



2024 한양 음성·언어인지과학 미니국제심포지엄

Mini HISPhonCog 2024

Mini Hanyang International Symposium
on Phonetics and Cognitive Sciences of Language 2024

*Department of English Language and Literature
Hanyang Institute for Phonetics & Cognitive Sciences of Language (HIPCS)
Hanyang Humanities Enhancement Center (H²EC)*

2024.3.21 (목) 15:00~18:30

인문관 3층 316호 (3rd floor)

참관문의(한양대 소속 무료): 02-2220-2507



Holger Mitterer
(U of Malta)



Jeremy Steffman
(U of Edinburgh)



Xinran Ren
(Sun Yat-sen U)



Jiyoung Jang
(HIPCS, Hanyang U)



Argyro Katsika
(UC Santa Barbara)



Sejin Oh
(HIPCS, Hanyang U)



Richard Hatcher
(HIPCS, Hanyang U)

<http://site.hanyang.ac.kr/web/hisphoncog>

HISPhonCog 2024 Program

March 21, Thursday, 2024

15:00-18:30

15:00-15:30	Jiyoung Jang (HIPCS) & Argyro Katsika (University of California, Santa Barbara) <u>Focus structure and articulatory strengthening in edge-prominence language: an articulatory study of Seoul Korean</u>
15:30-16:00	Sejin Oh (HIPCS) <u>Temporal aspects of a prevocalic glide in gestural terms in Korean</u>
16:00-16:30	Richard Hatcher (HIPCS) <u>Prosodic case-marking in the absence of morphological case markers in Korean</u>
16:30-17:00	Break
17:00-17:30	Xinran Ren (Sun Yat-sen University, China) <u>Third Language Perception Predicted by Acoustic and Perceptual Similarity between L3 and L1/L2 Vowels</u>
17:30-18:00	Jeremy Steffman (University of Edinburgh, UK) <u>Rate-dependent speech perception interacts with prosodic phrasing in Korean</u>
18:00-18:30	Holger Mitterer (University of Malta, Malta) <u>CAN SEGMENTAL DETAIL INFLUENCE PROSODIC ANALYSIS? THE CASE OF CONTRACTIONS IN ENGLISH</u>

Focus structure and articulatory strengthening in edge-prominence language: an articulatory study of Seoul Korean

Jiyoung Jang¹ & Argyro Katsika²

¹*Hanyang Institute for Phonetics & Cognitive Sciences of Language,*

²*University of California, Santa Barbara*

Abstract: Articulatory gestures under phrasal prominence undergo strengthening, becoming longer, larger, and faster (cf. Cho, 2006). Limited research, mainly from head-prominence languages, suggests that this articulatory strengthening does not simply correspond to a prominent vs. non-prominent distinction, but that it encodes focus structure instead, with the phonetic effects increasing not just from unfocused to focused units, but also—roughly—from broad focus to narrow focus, and then to contrastive focus (Breen et al., 2010; Hermes et al., 2008; Roessig & Mücke, 2019; Katsika et al., 2023). However, it is unclear whether focus structure is encoded in edge-prominence systems. Here, we turn to Seoul Korean, an edge-prominence language, in which the focused word is assumed to start an Accentual Phrase (AP) and exhibits prominence-induced strengthening.

Analyses of kinematic duration, displacement, and velocity from data of six native Seoul Korean speakers examine the degree of strengthening on focused AP-initial gestures. Results show that, in Korean, focus-induced strengthening reflects the focus structure, although kinematic dimensions differ in the number of focus types they distinguish. Yet, the order of encoded types remains consistent and similar to that found in head-prominence languages (Breen et al., 2010; Roessig et al., 2019; Katsika et al., 2023). The findings support the view that a hierarchy of prominence might emerge from the interface of prosodic structure with focus structure and suggest that this might be a property that holds across categories of prosodic typology.

