Nederlandse Vereniging voor Fonetische Wetenschappen

Dag van de Fonetiek

ABSTRACTS

16 december 2022

Utrecht
Drift 21 – Sweelinckzaal

PROGRAMMA

inloop	9.30-10.00		
keynote	10:00-10:45	Fonetiek in de klas	Marc van Oostendorp
pauze	10:45-11:00	koffie en thee	
1	11:00-11:20	Het Asta-project: automatische spraakherkenners voor Nederlandse dialecten	Martijn Bentum, Eric Sanders, Antal van den Bosch & Henk van den Heuvel
2	11:20-11:40	What are the mental representations of speech segments during speech preparation, as revealed by self-monitoring for speech errors?	Sieb Nooteboom & Hugo Quené
3	11:40-12:00	An acoustic analysis of West Frisian monophthongs	Martijn Kingma
4	12:00-12:20	An operationalization of causal factors in vowel shifts	Cesko C. Voeten, Meredith Tamminga & Joshua B. Plotkin
ALV	12:20-12:25		
lunchpauze	12:25-13:50		
5	13:50-14:10	Recalibration of lexical stress perception can be driven by visual beat gestures	Ronny Bujok, David Peeters, Antje Meyer & Hans Rutger Bosker
6	14:10-14.30	Can Rapid Prosody Transcription be replicated?	Riccardo Orrico, Stella Gryllia, Jiseung Kim & Amalia Arvaniti
7	14:30-14:50	Both Contextual and Talker-Bound F0 Information Affect Voiceless Fricative Perception	Orhun Uluşahin, Hans Rutger Bosker, James M. McQueen & Antje S. Meyer
8	14:50-15:10	Ghost segments in the Flemish Tussentaal	Mishko Bozhinoski

poster pitches	15:10-15:15	1-minute poster pitches	
posters / pauze	15:15-16:00	koffie en thee / posters	
9	poster 1	The perception of L2 French nasal vowels by L1 Dutch listeners	Hanna De Haes, Ellen Simon & Peter Lauwers
10	poster 2	Translating a Korean Poem into English – A Case Study on the Connections between Emotions and Phonetics	Mika Aya Soehnlein
11	poster 3	The influence of musical abilities on the processing of contrastive focus prosody in an L2: An eye-tracking study	Nelleke Jansen, Hanneke Loerts, Eleanor Harding, Deniz Başkent & Wander Lowie
12	poster 4	The speaker in speech – linguistic context in forensic speaker comparisons	Willemijn Heeren, Meike de Boer & Laura Smorenburg
13	16:00-16:20	F0 dynamics associated with prominence realisation in children with hearing impairment	Jérémy Genette & Jo Verhoeven
14	16:20-16:40	Phoneme Categorization after Speaking with a Bite Block	Xinyu Zhang & Esther Janse
15	16:40-17:00	Exploring articulation rate entrainment: Interactions over Zoom between typical and atypical speakers	Lotte Eijk & Sophie Meekings
borrel	17:00-18:00		

KEYNOTE

Fonetiek in de klas

Marc van Oostendorp

Radboud Universiteit Nijmegen & Meertens Instituut

De fonetiek is een vak vol interessante inzichten, spectaculaire resultaten en interessante manieren om ook als jongere zélf aan de slag te gaan. Waarom wordt daar dan zo weinig gebruik van gemaakt? Er zijn allerlei redenen om de wetenschap bij jongeren te brengen – omdat je wil dat de continuïteit gegarandeerd wordt doordat ook nieuwe generaties enthousiast raken, omdat je denkt dat het belangrijk is dat mensen op een verantwoordelijke manier met nieuwe technologie kunnen omgaan, of omdat je simpelweg de schoonheid van het vak wil delen. In deze presentatie pleit ik niet alleen voor meer fonetiek in het onderwijs, maar doe ook tips aan de hand hoe we dat zouden kunnen bereiken.

Het Asta-project: automatische spraakherkenners voor Nederlandse dialecten

Martijn Bentum, *CLS/CLST, Radboud Universiteit*
Eric Sanders, *CLS/CLST, Radboud Universiteit*
Antal van den Bosch, *Universiteit Utrecht*
Henk van den Heuvel, *CLS/CLST, Radboud Universiteit*

Het Meertens instituut heeft in de tweede helft van de 20ste eeuw in heel Nederland opnames gemaakt van verschillende dialecten en ongeveer driehonderd uur is handmatig getranscribeerd. Dit materiaal lijkt de ideale basis voor het ontwikkelen van *dialect-specifieke spraakherkenners*, maar er zijn ook aanzienlijke uitdagingen bij het verwerken van deze data.

De handmatige transcripties zijn gekoppeld aan de spraakopnames via een metadata-bestand. Helaas zijn de transcripties niet opgelijnd met de audio waardoor het onduidelijk is wanneer er wat gezegd wordt. Verder zijn de transcripties uitgeschreven in semi-conventionele spelling die is aangepast om de uitspraak in het dialect weer te geven. Hierdoor zijn de handmatige transcripties helaas niet altijd consistent en zijn ze moeilijker te koppelen aan automatische transcripties voor oplijning tussen audio en transcriptie.

