

## Shorter vowel duration as a potential word segmentation cue: A study with listeners of Taiwanese Southern Min

Shu-chen Ou<sup>1</sup> and Zhe-chen Guo<sup>2</sup>

<sup>1</sup>National Sun Yat-sen University, Taiwan; <sup>2</sup>Northwestern University, USA

The cross-linguistic tendency for listeners to exploit longer word-final vowels in speech segmentation is well-documented [1]. One explanation holds that these longer vowels have delayed syllable perceptual centers, which confer perceptual salience on the word-final boundaries [2]. This view predicts that shortening word-final vowels would reduce the word-finality percept and impede segmentation. This study investigates the use of vowel shortening by listeners of Taiwanese Southern Min (TSM), in which the predicted shortening effect may be overridden by the language-specific sound distribution.

In TSM, vowels can combine with a glottal stop coda [ʔ] to form checked syllables. Due to this coda segment, the vowels are phonetically realized with shorter duration and glottalization, providing perceptual cues to their contrasts with non-checked syllables [3]. Yet, because of TSM tone sandhi, the [ʔ]-ending syllables occur only word-finally and lose [ʔ] in non-final positions (e.g., [ti<sup>33</sup> ba<sup>23</sup>] ‘pork’ vs. [ba<sup>53</sup> kuã<sup>55</sup>] ‘jerky;’ the numbers indicate tonal pitch heights with 5 being the highest) [4]. Thus, an alternative prediction is that shorter vowels are perceived as word-final by TSM listeners and used to segment speech, especially when accompanied by a [ʔ] coda and glottalized.

We tested the competing predictions with an artificial language (AL) learning experiment [5]. Participants first listened to speech streams in which nonsense trisyllabic words of an AL (e.g., [megiba]) were concatenated without pauses. There were six AL words, all having a flattened pitch and equalized intensity with the final vowels manipulated (1) to be in an open syllable or a closed one with a [ʔ] coda and glottalized voice quality (referred to as “open” and “glottal,” respectively); (2) to be of the same base duration as the non-final vowels or shortened by a factor of 1.5 (“base” and “short,” respectively). The shortening factor was based on segment duration manipulation in previous AL studies [1, 6, 7] and both base and short durations were resynthesized from natural productions. The design resulted in four conditions (Figure 1). Next, participants identified the AL words in a two-alternative forced-choice test. There were 36 trials and hence 36 responses per participant. Higher response accuracy indicates better segmentation of the speech streams. Thirty native TSM listeners were recruited for each of the four conditions for a total of 120 participants.

Test response accuracy (Figure 2a) was analyzed with Bayesian mixed-effects logistic regression using uninformative priors and including Duration (short vs. base), Coda (glottal vs. open), and their interaction as fixed effects and participants and AL words as random effects. An effect was deemed statistically significant when the 95% highest density interval of its marginal posterior distribution excluded zero. The results (Figure 2b) indicated a significant Duration effect, suggesting that shortening the word-final vowels reduced segmentation performance. Neither Coda nor the interaction was significant.

The findings supported only the first prediction, aligning with the view that the use of vowel duration in segmentation is guided by the timing of syllable perceptual center. They provided further evidence that such a segmentation mechanism is grounded in perceptual salience of word boundaries (enhanced and reduced through vowel lengthening and shortening, respectively) and not simply in detection of recurrent patterns, since the consistently shortened final vowels in the AL can in theory provide useful segmentation cues. We found no support for the prediction based on TSM-specific sound distribution even when the shorter vowels were closed off by [ʔ] with glottalization. This may be attributed to the fact that the TSM tone sandhi causing the [ʔ] dropping in non-final positions also involves tonal changes, which were neutralized in our AL. An alternative possibility is that vowel glottalization in spontaneous TSM is implemented in several ways [8], which might not be all represented in our stimuli. Further work is needed to better understand the extent to which segmentation strategies are driven by language-specific duration patterns through a cross-linguistic comparison.

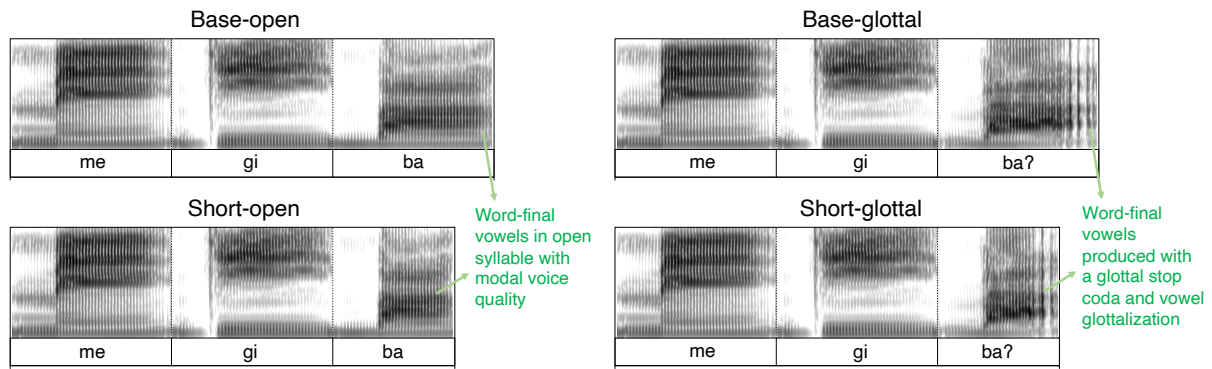


Figure 1. Broadband spectrograms of a sample artificial language word across the four experimental conditions.