References

- Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory Phonology*, 8, 519-548.
- Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7-9), 1044-1098.
- Hermes, A., Becker, J., Mücke, D., Baumann, S., & Grice, M. (2008). Articulatory gestures and focus marking in German. In *Proceedings of Speech Prosody 2008*, 457-460.
- Roessig, S., & Mücke, D. (2019). Modeling dimensions of prosodic prominence. *Frontiers in Communication*, 4, 44.
- Katsika, A., Jang, J., Krivokapić, J., Goldstein, L., & Saltzman, E. (2023). A hierarchy of prominence: The production and perception of focus in American English. In *Proceedings of the International Congress of Phonetic Sciences* (pp. 1696-1700).

Temporal aspects of a prevocalic glide in gestural terms in Korean

Sejin Oh¹, Sahyang Kim^{1,2}, Taehong Cho¹

¹ Hanyang Institute for Phonetics & Cognitive Sciences of Language, Hanyang University, ² Hongik University

Across languages, a glide /j/ may be analyzed differently: as part of a ‘segment sequence’, a ‘complex segment’, or a ‘diphthong’ (e.g., [1-5]). But characterizing its exact nature can be elusive due to different approaches employed by different researchers. The Korean glide /j/ presents such a case. Some findings concern its phonological behavior in forming an onset cluster (e.g., [6]), a complex segment (e.g., [7]), or a diphthong (e.g., [8]). Others analyze its acoustic characteristics: a brief F2 steady state, for example, is used to characterize /j/ as part of a complex segment [9]. But the same acoustic evidence might also support a diphthongal formation as the acoustic form often obscures the temporal relations of the actual articulatory gestures involved. We therefore directly examine temporal realizations of articulatory gestures.

In Articulatory Phonology (e.g., [10, 11]), gestures can be hypothesized to be timed simultaneously for a complex segment, but sequentially for a segment sequence. [12, 13] showed that for a complex segment, C-/j/ gestures are indeed timed together, so that their onset-to-onset lag is much less influenced by variation in C duration, compared to a segment sequence in which timing of /j/-onset relative to C-onset (onset-to-onset lag) is positively correlated with C duration. We adopt this approach to examine /C-/j/ gestural coordination in /mjV/, compared to the reference case of /mV/ gestures assumed to be timed simultaneously. The aim is to understand the temporal characteristics of the Korean /j/ in gestural terms, which will inform whether /j/ forms a complex segment or a segment sequence.

The results revealed that /mj/ demonstrates a longer onset lag and a shorter glide formation duration compared to /mi/, refuting the possibility of the Korean glide being part of a diphthong. Furthermore, we observe the nearly flat regression lines for both conditions, showing that variations in G1 duration have minimal impact on Onset lag. Crucially, however, this pattern is not consistent across all speakers, where some speakers (F11 and M15) exhibit an English segment sequence-like pattern for Korean /mj/, while showing onset-to-onset coordination for /mi/. The results imply that the surface timing of the C+glide gestures is not as invariant as the phonologically specified gestural coordination (sequential/simultaneous) would predict. This variation seems to accommodate the phonotactics that imposes temporal constraints on the onset (typically a singleton C but two Cs only with a glide overlapping with C1), reflecting the range of coarticulation permissible in the phonetic grammar of the language.

References

- [1] Anderson, S. R. (1976). Nasal consonants and the internal structure of segments. *Language*, 326-344.
- [2] Davis, S., & Hammond, M. (1995). On the status of onglides in American English. *Phonology*, 12(2), 159-182.
- [3] Lombardi, L. (1990). The nonlinear organization of the affricate. *Natural Language & Linguistic Theory*, 8(3), 375-425.
- [4] Riehl, A. (2008). *The phonology and phonetics of nasal-obstruent sequences* (Doctoral dissertation, Cornell University).
- [5] Sagey, E. C. (1986). *The representation of features and relations in non-linear phonology* (Doctoral dissertation, Massachusetts Institute of Technology).
- [6] Cheon, S. Y. (2002). Glides as consonants in Korean. *LANGUAGE RESEARCH*, 38(2), 619-646.
- [7] Kim, H. Y. (1998). Prenucleus glides in Korean. *Studies in the Linguistic Sciences*, 28(1), 113-35.
- [8] Kim, C. W., & Kim, H. Y. (1991). The character of Korean glides. *Studies in the Linguistic Sciences*, 21(2), 113-125.
- [9] Suh, Y., & Hwang, J. (2016). The Korean prevocalic palatal glide: A comparison with the Russian glide and palatalization. *Phonetica*, 73(2), 85-100. DOI:10.1159/000444189
- [10] Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215-249). New York: Cambridge University Press.
- [11] Iskarous, K., & Pouplier, M. (2022). Advancements of phonetics in the 21st century: A critical appraisal of time and space in Articulatory Phonology. *Journal of Phonetics*, 95(2), 1-28. doi:10.1016/j.wocn.2022.101195
- [12] Shaw, J. A., Durvasula, K., & Kochetov, A. (2019). The temporal basis of complex segments. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, 676-680. Canberra: Australasian Speech Science and Technology Association Inc.
- [13] Shaw, J. A., Oh, S., Durvasula, K., & Kochetov, A. (2021). Articulatory coordination distinguishes complex segments from segment sequences. *Phonology*, 38(3), 437-477. Doi:10.1017/S0952675721000269.