Om een eerste oplijningsbenadering te maken hebben we gebruik gemaakt van automatische spraakherkenning door een Nederlands Wav2vec2 model in combinatie met het Needleman-Wunch algoritme. Dit algoritme benadert een optimale oplijning tussen twee sequenties in dit geval tussen de handmatige en automatische transcripties. Deze oplijning wordt gecontroleerd met een hiervoor ontwikkeld webgebaseerde annotatietool. Met de resultaten van deze annotatie kan de spraakherkenner verbeterd worden voor specifieke dialecten.

Tot nu toe kan geconcludeerd worden dat het oplijnen van de oorspronkelijke handmatige dialecttranscripties met behulp van een standaard Nederlands Wav2vec2 model en het Needleman-Wunsch algoritme goed werkt, maar dat er handmatige filtering van de data nodig is om het materiaal geschikt te maken voor het trainen van dialect-specifieke spraakherkenners.

ASTA is een subproject van Werkpakket 3 (“Linguistics”) van het CLARIAH-PLUS Grootschalige Wetenschappelijke Infrastructuurproject, en wordt gefinancierd door NWO (projectnummer 184.034.023).

What are the mental representations of speech segments during speech preparation, as revealed by self-monitoring for speech errors?

Sieb Nooteboom & Hugo Quené

Utrecht institute of Linguistics OTS

Segmental speech errors can be detected by self-monitoring both before and after speech is initiated. This allows us to answer four questions about speech preparation: (1) What is the delay between self-monitoring internal and external speech? (2) What are the representations of speech sounds on which self-monitoring focuses during “early” and “late” error detection? (3) Why are some speech errors detected by self-monitoring before and others after speech initiation, and others again not at all?(4) What is the role of cognitive control in self-monitoring for speech errors? For answering those questions we use responses from 6 of our earlier SLIP experiments. We find that (1) the delay between “early” and “late” error detection is roughly 4 segments or nearly 500 ms. (2) Speech sounds are represented very differently before and after speech initiation. This is not predicted from articulatory phonology but supports the proposal that in internal speech sounds are represented as targets in auditory perceptual space. (3) Phonetic contrast between segments is a major factor determining whether speech errors are detected internally, externally or not at all. (4) Degree of conflict between competing items during speech preparation controls frequency of speech errors, but not necessarily of error detection.

An acoustic analysis of West Frisian monophthongs

Martijn Kingma

Universiteit van Amsterdam & Fryske Akademy

Previous literature on Modern West Frisian reports a fairly large vowel inventory, including 18 monophthongs consisting of nine vowel pairs distinguished by length. Yet, this extensive vowel system has hardly been subject to phonetic investigation. Here, I report some preliminary findings of an ongoing investigation into some acoustic properties (F1, F2 and duration) of the 9 short and 9 long monophthongs of this minority language. Based on corpus data from 2017-2019 of native speakers from the former municipality Boarnsterhim, realisations of 6 older males are compared with 6 younger female speakers. Vowels were measured preceding [t] or [s] in closed, stressed syllables of content words in spontaneous speech. Besides providing a preliminary acoustic description of Frisian monophthongs in spontaneous speech, some sound changes are looked at in more detail. The first analyses reveal at least one obvious change in progress: younger speakers show the tendency to diphthongise the long high vowels [iː yː uː] to [iˑə yˑə uˑə]. This change could impact the Frisian vowel inventory, as these diphthongs hold phonemic value (see the minimal pair *wiid* [ʋiːt] 'wide' vs. *wiet* [ʋiˑət] 'wet'). Tracking the origin and spread of this merger could provide new insights into how Frisian is changing, possibly under the influence of Dutch.

An operationalization of causal factors in vowel shifts

Cesko C. Voeten[123], Meredith Tamminga[1], Joshua B. Plotkin[2]

1) University of Pennsylvania, Department of Linguistics, Philadelphia, PA, USA
2) University of Pennsylvania, Department of Biology, Philadelphia, PA, USA
3) Fryske Akademy, Leeuwarden, The Netherlands

Ample phonetic literature has tried to determine causality behind sound change, oftentimes in vowel shifts. Proposed causes may be divided into *internal* factors (e.g. push chains maintaining margins of security, drag chains maintaining symmetry; [1]) and *external* ones (gender, social network, etc.; [2]). The usual approach to establishing such effects (e.g. [3]) is to demonstrate that such factors significantly correlate with change progression over time. However, such a *correlation* is insufficient to establish *causation*. Thus, do internal and external factors really *cause* change, or are they simply along for the ride?

We present a method to answering this question based on natural selection in biology. We operationalize causality as a deterministic pressure driving language change ('directional selection'), which contrasts with the cumulative effect of random variation ('stochastic drift'). [4] demonstrated that this distinction makes it possible to disambiguate *caused* changes from historical accidents, based on results from English do-support (Fig. 1). We present ongoing work adapting this model from discrete syntactic features to continuous measures in F1/F2 space. We use this to infer selection pressure within the vowels in the Philadelphia Neighborhood Corpus, a corpus full of vowel changes spanning a course of ~100 years ([5]).