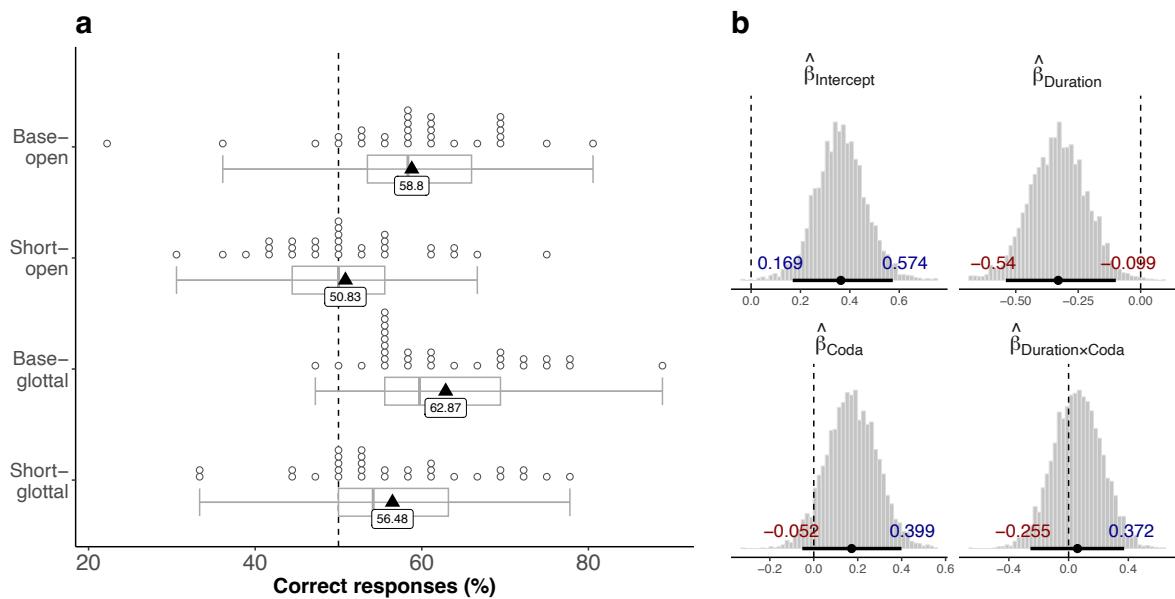


Figure 2. (a) Individual participants' response accuracy (circle) and mean accuracy (triangle) by condition. The edges of each boxplot mark the interquartile range and the whiskers indicate  $\pm 1.5$  interquartile range. Dashed vertical line represents the chance level (50%). (b) Marginal posterior distributions of the fixed-effect predictors. The interval under each distribution shows the 95% highest density interval and the mean value (circle). Dashed vertical line marks zero.

## References

- [1] M. D. Tyler and A. Cutler, "Cross-language differences in cue use for speech segmentation," *J Acoust Soc Am*, vol. 126, no. 1, pp. 367–376, 2009.
- [2] L. White, S. L. Mattys, L. Stefansdottir, and V. Jones, "Beating the bounds: Localized timing cues to word segmentation," *J Acoust Soc Am*, vol. 138, no. 2, pp. 1214–1220, 2015.
- [3] W. Zhang and Y.-A. Lu, "The role of duration in the perception of checked versus unchecked tones in Taiwanese Southern Min," in *Proceedings of the 20th International Congress of Phonetic Sciences*, R. Skarnitzl and J. Volin, Eds., Guarant International, 2023, pp. 226–230.
- [4] R.-F. Chung, *The segmental phonology of Southern Min in Taiwan*. Taipei: Crane Publishing, 1996.
- [5] J. R. Saffran, E. L. Newport, and R. N. Aslin, "Word segmentation: The role of distributional cues," *J Mem Lang*, vol. 35, no. 4, pp. 606–621, 1996.
- [6] L. White, S. Benavides-Varela, and K. Mády, "Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues?," *J Phon*, vol. 81, p. 100982, 2020.
- [7] S.-C. Ou and Z.-C. Guo, "The differential effects of vowel and onset consonant lengthening on speech segmentation: Evidence from Taiwanese Southern Min," *J Acoust Soc Am*, vol. 149, no. 3, pp. 1866–1877, Mar. 2021.
- [8] H.-H. Pan, "Glottalization of Taiwan Min checked tones," *J Int Phon Assoc*, vol. 47, no. 1, pp. 37–63, 2017.