

## Cross-linguistic word-medial stop lenition: A Functional PCA approach

Seung Suk Lee<sup>1</sup>, Morgan Sonderegger<sup>1</sup>

<sup>1</sup>McGill University

**Introduction.** Word-medial lenition is pervasive cross-linguistically, but it remains unclear how similar the phenomenon described as “lenition” is in different languages. Two common acoustic measures of stop lenition are consonant duration and the relative change in intensity from neighboring vowels which captures the shallower dip in intensity during the stop due to an incomplete closure [1,2,3,4,5] associated with lenited stops. Word position has a small effect on these measures compared to others (e.g. speaking rate), so careful statistical modeling is needed. Word-medial stops have been found to be more lenited, measured by intensity dip, compared to word-initial ones in a few languages in studies with careful statistical controls [1,2,3,4,5], and to have shorter duration across many languages in a study with minimal controls [6]. This effect of word position on stop realization is termed ‘continuity lenition’ in [7]—stops are lenited to signal the continuity of the current prosodic constituent (i.e., word)—suggesting *this phenomenon may be present across languages*, regardless of whether “lenition” has been documented. Previous studies differ in how “degree of lenition” is operationalized [1,2,3,4,5], making it difficult to compare results of different languages, and meaning that there has not been a large cross-linguistic study of stop lenition. This study tests the idea above. We propose and evaluate a data-driven intensity-dip measure, using FPCA [8], to measure stop lenition across 9 languages. Across force-aligned speech corpora [9,10,11], we use mixed-effects regression models to investigate how languages differ in the effect of word position on both stop duration and degree of intensity-dip.

**Methods.** We used [12] to extract acoustic measurements and query relevant tokens from forced-aligned speech corpora of French, German, Polish, Russian, Spanish, Swedish, and Turkish (from [9]), Seoul Korean [10], and American English [11], totaling 1.46m tokens. We measured intensity for every 10 ms from [V<sub>1</sub>CV<sub>2</sub>] instances where C was either a nasal or an oral stop varying in laryngeal categories (lenis/aspirated/fortis for Seoul Korean, voiced/voiceless for others). [C] was always a word-initial or medial syllable onset as nasals were included as a “fully lenited” baseline. Following landmark registration [13] to time normalize and smooth the intensity contours, and within-token centering, we performed FPCA to identify the component that captures the most variation in intensity contour shape. We fit two mixed effects regression models for each language, for stop duration and PC1, to investigate how nasal, voiced, and voiceless stops in each language vary in PC1 and duration, as a function of word position, after controlling for major covariates such as speech rate, place of articulation, neighboring vowel height, and syllable stress [1,2,3,4,5].

**Results.** **Fig. 1** shows PC1, which accounts for 40% of the variation, and seems to capture the intensity “dip”. Two tests further supported that PC1 is a reasonable measure of stop lenition. First, PC1 was highly correlated with the hand-adjusted measure used in [5] to measure lenition of lenis stops in the same dataset for Seoul Korean. Second, PC1 generally patterns as nasal < voiced < voiceless stops across languages (nasal < lenis < aspirated < fortis for Korean), as shown by model predictions in **Fig. 2**. **Fig. 3** shows the model estimated marginal mean differences in duration and PC1, between word-initial and medial positions for each stop category (nasal, voiced, and voiceless). Medial **voiced stops (1)** were consistently shorter and more lenited than initial ones across all languages. Medial **nasal stops (2)** were generally shorter, but in English, German, and Swedish, they were less lenited. The effect on PC1 was the largest in Seoul Korean, which is expected as nasals are domain-initially denasalized (less-lenited) [14]. Unlike voiced and nasal stops, the difference between medial and initial **voiceless stops (3)** was more language specific, with only 4/9 languages showing the expected pattern. Our results suggest that ‘continuity lenition’ may exist in *all* languages, at least for voiced stops, as a phonetic tendency, and thus could be used as a cue in speech segmentation [6].

Fig. 1 Black line shows the average intensity contour for [VCV]. Colored lines show the average line  $\pm$  quarters of standard deviation of PC1.

Fig. 2 Model estimated marginal mean standardized PC1 for nasal, voiced (or lenis for Korean), and voiceless (or aspirated/fortis for Korean) stops across languages.