Preliminary results demonstrate that some of the known ongoing changes in Philadelphia, such as pre-fortis [aɪ]>[ʌɪ], are the result of selection pressure, while others are historical accidents. We furthermore present in-progress quantitative analyses of internal and external factors that drive those changes that were subject to selection pressure. We discuss these findings in light of current theories of sound change.

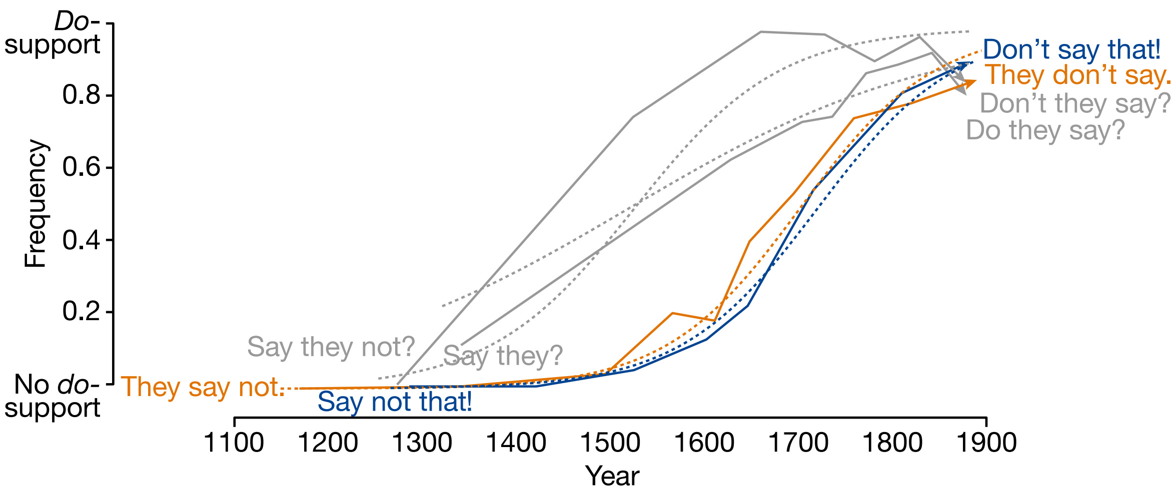

Figure 1: English do-support ([4]). All four changes show significant trends with time, but only two (orange & blue) were driven by selection; the other two (gray) were historical accidents.

References

[1] Labov, W. (1994). *Principles of linguistic change.* Vol. I: internal factors. Blackwell.
[2] Labov, W. (2001). *Principles of linguistic change.* Vol. II: social factors. Blackwell.
[3] Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1(3), 199-244.
[4] Newberry, M. G., Ahern, C. A., Clark, R., & Plotkin, J. B. (2017). Detecting evolutionary forces in language change. *Nature*, 551(7679), 223-226.
[5] Labov, W., Rosenfelder, I., & Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, 89(1), 30-65.

Recalibration of lexical stress perception can be driven by visual beat gestures

Ronny Bujok[1], David Peeters[2], Antje Meyer[1], Hans Rutger Bosker[1,3]

[1]*Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands*
[2]*Department of Communication and Cognition, TiCC Tilburg University, Tilburg, The Netherlands*
[3]*Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands*

Auditory speech is highly variable. Listeners may therefore use the visual modality to disambiguate ambiguous speech sounds. For instance, when repeatedly presented with an ambiguous sound /a?a/ midway between /aba/ and /ada/, paired with a video of a talker producing either /aba/ or /ada/, listeners recalibrate their perception of a later presented auditory /aba - ada/ continuum (Bertelson et al., 2003). Here we tested whether recalibration can also occur for lexical stress perception. In Experiment 1 participants were exposed to an ambiguously stressed token of /ka.nɔn/, perceptually midway between Dutch CAnon [strong-weak (SW); "canon"] and kaNON [weak-strong (WS); "cannon"], disambiguated either by a beat gesture aligned to the first or second syllable (visual). In a later test phase participant categorized an auditory CAnon – kaNON lexical stress continuum. The results revealed that participants' responses in the test phase shifted in the direction of the disambiguating beat gestures they saw in the exposure phase. In Experiment 2 participants were exposed to a different ambiguous word (/vo:r.na:m/) but tested on the same CAnon – kaNON continuum, to test if the effect would also generalize to different words. However, results show that participants were not able to generalize. Ongoing work is investigating whether this generalization is modulated by acoustic distance. Nonetheless, the effect was clearly present across multiple auditory steps in Experiment 1. Therefore, we suggest that beat gestures can recalibrate lexical stress perception and thus have a long-lasting effect on auditory perception.

Can Rapid Prosody Transcription be replicated?