Prosodic case-marking in the absence of morphological case markers in Korean

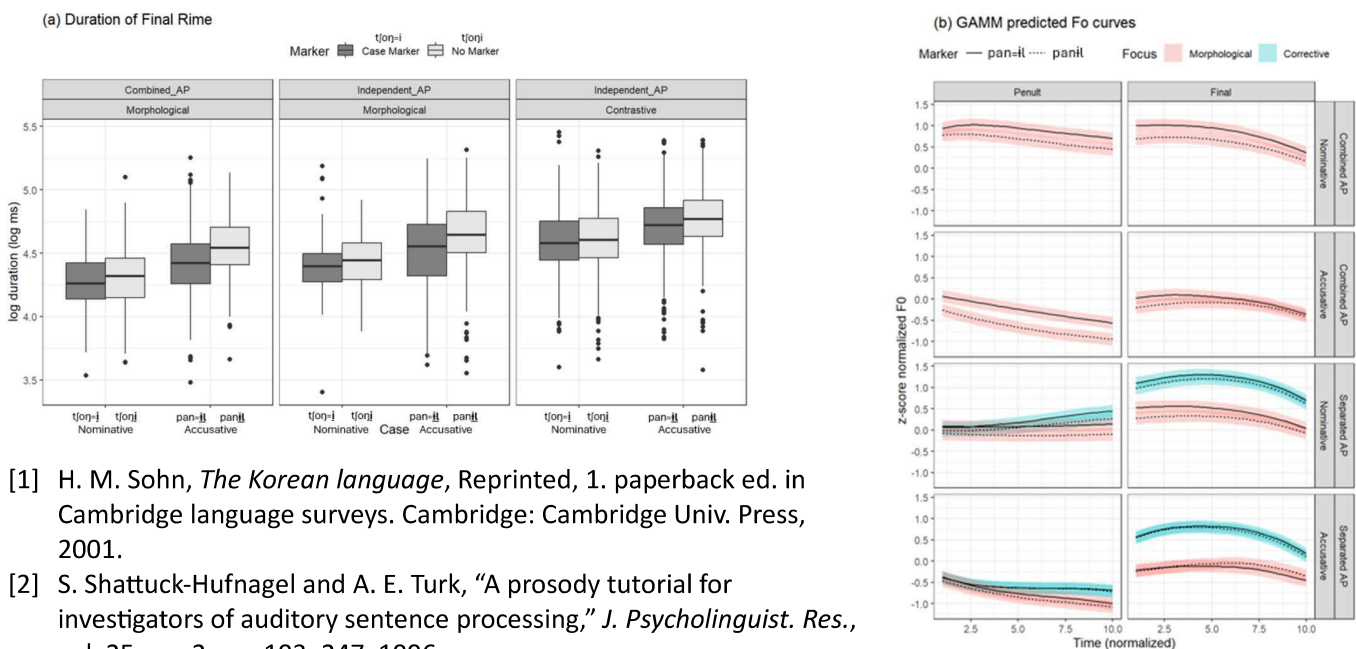
Richard Hatcher¹, Sahyang Kim^{1,2}, Taehong Cho¹

¹Hanyang Institute of Phonetics and Cognitive Science of Language, Hanyang University, ²Hongik University

Introduction: Korean's morphological system richly employs particles for grammatical functions, including case marking [1]. However, casual speech often sees the omission of these markers, leaving a potential gap in signaling grammatical roles. This study hypothesizes that prosody—through tone-segment alignment, magnitude of fundamental frequency (F0) extrema, and durational patterns—compensates for this absence, thus maintaining the morphosyntactic structure intact [2], [3], [4].

