NVFW
Nederlandse Vereniging voor Fonetische Wetenschappen

Dag van de Fonetiek

ABSTRACTS

15 december 2023

Utrecht
Drift 21 – Sweelinckzaal

Programma Dag van de Fonetiek, 2023		
09:30	10:00	inloop
10:00	11:00	Hans Rutger Bosker *Handy prosody: how hands can help you hear*
11:00	11:20	pauze
11:20	11:45	Johanneke Caspers*, Wie is er aan de beurt? De rol van spraakmelodie bij beurtwisselingen in het Nederlands.*
11:45	12:10	Cesko C. Voeten*, Fonetische natuurlijkheid als drijfveer van verandering?*
12:10	12:20	ALV
		lunchpauze
14:10	14:35	Stella Gryllia & Amalia Arvaniti, *Compositionality in intonation: Are tunes composed of independent elements?*
14:35	15:00	Michaela Watkins, Silke Hamann & Paul Boersma*, How to Measure Real Pitch Jumps in Praat: An Example Study from Korean*
15:00	15:25	Ahdiyeh Alipour & Tom Lentz, *The uncanny valley effect for prosody*
15:25	15:45	pauze
15:45	16:10	Laura Smorenburg & Willemijn Heeren, *Idiosyncratic and linguistic information in /s/ in telephone speech*
16:10	16:35	Elizaveta Ivanova, Anastasia Sukmanova, Marina Norkina & Alisa Kosikova, *Phonological challenges in children With Developmental Language Disorder (DLD) across Slavic languages*
		borrel

KEYNOTE

Handy prosody: how hands can help you hear

Hans Rutger Bosker

SPEAC lab, Donders Institute, Radboud University

Speech conveys both segmental information about vowels and consonants as well as suprasegmental information about for instance intonation, rhythm, and lexical stress; also known as the prosody of speech. However, in face-to-face conversations, we do not only exchange sounds; we also move, nod, and gesture to the rhythm of our speech. In this keynote, I will demonstrate how the timing of hand gestures contributes to audiovisual prosody perception, with a focus on lexical stress. For instance, evidence for a ‘manual McGurk effect’ showcases how even relatively simple flicks of the hands can guide whether you hear "PLAto" or "plaTEAU". Moreover, human listeners are shown to actively weigh various multisensory cues to prosody depending on the listening conditions at hand. Thus, these findings emphasize that prosody is a multimodal linguistic phenomenon, with the voice, lips, and even hands conveying prosody in concert.

Wie is er aan de beurt? De rol van spraakmelodie bij beurtwisselingen in het Nederlands

Johanneke Caspers

Leiden University Centre for Linguistics

In natuurlijke conversatie wisselen sprekers elkaar soepel en snel af. Uit onderzoek blijkt dat er tussen opeenvolgende sprekers gemiddeld 0-200 ms stilte zit, terwijl het plannen van een beurt ca. 600-1500 ms kost. Dit zou kunnen betekenen dat de rol van uitingsfinale zinsmelodie in het beurtwisselingssysteem beperkt is tot het markeren van de plaats waar de beurt eventueel kan worden overgenomen, maar dat het type finale grenstoon - stijgend (H%) of laag (L%) - er weinig toe doet, omdat de planning van een beurt eerder moet beginnen. Er is enige evidentie dat het ontbreken van een hoge of lage grenstoon (in plaats daarvan eindigt de uiting in een vlakke middelhoge toon, %) werkt als beurthoud-cue.

Meermalen is aangetoond dat er zoiets bestaat als vraagintonatie, en dat een finale stijging (H%) een declaratieve uiting kan markeren als vraag. Dit zou betekenen dat er mogelijk toch een rol is voor H% als onafhankelijk signaal om de beurt over te dragen, en dat de gerapporteerde zeer korte pauzes tussen sprekers het gevolg zijn van een te oppervlakkige analyse van de beurten in de gebruikte datasets. In het NWO-project ‘Who’s next?’ wordt onderzocht wat de rol van de drie finale grenstonen (H% stijgend, L% laag, % middelhoog) in het beurtwisselingssysteem van het Nederlands is, te beginnen met een gedetailleerde annotatie van een corpus gesproken dialoog (wat betreft intonatie, beurttype, lexicosyntactische en pragmatische structuur), gevolgd door experimenten met gemanipuleerde stukken dialoog en eye-tracking.