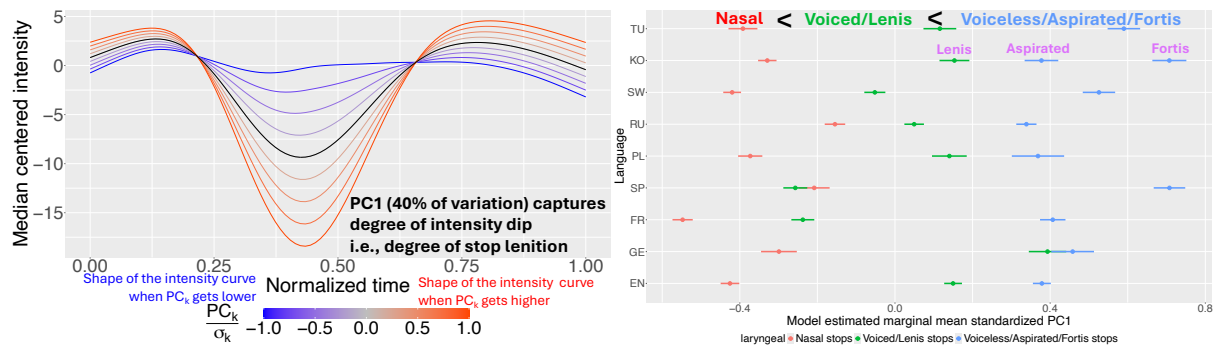
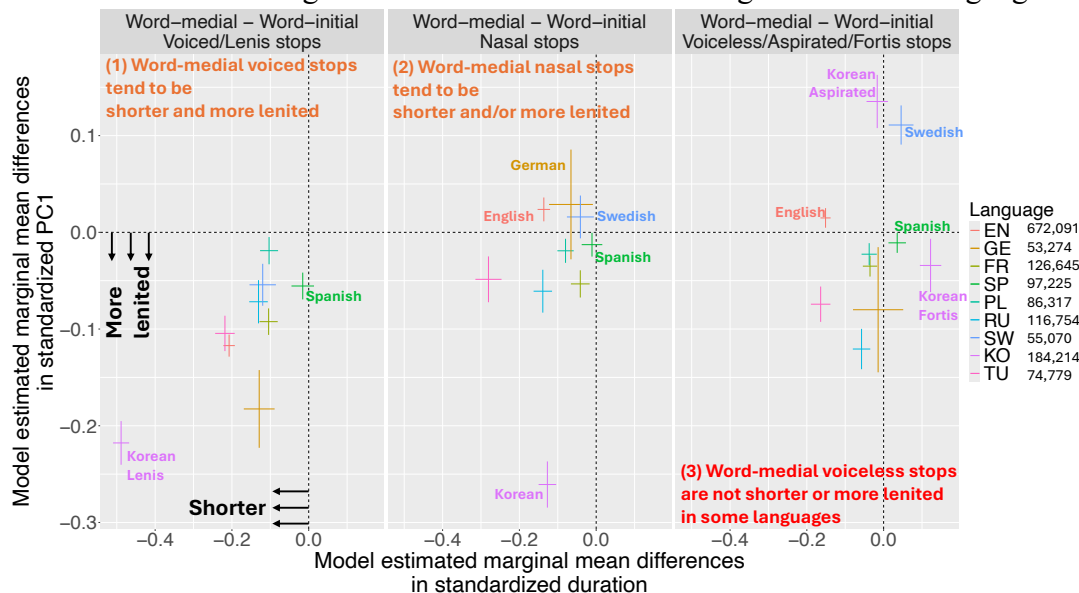


Fig. 3 Model estimated marginal mean differences in standardized duration and PC1. A point in the third quadrant ( $x < 0, y < 0$ ) indicate that word-medial tokens were shorter and more lenited than initial ones in that language. Cross marks indicate 95% confidence intervals for duration and PC1. Numbers in the legend are number of tokens investigated for each language.



**References.** [1] Kingston, J. (2008). Lenition. *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*. [2] Ennever, T., et al. (2017). A replicable acoustic measure of lenition and the nature of variability in Gurindji stops. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*. [3] Katz, J. and Pitzanti, G. (2019). The phonetics and phonology of lenition: A Campidanese Sardinian case study. *LabPhon 10(1)*. [4] Cohen Priva, U. and Gleason, E. (2020). The causal structure of lenition: A case for the causal precedence of durational shortening. *Language 96*. [5] Lee, S. (2025). *Simultaneous Encoding and Decoding of Laryngeal Contrasts and Prosodic Structure in Seoul Korean*. UMass Ph.D. thesis. [6] Blum, F., et al. (2024). Consonant lengthening marks the beginning of words across a diverse sample of languages. *Nature Human Behaviour*. [7] Katz, J. (2016). Lenition, perception and neutralisation. *Phonology*. [8] Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis*. [9] Schultz, T., et al. (2013). GlobalPhone: A multilingual text & speech database in 20 languages. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. [10] Yun, W., et al. (2015). The Korean Corpus of Spontaneous Speech. *Phonetics and Speech Sciences*. [11] Panayotov, V. et al. (2015). Librispeech: An ASR corpus based on public domain audio books. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* [12] McAuliffe, Michael, Elias Stengel-Eskin, Michaela Socolof, and Morgan Sonderegger (2017). Polyglot and Speech Corpus Tools: a system for representing, integrating, and querying speech corpora. In *Proceedings of Interspeech 2017*. [13] Gubian, M., et al. (2015). Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts. *Journal of Phonetics*. [14] Jang, J., et al. (2018). Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences. *The Journal of the Acoustical Society of America*.