Methods: Utilizing speech data from 21 native Korean speakers, this investigation contrasts segmentally neutralized noun phrases with and without nominative or accusative markers in varied focus conditions. Participants were presented with 16 pairs of test words, embedded in sentences under two focus conditions: broad and corrective. To aid in elicitation and ensure clarity of context, each test word was accompanied by a visual representation of its meaning. The study initially analyzed a total of 4,032 tokens, derived from 21 speakers who each produced 16 pairs of words, differentiated by two morphological structures, across two focus contexts, and repeated three times; however, only 3,909 tokens were ultimately used for analysis after excluding those containing speech errors.

Results: The findings reveal significant prosodic differentiation between cases, with prosodic cues modifying both the temporal and tonal dimensions of speech to signal case roles. The influence of morphological structure on the duration of the final rime interacts with focus type, where, despite corrective focus exhibiting longer overall durations, the effect of morphological structure is noticeably diminished. Similarly, the effect was stronger for targets in the accusative case where the case marker consists of two segments, =*il*. In regard to f0, words with case markers exhibited overall higher contours and in some cases a later peak, highlighting a prosodic distinction tied to morphological marking. As with duration, the effect of morphological structure was attenuated but retained significance under corrective focus. The study reveals Korean's adept use of prosody to signal grammatical roles when morphological markers are absent, with these distinctions most pronounced in broad focus contexts, ensuring clear and effective communication.



- [1] H. M. Sohn, *The Korean language*, Reprinted, 1. paperback ed. in Cambridge language surveys. Cambridge: Cambridge Univ. Press, 2001.
- [2] S. Shattuck-Hufnagel and A. E. Turk, "A prosody tutorial for investigators of auditory sentence processing," *J. Psycholinguist. Res.*, vol. 25, no. 2, pp. 193–247, 1996.
- [3] I. Feldhausen, "The prosody-syntax interface in Catalan," University of Pottsdam, 2008. Accessed: Jan. 15, 2024.
- [4] P. Strycharczuk, "Phonetic Detail and Phonetic Gradience in Morphological Processes," in *Oxford Research Encyclopedia of Linguistics*, 2019. doi: 10.1093/acrefore/9780199384655.013.616.

Third Language Perception Predicted by Acoustic and Perceptual Similarity between L3 and L1/L2 Vowels

Xinran Ren (Sun Yat-sen University) & Peggy Mok (The Chinese University of Hong Kong)

A frequently asked question regarding L3 acquisition is which existing language system, L1, L2 or L1+L2, is transferred to L3. This study does not only aim at finding out the source of transfer to L3, but more importantly, it also tests the role of two factors, acoustic and perceptual similarity, in predicting the discrimination accuracy of L3 vowels. These two factors play an important role in several L2 perception models, such as PAM-L2 (Best and Tyler 2007) and L2LP (Escudero 2005), while not many studies extended the perceptual assimilation model to L3, and acoustic similarity has rarely been investigated in L3 studies.

22 college-aged L3 learners (8M, 15F) participated in a perceptual mapping task (with all possible vowels in Cantonese and English as choices), an identification task of target Korean vowels /u,o,ʌ/ and a reading task containing L1 and L2 vowels. To measure acoustic similarity between L3 and L1/L2 vowels, cross-language discriminant analyses (LDA) were conducted using F1, F2, F3 and duration normalized by Lobanov's z-score. For this aim, we compared the acoustic properties between the L3 learners' L1 and L2 vowel productions and the native Korean vowel tokens that were used as stimuli in the identification task. Perceptual similarity was measured through perceptual mapping percentages. Subsequently, Levy (2009)'s overlap scores were adopted to quantify acoustic and perceptual similarity. It represents the proportion of overlap between L3 target and L1 and L2 vowels, so a higher score means more overlap.