Fonetische natuurlijkheid als drijfveer van verandering?

Cesko C. Voeten[1234]

[1]*University of Pennsylvania, Department of Linguistics*
[2]*University of Pennsylvania, Department of Biology*
[3]*Universiteit van Amsterdam, afdeling Literatuur- en Taalwetenschap*
[4]*Fryske Akademy*

Een belangrijk onderdeel van het 'constraints problem' van Weinreich, Labov, & Herzog (1968) is de vraag welke taalveranderingen mogelijk en onmogelijk zijn. In de fonetiek worden de mogelijkheden beperkt, volgens Garrett & Johnson (2013), door fonetische *biases*, die voortkomen uit de eigenschappen van het productie- en perceptieapparaat. In hun visie leiden die *biases* tot taalverandering, indien ze (waarom dan ook) sociale relevantie verkrijgen. Hoe werken die *biases*? *Drijven* die verandering – in termen uit de biologie, veroorzaken ze *selectiedruk* – of zijn ze slechts een beperking op de weg die een taalverandering van nature (uit zgn. *stochastische verschuiving*) zou moeten vinden? Ik onderzoek die vraag in een specifieke casus uit het Engels van Philadelphia.

Één van de kenmerken van het Philadelphia-accent is *American Raising*: een centralisering van [aɪ] tot [əɪ] voor stemloze obstruenten. Voor deze verandering zijn twee verklaringen geponeerd (Davis & Berkson 2021): *pre-fortis clipping* (vocaalverkorting die maakt dat de [a] niet helemaal bereikt wordt) en *offglide peripheralization* (de [ɪ] wordt [i] en de [a] coarticuleert mee). Ik onderzoek American Raising in het Philadelphia Neighborhood Corpus (408 sprekers geboren in 1880-1994) d.m.v. een statistisch model oorspronkelijk uit de populatiegenetica (zie Nourmohammad et al 2017) dat onderscheid maakt tussen selectie en stochastische verschuiving. Resultaten tonen sterke selectiedruk in de genormaliseerde F1/F2-ruimte, die bovendien sterker is bij vrouwen dan bij mannen. Ik vind echter geen evidentie voor bijkomstige selectie op kortere vocaalduren in deze allofooncontext, noch voor sterkere selectie in offglides dan in onglides. Het gevonden verschil tussen mannen en vrouwen leidt tot de conclusie dat *sociale* factoren een drijfveer (kunnen) zijn van taalverandering. Over de rol van *fonetische* factoren is, gezien de nulresultaten voor die factoren, echter geen conclusie mogelijk.

Referenties

Davis, S., & Berkson, K. (2021). American Raising: An Introduction. *Publication of the American Dialect Society*, 106(1), 1-12.

Garrett, A., & Johnson, K. (2013). Phonetic bias in sound change. In: Yu, A. C. L. (Ed.). *Origins of sound change: Approaches to phonologization* (pp. 51-97). Oxford University Press.

Nourmohammad, A., Rambeau, J., Held, T., Kovacova, V., Berg, J., & Lässig, M. (2017). Adaptive evolution of gene expression in Drosophila. *Cell Reports*, 20(6), 1385-1395.

Weinreich, U., Labov, W., & Herzog, M. (1968). *Empirical foundations for a theory of language change*. University of Texas Press.

Compositionality in intonation: Are tunes composed of independent elements?

Stella Gryllia & Amalia Arvaniti

Radboud University

A recurrent issue in the study of intonation relates to whether contours should be treated as gestalts [1, 2] or composites of independent elements [3, 4]. We contribute to this debate by examining a corpus of 2135 Greek wh-questions, elicited from 18 speakers using a discourse completion task (DCT). DCTs involved two scenarios: Scenario A presented a situation ending with an information-seeking question, while Scenario B presented a situation in which the wh-question was used as an implicit statement. The expected tune for Scenario A is autosegmentally analysed as a L*+H pitch accent on the utterance-initial wh-word, followed by a L- phrase accent and a H% boundary tone; for Scenario B the expected tune is analysed as L+H* L- L% [5]. We applied functional principal component analysis (FPCA), a data-driven method that breaks down curves into components capturing independent modes of curve variation. FPCA was followed by LMEMs on the principal component coefficients. The results show that the pitch movement associated with each of the posited tonal elements is captured by a different PC: PC1 captures the shape of the fall (as a consequence of peak height and alignment), PC2 captures the extent of the initial rise and subsequent peak alignment of the pitch accent (the difference between L*+H and L+H*), and PC4 the difference between a final rise (H%) and low, flat pitch (L%). Given that each PC presents an independent mode of variation, we can conclude that tunes are composites of independent elements. These results provide prima facie evidence for tune compositionality.

