articulation and burst presence/absence, warranting further assessment of bursts across place of articulation.

This analysis demonstrates the complexity of phonetic detail within and across speakers, and the difficulty in fitting such detail into concise phonological categories. In Korean, place of articulation place of articulation, rather than the specified laryngeal category, may yield more insights into the nature of the burst rather than comparing across the three larger categories, but this requires a separate study. It is therefore not recommended to map markedness onto phonetic detail (or vice versa).

References

Bang, H-Y., Sonderegger, M., Kang, Y., Clayards, M., Yoon, T-J., (2017), The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics, 66*, 120-144.

Kang, Y., Schertz, J., Han, S. (2022). *The Phonology and Phonetics of Korean stop laryngeal contrasts.*

Lombardi, L. (1994), *Laryngeal Features and Laryngeal Neutralization.* New York & London: Garland.

Yun et al. (2015). The Korean Corpus of Spontaneous Speech, *Phonetics and Speech Sciences*, *7*(2), 103-109.

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics 54*, 35-50.

Han, J.I. (2000). Intervocalic stop voicing revisited. *Speech Sciences*, *7*, 203-216.