The acoustic and perceptual overlap scores were calculated according to the results of cross-linguistic acoustic comparisons and perceptual mappings. As is shown in the following table, the vowel contrasts /u-o/ and /o-ʌ/ have higher acoustic and perceptual scores than /u-ʌ/, which successfully predicted a more accurate discrimination of /u-ʌ/ than /u-o/ and /o-ʌ/. The comparisons between different vowel contrasts were confirmed by paired-samples t-tests. Therefore, in line with previous studies on L2 perception, the results in the current study also demonstrate that a larger acoustic or perceptual overlap between L3 and L1/L2 will lead to more difficulties in L3 vowel discrimination. The results will be further discussed from a perspective of L3 acquisition.

Overlap scores	NK Contrasts		
	/u-o/	/o-ʌ/	/u-ʌ/
Acoustic	34.8	31.2	8.0
Perceptual	64.8	48.6	31.5

Discrimination Accuracy: /u-o/ ≈ /o-ʌ/ < /u-ʌ/

References:

- Best, C. T. & Tyler, M. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In *Language Experience in Second-Language Speech Learning: In Honor of James Emil Flege*. Edited by O. Bohn and M. J. Munro. Amsterdam: John Benjamins, pp. 13–34.
- Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Netherlands Graduate School of Linguistics.
- Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, 126(5), 2670–2682.

Rate-dependent speech perception interacts with prosodic phrasing in Korean

Jeremy Steffman, Sahyang Kim, Taehong Cho & Sun-Ah Jun

Introduction: Rate-dependent speech perception is the phenomena by which temporal cues in speech are perceived relative to context. A preceding fast context causes subsequent cues to be perceived as relatively slow/long (e.g. an ambiguous voice onset time (VOT) will be perceived as longer); a preceding slow context makes the same speech sound relatively fast (e.g., a VOT value will be perceived as shorter). These effects are usually understood via auditory *contrast* mechanisms [1,2]. However, prosodic phrasing/grouping, which is often cued by local duration modulations around a prosodic juncture, has also been shown to generate predictive or expectation-based effects on temporal cue processing, based on prosodic strengthening [2,3]. In this study we test how speech rate context *interacts with prosodic phrasing* in Korean, exploring how rate effects operate across prosodic contexts. We test the perception of the contrast between aspirated stops (/k^h/ with long VOT and a short following vowel) and fortis stops (/k*/ with short VOT and a long following vowel). **Korean prosodic phrasing** induces “domain-initial strengthening” in speech production [4,5]. For aspirated stops, this entails VOT lengthening. For fortis stops, this entails slight VOT shortening, or little change. For *both* stop categories, the subsequent vowel in a CV sequence is lengthened. [3] showed that these patterns have perceptual consequences. When there was a prosodic contextual cue that would lead to an expectation of domain-initial strengthening, the vowel in the CV sequence was perceived as shorter than when there was no such cue (i.e. due to expected phrase-initial vowel lengthening, a vowel “needed to be longer” to cue /k*/). Crucially, this effect of prosodic grouping was observed with temporal context controlled and only F0 cuing prosodic structure. **Method:** 32 speakers of Seoul Korean categorized a VOT continuum as /k*/ (tense) or /k^h/ (aspirated), in the carrier phrase shown in Figure 1. Six conditions crossed **rate** (normal vs. slow) and prosodic **phrasing** (no boundary vs. boundary vs. boundary+pause). Rate manipulation resulted from the linear expansion of pre-target material (distal context and local context in Fig. 1). The phrasing manipulations involved lengthening only the pre-target syllable, resulting in an intonational phrase (IP) boundary condition, and creating a stronger boundary condition by introducing a pause (IP+pause condition). Responses (252 per participant) were analyzed with Bayesian mixed-effects logistic regression (maximal by-participant random slopes). In Fig. 2 we report estimates and “pd” (prob. of direction) which when > 97.5 is taken as evidence for an effect. We made several **predictions:** 1) There should be a typical rate-based contrast effect in the no boundary condition. 2) Based on [3] there should be a domain-initial strengthening effect across prosodic boundary conditions in terms of /k*/ responses, ordered as no boundary > IP > IP+pause. It is an exploratory question if rate effect persists in the IP conditions, or interact with phrasing effects. **Results:** Prosodic strengthening effects: an interaction between rate and prosodic boundary was observed (pd = 100), reflecting a difference in phrasing conditions, *only* when rate was slow (Fig. 2B, “slow” rate bars ascend left to right). This reflects the predicted phrasing effect. Speech rate effects: A credible rate-based contrast effect is found in the no boundary condition: slow rate makes VOT sound relatively short and *increases* /k*/ responses. No rate effect is found in the IP condition, and an unexpected reversal of the rate effect is present in the IP+pause condition. We speculate that this effect results from a **re-interpretation of rate slow-down cues in the presence of a strong prosodic boundary**, i.e. when a pause is present. No effect of rate in the IP condition is hypothesized to result from a competition of rate and phrasing influences, though this remains to be further tested. **In summary:** We find that 1) prosodic phrasing effects obtain (which defy local contrast effects: longer pre-boundary vowels in slow + IP vs. slow + no boundary favors perception of longer-VOT /k^h). However, phrasing effects only occur at slower speech rates. 2) Canonical rate effects obtain when there is no prosodic boundary, but disappear, or are reversed, when a prosodic boundary (i.e. with a pause) is present. These results thus provide novel insight into how rate-dependent perception interacts with phrasing, and show that linguistic structure (prosodic boundaries) plays a mediating role.