References

[1] Hirst, D., & Di Cristo, A.1998. A survey of intonation systems. In D. Hirst & A. Di Cristo (Eds.), *Intonation Systems a Survey of Twenty Languages*, 1-44.

[2] Xu, Y. 2005. Speech melody as articulatorily implemented communicative functions. *Speech Communication*, 46(3-4), 220-251.

[3] Pierrehumbert, J. & Hirschberg, J. B. 1990. The meaning of intonational contours in the interpretation of discourse. In P.R. Cohen, J. Morgan & M.E. Pollack (Eds.), *Intentions in Communication*, 271-311.

[4] Ladd, D. R. 2008. *Intonational Phonology*. Cambridge University Press.

[5] Baltazani, M., Gryllia, S., & Arvaniti, A. 2020. The Intonation and Pragmatics of Greek wh- Questions. *Language and Speech*, 63(1), 56–94. https://doi.org/10.1177/0023830918823236

How to Measure Real Pitch Jumps in Praat: An Example Study from Korean

Michaela Watkins, Silke Hamann, Paul Boersma

University of Amsterdam

Pitch doubling and halving occur when the F0 pitch contour does not follow a consistent curve but rather 'jumps' in pitch are observed. Often these pitch jumps are deemed as pitch tracking 'errors', due to the initial seemingly inconsistent nature of the jumps. This phenomenon often co-occurs with creak, thus leading to further complications in automatic measurements in pitch-tracking software.

However, a visual investigation into Seoul Korean instead suggests that doubling and halving effects in 'fortis' plosives are not 'errors', but rather constitute an inherent characteristic of this kind of stops. Previous research often tried to 'correct' for such pitch tracking jumps by ignoring tokens below a certain F0 threshold, and as a result likely did not provide a complete overview. The 'true' pitch contour of Korean fortis stops can be established through adjustment of specific settings in Praat, upon which it becomes clear that Korean fortis stops often come with creaky voice and *true* pitch halving during the initial part of the following vowel.

We conclude that (1) pitch jumps should not automatically be assumed as software or measurement errors, (2) Praat is able to handle pitch jumps and visualize them systematically, and (3) selectively leaving out tokens to avoid pitch jumps is unnecessary and can even misrepresent the phonetic data.

The uncanny valley effect for prosody

Ahdiyeh Alipour & Tom Lentz

Tilburg University

The incorporation of human-like traits into intelligent verbal agents (IVAs) can strengthen their appeal as users are inclined to anthropomorphize technology. Nevertheless, non-human objects with close-to-human characteristics can evoke the uncanny valley effect (UVE), a feeling of eeriness. This study investigated whether very human-like prosody used by IVAs can evoke a UVE, looking at perceived eeriness, robomorphism, anthropomorphism and trust in/of IVAs. Two experiments were conducted in which participants (N = 88) listened to a virtual museum guide. The guide's voice was a human voice, the pitch of which was manipulated to sound more robotic, by reducing the variation around the declination over the utterance (first experiment) or around the average (second experiment). The reduction was 0% (human), 33%, 66% or 100% (most robotic); participants listened to each level once. There was a linear relationship between degree of prosody manipulation and participants' perceptions on the variables mentioned above, in the expected directions (experiment 1: $0.31 < |b| < 0.57$, $p < 0.001$; experiment 2: $0.27 < |b| < 0.74$, $p < 0.001$), but no evidence of a UVE. However, a few individuals did show a clear and consistent UVE, the significance of which cannot be estimated as each level was presented only once. A follow-up study using a finer-grained difference and more measurements per participants is underway; results will be discussed. This study's insights can contribute to better design and implementation of IVAs by tackling potentially unfavorable emotional and behavioral reactions to human-like voices.