Riccardo Orrico, Stella Gryllia, Jiseung Kim, Amalia Arvaniti

Radboud University

Prominence is studied using Rapid Prosody Transcription (RPT), in which naïve participants hear utterances and mark the words they consider prominent. One such study, Arvaniti et al. (2022; *Speech Prosody*), examined the prominence scores of 281 accents – independently coded as H* or L+H* using phonetic criteria and as contrastive or non-contrastive using pragmatic criteria – provided by Standard Southern British English (SSBE) speakers. Individual participant responses fell into three patterns: responses based on acoustic prominence (favouring L+H* accents), pragmatic meaning (favouring contrastive accents), or both (marking both L+H*s and contrastive accents as prominent). We tested whether RPT results and these response patterns are replicable and explainable by Empathy Quotient (EQ), which enhances attention to meaning leading to higher scores for contrastive accents, and Autism Quotient (AQ) or musicality (measured by mini-PROMS), which enhance attention to phonetic detail leading to higher scores for L+H*s. Sixty two SSBE speakers participated in RPT using the method and materials of Arvaniti et al. (2022). The aggregate results were replicated: contrastive L+H*s were significantly *more* likely and non-contrastive H*s significantly *less* likely to be considered prominent, while non-contrastive L+H*s and contrastive H*s had similarly low scores. Individual participant responses were not affected by AQ. However, participants with high EQ did prioritize pragmatic over phonetic cues (i.e. they favoured contrastive accents independently of shape). Finally, those scoring high in musicality were most sensitive to phonetic differences particularly when they combined with pragmatics (leading to very high scores for contrastive L+H*s and very low scores to non-contrastive H*s).

Both Contextual and Talker-Bound F0 Information Affect Voiceless Fricative Perception

Orhun Uluşahin[1], Hans Rutger Bosker[1,2], James M. McQueen[1,2], Antje S. Meyer[1]

[1] *Max Planck Institute for Psycholinguistics*
[2] *Donders Institute for Brain, Cognition and Behaviour*

Speech perception is sensitive to context. An example of this is the contrastive effect of fundamental frequency (F0) on the perception of voiceless fricatives' spectral center of gravity (CoG) (e.g., Niebuhr, 2017). However, whether knowledge about a talker's characteristic mean F0 can produce similar effects remains unknown. The present study therefore investigated the effects of contextual (Exp 1) and talker-bound (Exp 2) F0 information on the perception of the voiceless fricatives /s/ and /ʃ/. In Experiment 1, in a 2AFC task, native Dutch listeners (N=10) categorized target words as the Dutch words "sok" /sɔk/ or "sjok" /ʃɔk/ embedded in a carrier sentence ("Nu komt het woord...") in 3 intermixed F0 conditions. The fricatives were tokens from a synthetic 8-step fricative continuum from /s/ to /ʃ/. The carrier sentence was pitch shifted ±4 semitones to create High-F0 and Low-F0 context conditions, alongside a Mid-F0 (i.e., non-shifted) control condition. Ambiguous fricatives were perceived as more /s/-like in Low-F0 sentences compared to High-F0 sentences. In Experiment 2, new participants (N=32) first listened to 20 minutes of speech (exposure) from the same talker whose voice had been consistently pitch-shifted up (High-F0 group) or down (Low-F0 group) ±4 semitones. Afterwards, a 5-step subset of the original 8-step fricative continuum was used in a 2AFC task where participants categorized stimuli without carrier sentences as "sok" or "sjok". The continuum was again perceived as more /s/-like for the Low-F0 group compared to the High-F0 group. Together, the findings suggest that listeners use not only the immediate context but also previously established knowledge about talkers' typical F0 to interpret incoming speech sounds.

References

Niebuhr, O. (2017). On the perception of "segmental intonation": F0 context effects on sibilant identification in German. *EURASIP Journal on Audio, Speech, and Music Processing*, *2017*(1), 19. https://doi.org/10.1186/s13636-017-0115-3

Ghost segments in the Flemish Tussentaal

Mishko Bozhinoski
Supervised by Paul Boersma

University of Amsterdam

The Flemish informal standard (or *Tussentaal*) exhibits a feature present in most Low Franconian variants of Flanders: the elision of final coronal plosives in words such as *wat*, *niet* and *met*, whereby they are realised without their final stop: [βɑ], [ni] and [mɛ]. Notably, these words participate in a sandhi process that triggers the devoicing of any fricatives that follow, similar to how those fricatives would devoice if the elided coronal were still present. Thus, phrases such as <wat vlees> and <met zand>, in which the second word normally starts with a voiced fricative, are realised as [βɑ fle:s] and [mɛ sɑnt]. The elision of this final coronal happens in some words (*met, gaat, wat)* and not in others (*wet, tot, uit)*. While previous literature has acknowledged the existence of this ghost segment (such as Camerman, 2007; De Schutter, 1999; Rys, 2020), there had been little to no synchronic or diachronic account for its distribution prior to this study. The present study therefore set off to do the following: (1) document the presence of the feature in the Tussentaal, and (2) attempt to account for its distribution.