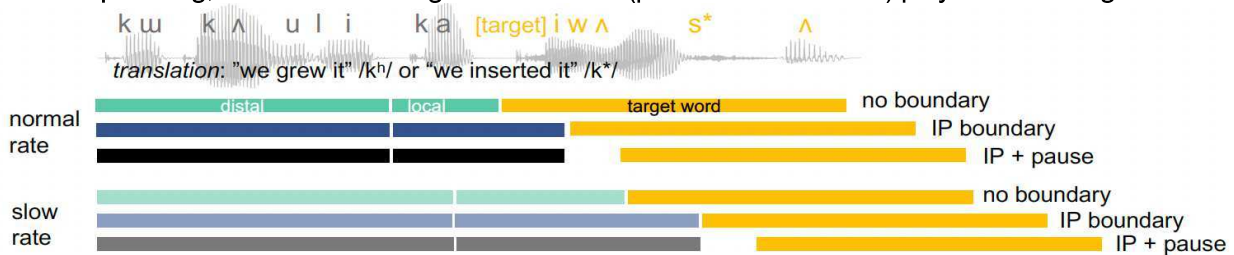


Figure 1: Schematic speech rate conditions and prosodic boundary conditions in the stimuli.

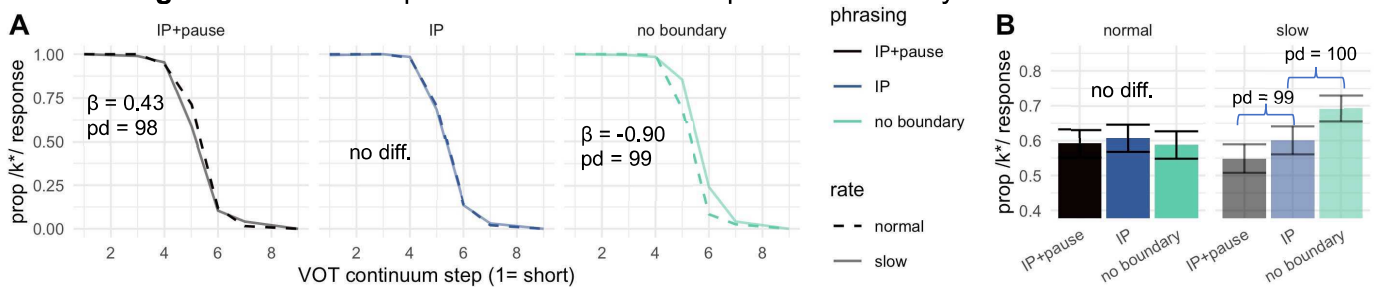


Figure 2: Categorization for the continuum (A) and pooled for steps 3-8 to show context effects (B)