Keywords: Prosody, Intelligent Virtual Assistants, Prosody, Human-Robot Interaction, Anthropomorphism, Uncanny Valley Effect

Idiosyncratic and linguistic information in /s/ in telephone speech

Laura Smorenburg[1] & Willemijn Heeren

Leiden University Centre for Linguistics
Leiden University

Telephone speech can be characterised by a limited bandwidth, varying bitrates (in the case of mobile signals), and different speech behaviour. For vowels, formants that are situated near lower signal cut offs are shown to be affected (e.g., Künzel 2001). Investigating the effects of telephone filters is particularly relevant for forensic speaker comparisons, where wiretapped telephone conversations are commonly analyzed. However, not much is known about the effect of telephone filters on different consonants.

In the current work, we investigated fricative /s/, which has been described to contain relatively high amounts of idiosyncratic information (Kavanagh, 2012; Van den Heuvel, 1996), even in narrowband telephone recordings where most of its high-frequency spectral characteristics are compromised (Smorenburg & Heeren, 2021). We annotated >100 /s/ tokens for 60 speakers of British English from the West Yorkshire Regional English Database (WYRED: Gold, 2020). These participants performed a forensic speech task in which they conversed over the telephone with an 'accomplice' and were recorded both over a microphone placed in front of them and wiretapped over the landline.

Results show that linguistic information (effects of phonetic context and syllabic position) is compromised in telephone recordings compared to microphone recordings, but that some linguistic environments still show more between-speaker variability. Specifically, when /s/ is followed by labial sounds, speaker-classification accuracy was higher. This indicates that coarticulation contains idiosyncrasies that can be used in forensic speaker comparisons.

References

Gold, Erica (2020). *WYRED - West Yorkshire Regional English Database 2016-2019.* [Data Collection]. Colchester, Essex: UK Data Service. 10.5255/UKDA-SN-854354

Künzel, H. J. (2001) Beware of the 'telephone effect': the influence of telephone transmission on the measurement of formant frequencies. Forensic Linguistics 8: 80–99. http://dx.doi.org/10.1558/sll.2001.8.1.80

Smorenburg, L. and Heeren, W. (2021) The distribution of speaker information in Dutch fricatives /s/ and /x/ from telephone dialogues. *Journal of the Acoustical Society of America 147*(2): 979–989. https://doi.org/10.1121/10.0005845

[1] *Now at UiL OTS, Utrecht University*

Phonological challenges in Children With Developmental Language Disorder (DLD) Across Slavic Languages

Elizaveta Ivanova, *University of Surrey, UK*
Anastasia Sukmanova, *HSE University, National Research University, Russia*
Marina Norkina, *Sirius University of Science and Technology, Russia*
Alisa Kosikova, *Saint Petersburg State University, Russia*

Developmental Language Disorder (DLD) causes significant challenges in expressive and/or receptive language and affects around 7% of the population (Bishop, 2017). While many studies have investigated various aspects of DLD, most of them focused on English-speaking regions, leaving many other regions underrepresented in the DLD framework. Meanwhile, the inclusion of diverse language groups is crucial for understanding the overall homogeneity of DLD, discussing standardized diagnostic tools and potentially developing better interventions. Here we present our systematic review of DLD in the Slavic language group that included all original studies in English and Russian databases (PROSPERO ID - CRD42021235107). While we have been looking at several linguistic processing and comprehension skills (phonology, morphology, syntax, semantics and pragmatics), for this conference, we highlight the results specific to phonological processing.

Our final sample includes 8 phonological studies of Russian (4), Slovak (1), Serbian (2), and Czech (1) languages, covering a range of children from 3 to 14 years old. Across these studies, common issues include misarticulated/omitted sounds, diphthong simplification, and deficiencies in phonological memory, strongly associated with rhyme, phonemic awareness, and grapheme recognition. Overall, deficits in phonology across Slavic languages tend to co-occur together with other difficulties (e.g. grammar and semantics), may result from deficits in motor development or correlate with other processes (such as reading). Therefore, the phonological deficits in Slavic children with DLD should be viewed as one of the facets of the heterogeneous multifaceted DLD.