Using a Python web-scraping script and the corpus search engine *OpenSonar* (Oostdijk et al., 2018), up to (where possible) ten audio tokens each of the fifty most common Dutch words according to *A Frequency Dictionary of Dutch* (Tiberius & Schoonheim, 2014) that contain final -t or -d were collected from the Corpus Gesproken Nederlands (Eerten, 2007), a spoken corpus of Dutch. The scraped data pool was filtered to include only speakers of Belgian Dutch. The joint results of an algorithmic approach using *Praat* and a manual annotation approach reveal that final coronals are phonetically elided post-vocalically in the present tense inflectional morpheme /-t/ and most function words. All exceptions found were words that have undergone a historical process of word-final schwa deletion, suggesting that final coronal plosive deletion may be older. And indeed, evidence of final coronal plosive deletion was found in texts written in the 13th century, before word-final schwa deletion had taken place.

References

Camerman, Filip. (2007). *Antwerps schrijven: Spelling en grammatica van het 21e-eeuwse Antwerps* (pp. 44-46). de Vries-Brouwers.
De Schutter, G. (1999). Het Antwerps. Een schets van zijn evolutie tussen 1898 en 1998. In J. Kruijsen & N. Van der Sijs (Eds.), *Honderd jaar stadstaal* (pp. 301–315). Contact.
Eerten, L. van. (2007). Corpus Gesproken Nederlands. *Nederlandse taalkunde*, *12*(3), 193-215.
Oostdijk, Nelleke and Reynaert, Martin and Hoste, Véronique and Schuurman, Ineke (2018). OpenSoNaR [Online service]. Available at the Dutch Language Institute: http://hdl.handle.net/10032/tm-a2-c7.
Rys, K. (2020). 1.5.2.6 Deletie van t. *Algemene Nederlandse Spraakkunst*. https://e-ans.ivdnt.org/topics/pid/topic-16007638353447182 (accessed 13 augustus 2022).
Tiberius, C., & Schoonheim, T. (2014). *A frequency dictionary of Dutch: Core vocabulary for learners*. Routledge, Taylor & Francis Group.

F0 dynamics associated with prominence realisation in children with hearing impairment

Jérémy Genette[a], Jo Verhoeven[a,b]

[a] *Universiteit Antwerpen, Department of Linguistics, CLIPS Computational Linguistics and Psycholinguistics, Antwerp, Belgium*
[b] *City, University London, Department of Language and Communication Science, London, United Kingdom*

Prominence in speech is typically realised by means of greater amplitude, duration and F0 of the vowel nucleus in the syllable that carries word/sentence stress. It is well-established that F0 is the more important physical dimension. The present study reports the analysis of F0 in word realisation of two groups of children differing in hearing status. One group consisted of children with hearing impairment, while the other group consisted of age-matched children with normal hearing. The hearing-impaired children had been fitted with either a cochlear implant or a conventional hearing aid. Children had participated in a (non-)word imitation task which consisted of the repetition of monosyllables containing one of the monophthongs of Belgian Standard Dutch. Measurement and analysis of F0 in the vowel nuclei revealed interesting differences between the groups. The children with hearing impairment had the highest overall F0. In terms of the dynamics of F0 associated with prominence, all children correctly realised an underlying prominence-lending rise-fall pattern which at the phonetic level manifested itself as a falling pitch movement. In addition, the F0 contour in children with a conventional hearing aid was steepest, while it was shallowest in children with a cochlear implant. The contour in children with normal hearing was situated between the two previous groups. The observed group differences are attributed to the acoustic information provided by the type of device the hearing-impaired children are equipped with.

Phoneme Categorization after Speaking with a Bite Block

Xinyu Zhang, Esther Janse

Radboud University

The processes of speech production and speech perception are closely intertwined. Existing research on how articulator displacement affects speech perception usually has the setup where participants perform a perception task while the configurations of their articulators are changed (e.g., external force applied on the jaw in Nasir & Ostry 2009) i.e., the articulator manipulation happened during the perception task. We investigate whether and how the experience of having spoken with an altered articulatory configuration has any consequences for phoneme categorization after the manipulation has been removed. Furthermore, we ask whether having heard one's own production during the manipulated production affects this potential shift.

Participants were randomly assigned to a bite-block or a no-bite-block group. The bite block inhibited tongue height movement. All participants first read non-words without the bite block, after which they were tested on their categorization of the phonemes /ɪ/ and /ɛ/ embedded in monosyllabic Dutch words (perception pre-test). Then, dependent on their group assignment, participants would either speak with or without the bite block. An additional group assignment determined whether participants (in both bite-block and no-bite-block groups) would either hear their own speech productions or not, i.e., speaking either with ordinary auditory feedback or with speech-shaped noise to mask their auditory feedback. A post-test identical to the perception pre-test followed for all groups. By comparing categorization responses between pre-test and post-test for all bite block and auditory feedback combinations, we will be able to answer our research questions. Results will be discussed.