References [1] Newman, R. S., & Sawusch, J. R. (1996). Perceptual normalization for speaking rate: Effects of temporal distance. *Perception & Psychophysics*, 58(4), 540-560. [2] Mitterer, H., Cho, T., & Kim, S. (2016). How does prosody influence speech categorization?. *JPhon*, 54, 68-79. [3] Steffman, J., Kim, S., Cho, T., & Jun, S.-A. (2022). Prosodic phrasing mediates listeners' perception of temporal cues: Evidence from the Korean Accentual Phrase. *JPhon*, 94, 101156. [4] Cho, T., & Jun, S. A. (2000). Domain-initial strengthening as enhancement of laryngeal features: Aerodynamic evidence from Korean. *UCLA working papers*, 57-70. [5] Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *JPhon*, 29(2), 155-190.]

CAN SEGMENTAL DETAIL INFLUENCE PROSODIC ANALYSIS? THE CASE OF CONTRACTIONS IN ENGLISH

Holger Mitterer^{1,3}, Sahyang Kim², Taheong Cho³

¹University of Malta, ²Hongik University ³Hanyang University
holger.mitterer@um.edu.mt, sahyang@gmail.com, tcho@hanyang.ac.kr

The current project rests on two well-attested phenomena. First of all, prosodic structure influences sentence comprehension (Steinhauer et al., 1999; Weber et al., 2006). Secondly, prosodic structure is phonetically reflected not only on suprasegmental dimensions (e.g., f₀, duration, and amplitude) but also on segmental dimensions (e.g., coarticulation, segmental strengthening or reduction) (Cho et al., 2017). This leads to the question whether the listener makes use of the prosodic-structurally conditioned segmental detail to compute the prosodic structure of a given utterance, which then in turn would influence sentence comprehension.

In two previous projects, we investigated this issue. Mitterer et al. (2020) tested the role of the glottal stop in Maltese, which, next to its role as phoneme, also appears as a prosodic boundary marker. The results showed that listeners use the glottal stop to decide between an early or late closure, even though the glottal stop must be processed as segmental information (due to its phonemic status). In Mitterer et al. (2024), we exploited the German *verum* contrast on the auxiliary *haben* (Engl. ‘to have’), which indicates agreement with a yes/no question. Instead of using a pitch accent on the auxiliary, it was produced in its full phonetic form [habən], which, according to a corpus study, is an infrequent, and hence marked, pronunciation ([habm] being a more common pronunciation). Using a web-based experiment with a “video-game” like set-up, the results showed that listeners were able to use segmental detail as a cue to prosodic structure; however, the effect was relatively small and at least partly due to learning of experimental contingencies.

In the current project, we make use of contractions in English. Consider the following three replies to the question: “The aliens haven’t shot the robot, right?”, which vary in the form of the auxiliary *have*.

- (1a) *Well, they HAVE shot the robot.* (full, pitch accent)
- (1b) *Well, they’ve shot the paddle.* (contracted, no pitch accent)
- (1c) *Well, they have shot the paddle.* (full, no pitch accent)

In Answer 1a), the pitch accent on *have* indicates a contrast with the question, implicating that the robot has been shot after all, while 1b) indicates that, as implied in the question, the robot was shot. The interesting case is 1c) in which the lack of contraction might indicate prosodic weight on the auxiliary, indicative of prosodic weight, which might lead to the inference that the speaker intends contrast (as in 1a). The first two experiments showed that a paradigm that avoids potential learning (by sometimes using contextually inappropriate prosodic means) is still able to reveal the influence of prosody on sentence comprehension, but that this effect could only be driven by the supra-segmental pitch accent (see 1a) but not by segmental means alone (1c).

However, the rendition of the uncontracted *have* had a reduced /h/ in the first two experiments. Experiments 3 therefore used a version with a full /h/ in versions of 1c) and now listeners were able to use this as a segmental cue to prosodic structure. Experiment 4 ruled out that this was due to the increased duration with a full /h/. The results hence show that segmental detail is used as a cue to the prosodic structure. Moreover, since we use a mouse-tracking task, this influence arises at an early stage of processing.