Exploring articulation rate entrainment: Interactions over Zoom between typical and atypical speakers

Lotte Eijk & Sophie Meekings

Department of Psychology, University of York, Heslington, United Kingdom

Speakers are known to change their speech to become more similar to the speech of their interlocutor [e.g. 2]. This phenomenon is extensively investigated in typical speakers and is known as entrainment, or e.g., alignment, accommodation. Entrainment has been found to help speakers in multiple ways, e.g., better task success or higher likeability [3, 4]. However, this process has received less attention in atypical speakers, while entrainment may indeed help these speakers [1]. In this study, we explore whether typical speakers entrain more to atypical speakers than to typical speakers when they are in interaction with an interlocutor over Zoom. More specifically, we study articulation rate of typical speakers in conversation with other typical speakers and with speakers who stutter. We furthermore elaborate on the challenges of these analyses using speech recorded via Zoom.

Forty participants (twenty pairs) participated in this study. Ten pairs of typical speakers and ten different pairs of typical-atypical speakers performed two tasks. First, they performed a picture description by themselves, and then interacted with another speaker to find the differences between pictures in three rounds of the Diapix task [5] over Zoom, lasting around 20 to 30 minutes on average.

We will elaborate on the pre-processing steps and challenges of working with speech recorded over Zoom. Moreover, we will compare articulation rate within speaker between the picture description and the Diapix task, and between the speakers of a pair during the Diapix task. Statistical analyses are being conducted and results will be ready before the Day of the Phonetics.

References

[1] Borrie, S. A., & Liss, J. M. (2014). Rhythm as a coordinating device: Entrainment with disordered speech. *Journal of Speech, Language, and Hearing Research, 57*(3), 815-824.
[2] Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*(4), 2382-2393.
[3] Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language, 76*, 29-46.
[4] Schweitzer, A., & Lewandowski, N. (2013). Convergence of articulation rate in spontaneous speech. In *INTERSPEECH*, pp. 525-529.
[5] Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., & Bradlow, A. R. (2010). The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and speech, 53*(4), 510-54

POSTERS

The perception of French nasal vowels by Belgian Dutch listeners

Hanna De Haes, Prof. Ellen Simon, Prof. Peter Lauwers

Ghent University

Several models have been proposed to study the production and/or perception of non-native phonemes by L2 learners (cf. L2LP, Escudero, 2005; PAM-L2, Best & Tyler, 2007; SLM-r, Flege & Bohn, 2021). These models generally assume that the identification of L2 phonemes depends on their phonetic similarity with phonemes of the L1. In this light, the perception of nasal vowels by L2 learners is a particularly interesting subject, because unlike nasalized vowels, nasal vowel phonemes are quite rare in the world's languages. In fact, the phonemic contrast between oral and nasal vowels exists in only 22,6 % of them (Maddieson & Precoda, 1990). To our knowledge, little research has focused on the perception of nasal vowels by L2 learners (but see e.g., Detey & Racine, 2015; Inceoglu, 2016). This paper aims to contribute to our knowledge of the acquisition of nasal vowels by investigating how Belgian Dutch listeners perceive French nasal vowels. While Dutch has no nasal vowels in its inventory, present-day Hexagonal (Parisian) French has three, being /ɛ̃/, /ɑ̃/ and /ɔ̃/.

To this purpose, twenty-one native speakers of Belgian Dutch who were intermediate learners of French performed two categorization tasks, namely a cross-linguistic task and a French task. In the cross-linguistic task, participants matched French nasal vowels to their closest Dutch equivalents and rated these vowels on a category goodness scale. In the second (French) task, they classified French nasal vowels. The target stimuli were French CV and CVC sequences containing either /ɛ̃/, /ɑ̃/ or /ɔ̃/ as their syllable nucleus. Stimuli were produced by two native speakers of Hexagonal French.

The results of the French categorization task indicated that the French nasal vowel /ɔ̃/ was identified significantly more often than /ɛ̃/ and /ɑ̃/. This can be linked to a chain shift affecting nasal vowels of present-day French. Moreover, participants rarely categorized French nasal vowels as oral vowels, but displayed asymmetric patterns of confusion between French nasal vowels. Finally, the outcomes of the cross-linguistic categorization task revealed perceptual assimilation patterns that clarify the asymmetric patterns of confusion. These perceptual assimilation patterns are discussed in the light of two theoretical models, namely the Perception Assimilation Model for L2 listeners (PAM-L2; Best & Tyler, 2007) and the Second Language Linguistic Perception model (L2LP; Escudero, 2005).

References

Best, C.T., & Tyler, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In: O.-S. Bohn & Munro M.J. (Eds.), *Language experience in second language speech learning: in honor of James Emil Flege* (pp. 13-34). John Benjamins.

Detey, S., & Racine, I. (2015). Does perception precede production in the initial stage of French nasal vowel quality acquisition by Japanese learners? A corpus-based discrimination experiment. In: *Proceedings of ICPHS* 2015, Glasgow, 10-14 August.

Escudero, P. (2005). *Linguistic perception and second language acquisition: explaining the attainment of optimal phonological categorization*. PhD dissertation. University of Utrecht.

Flege, J., & Bohn O. (2021). The revised speech learning model (SLM-r). In: R. Wayland (Ed.), *Second language speech learning: theoretical and empirical progress* (pp. 3-83). Cambridge University Press.

Inceoglu, S. (2016). Effects of perceptual training on second language vowel perception and production. *Applied Psycholinguistics*, 37(5), 1175-1199.

Maddieson, I. & Precoda K. (1990). Updating UPSID. *UCLA Working Papers in Phonetics*, 74, 104-111.

Translating a Korean Poem into English
A Case Study on Exploring the Connection between Phonetics and Emotions

Mika Aya Soehnlein[1]

David Crystal showed in his book "Sounds Appealing" that there is at least some kind of symbolic meaning to vowels and consonants. He does not further elaborate on it, instead asking: "It doesn't make sense to ask 'What is the meaning of /i:/ or /u:/?'."[2] This is, however, the intention of this research. To explore the connection between emotions and phonetics. To see whether there is any connection at all, and if a potential connection could be universally applicable or if it is specific to each language and culture.

To find an answer to this, there were three steps needed. First, the phonetic inventory of both languages, Standard South Korean[3] and General American[4], was analysed and compared to have basis on which to work on. In a second step, pre-existing research was used to investigate how emotions are represented in speech in order to see which phonemes seem to correlate with which emotions. In a third step, this was used in a practical application of translating a poem to see whether this actually holds true. The medium of poetry was chosen as it is a literary medium that was originally meant to be performed and therefore, relies more heavily on phonetics and phonology to communicate its message. The poem was translated twice. First, the traditional way and in a second translation, the research was used to match the emotional valence of source and target word as best as possible to translate the poem. Both translations were then compared with each other.

The pre-existing research clearly shows that there is a connection between emotions and phonetics and because some emotions are more similar to each other, the distinction between categories is not fully clear cut. On the basis of this pre-existing research, seven emotions were split into three bigger groups: Group 1 consisting of anger, surprise, joy, and happiness; Group 2 consisting of sadness, fear, and disgust; and a third group Neutral, comprising of a neutral state of being as well as other emotions for which there was not enough data available to clearly put them into either group. Comparing both translations, they do have equally good results concerning trying to match the emotional valence of the source. This begs the question why the first translation was already this close to the source. Is it in my specific case because of my knowledge of English or was the gut instinct right because there is an inherent emotional quality to phonemes that the brain latches onto subconsciously, this in turn then influencing the decision-making process? Right now, it is too time consuming to match the emotional valence of every word from the source and every potential word of the target language. However, if a software could be programmed that could in split seconds do the work for you, then it could certainly majorly help with preselecting useful words – making the whole translation process quicker and the quality better.

---

[1] 1st-year RMA Linguistics Student at Utrecht University. This Bachelor Thesis was written at the University of Tuebingen, the supervisor was Stu Watts

[2] David Crystal, *Sounds Appealing* (London: Profile Books, 2018), 108.

[3] Jiyoung Shin, "Vowels and Consonants," *The Handbook of Korean Linguistics*. Edited by Lucien Brown and Jaehoon Yeon (Chichester: Wiley Blackwell, 2015) 4-6.

[4] Alan Cruttenden, and Alfred C. Gimson, *Gimson's Pronunciation of English*, 18th edition (London: Routledge, 2014).

The influence of musical abilities on the processing of contrastive focus prosody in an L2: An eye-tracking study

Nelleke Jansen[1,2], Hanneke Loerts[1], Eleanor Harding[2], Deniz Başkent[2], Wander Lowie[1]

[1]Department of Applied Linguistics, Center for Language and Cognition, Faculty of Arts, University of Groningen, Oude Kijk in 't Jatstraat 26, 9712 EK Groningen, The Netherlands
[2]Department of Otorhinolaryngology/Head and Neck Surgery, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands

In Dutch and English, contrastive focus is marked by a pitch accent, and native listeners use this cue in perception to anticipate upcoming information (Mulders & Szendrői, 2016; Perdomo & Kaan, 2021). However, eye-tracking evidence suggests Dutch adults have difficulty with using prosodic cues for anticipation in English (Ge et al., 2021), possibly due to small differences between Dutch and English focus cues and a higher demand on processing resources in L2 perception. Prosody perception abilities have been associated with individual differences in musical abilities (see Jansen et al., 2022). We investigated whether musical abilities influenced the processing of contrastive focus prosody by 45 Dutch adult L2 English users, using a visual-world eye-tracking paradigm. Participants listened to sentences with the particle *only*, which associates with the direct or indirect object that receives a focus accent, e.g. *I only gave a SPOON to the girl. I didn't give a FORK to the girl*. Meanwhile, they viewed pictures showing objects and characters mentioned. We investigated to what extent participants anticipated the focus alternative (*fork*) in the second clause, indicating they had correctly interpreted the accented word (*spoon*) as the contrasted element. We analysed anticipatory fixations and tested the influence of musical abilities based on the Short-PROMS (Zentner & Strauss, 2017). We hypothesised that L2 listeners with higher scores would show more anticipatory fixations on the focus alternative and fewer on the competitor image. Initial analyses using linear regression models support our hypotheses, indicating that individuals with stronger musical abilities show a faster interpretation of focus-marking pitch accents during L2 speech processing. These findings suggest that having stronger perceptual resources underlying both music and speech processing (e.g., Patel, 2011) can even influence prosody-to-meaning mapping in an L2.

References

Ge, H., Mulders, I., Kang, X., Chen, A., & Yip, V. (2021). Processing focus in native and non-native speakers of English: An eye-tracking study in the visual world paradigm. *Applied Psycholinguistics*, *42*(4), 1057-1088.

Jansen, N., Harding, E., Loerts, H., Başkent, D., & Lowie, W. (2022). The relation between musical ability and sentence-level intonation perception: A meta-analysis comparing L1 and non-native listening. In *Proc. Speech Prosody 2022* (pp. 713-717).

Mulders, I., & Szendrői, K. (2016). Early association of prosodic focus with *alleen* 'only': Evidence from eye movements in the visual-world paradigm. *Frontiers in Psychology*, *7*, 150.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, *2*, 142.

Perdomo, M., & Kaan, E. (2021). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research, 37*(2), 349–375.

Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the Short-PROMS and the Mini-PROMS. *Annals of the New York Academy of Sciences*, *1400*(1), 33-45.

The speaker in speech – linguistic context in forensic speaker comparisons

Willemijn Heeren, Meike de Boer, Laura Smorenburg*

Leiden University Centre for Linguistics, Leiden University
**now at Utrecht Institute for Linguistics OTS, Utrecht University*

In forensic speaker comparisons, one or more disputed speech samples of an unknown speaker are typically compared to one or more speech samples of a suspect. One concern in such analyses is the effect of phonetic-linguistic context on the speech sounds sampled for comparison; many phonetic studies show that contexts affect the acoustic realization of speech sounds. For example, stressed syllables evoke more canonical pronunciations, utterance-initial speech sounds are typically produced with more effort, and talking over the telephone changes speech behaviour.

To better understand the relevance of these acoustic effects for forensic speaker comparisons, we have been studying how segmental features that characterize speakers' voices depend on what speakers are saying (NWO VIDI project 276-75-010). More specifically, how does the speaker-specificity of speech sounds depend on its direct context (Heeren, 2020, Smorenburg & Heeren, 2020, 2021), on the speech channel (Smorenburg & Heeren, 2022), and on the language spoken (De Boer & Heeren, 2020; De Boer, Quené & Heeren, 2022). In this overview presentation we would like to share some of our main project results obtained on existing research databases (CGN, Oostdijk, 2000; D-LUCEA, Orr & Quené, 2017; WYRED, Gold et al., 2018), and show how our findings are relevant for and translate to forensically-realistic speech data taken from the NFI-FRIDA corpus (Van de Vloed et al., 2020).

References

Boer M.M. de, Quené H. & Heeren W.F.L. (2022), Long-term within-speaker consistency of filled pauses in native and non-native speech, *JASA Express Letters* 2(3): 035201.

Boer M.M. de & Heeren W.F.L. (2020), Cross-linguistic filled pause realization: the acoustics of uh and um in native Dutch and non-native English, *Journal of the Acoustical Society of America* 148(6): 3612-3622.

Gold, E., Ross, S., & Earnshaw, K. (2018). The "West Yorkshire Regional English Database": Investigations into the generalizability of reference populations for forensic speaker comparison casework. In *Proceedings of INTERSPEECH* (Vol. 2018–Sept., pp. 2748–2752)

Heeren W.F.L. (2020), The Effect of Word Class on Speaker-dependent Information in the Standard Dutch Vowel /aː/, *Journal of the Acoustical Society of America* 148(4): 2028-2039.

Oostdijk, N. (2000). "Het corpus Gesproken Nederlands," (The Spoken Dutch Corpus), *Nederlandse Taalkunde* 5, 280–284.

Orr, R., and Quené, H. (2017). " D-LUCEA: Curation of the UCU Accent Project data," in *CLARIN in the Low Countries*, edited by J. Odijk and A. van Hessen (Ubiquity Press, Berkeley), pp. 177–190.

Smorenburg B.J.L. & Heeren W.F.L. (2020), The distribution of speaker information in Dutch fricatives /s/ and /x/ from telephone dialogues, *Journal of the Acoustical Society of America* 147(2): 949-960.

Smorenburg B.J.L. & Heeren W.F.L. (2021), Acoustic and speaker variation in Dutch /n/ and /m/ as a function of phonetic context and syllabic position, *Journal of the Acoustical Society of America* 150(2).

Smorenburg L. & Heeren W.F.L. (2022), The effect of linguistic contexts on the acoustics and strength-of-evidence of /s/. *Proceedings 30th annual conference of the International Association for Forensic Phonetics and Acoustics*, p. 21-22.

Van der Vloed, D., Bouten, J., Kelly, F., & Alexander, A. (2018). NFI-FRIDA–Forensically realistic interdevice audio database and intial experiments. In *27th Annual Conference of the International Association for Forensic Phonetics and Acoustics* (IAFPA) (pp. 25-27).