The effect of shortening onset consonants on speech segmentation by Taiwanese Southern Min listeners

Shu-chen Ou[1], Zhe-chen Guo[2]

ID: 268

[1]Department of Foreign Languages and Literature, National Sun Yat-sen University, Taiwan, sherryou@mail.nsysu.edu.tw

[2]Department of Linguistics, University of Texas at Austin, USA, zcadamguo@utexas.edu

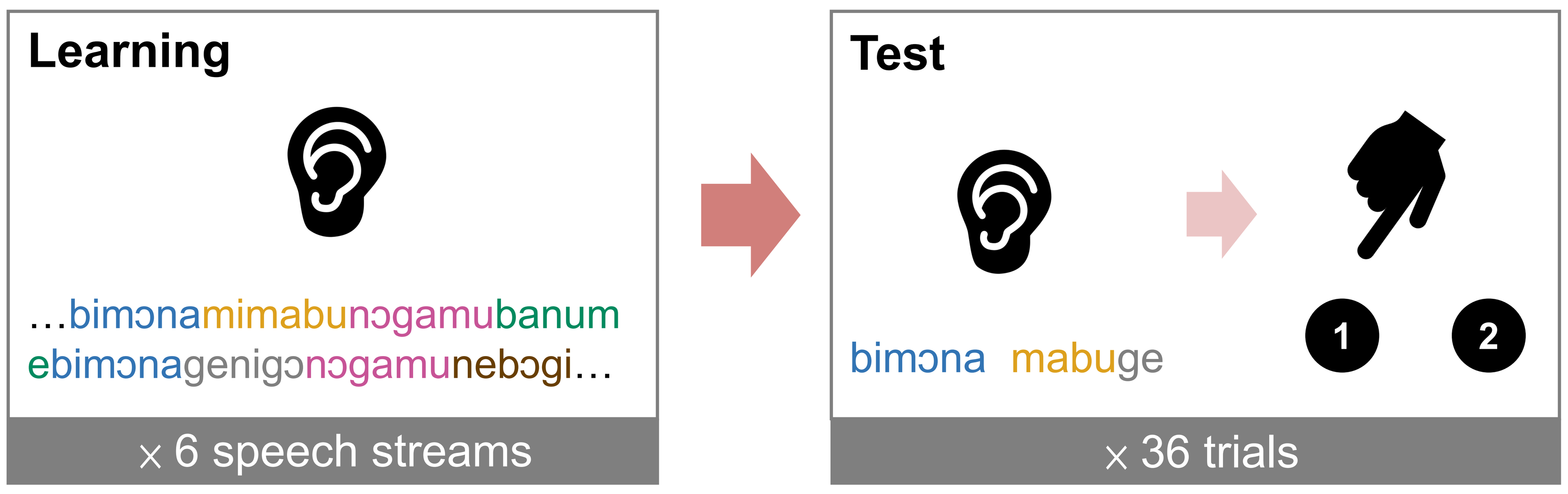
This study examined whether and how shortening voiced syllable-onset consonants impacts listeners' speech segmentation. Cross-linguistically, listeners tend to use longer onset consonants to locate word beginnings, possibly because of the greater phonetic richness and hence auditory-perceptual salience of these consonants. An alternative account, however, holds that it is the increased consonantality that guides the use of longer onsets as word beginning cues. Since shortening voiced stops and nasals enhances their consonantality, we tested the competing explanations in an artificial language learning experiment where shorter voiced onsets provided a potential segmentation cue. Taiwanese Southern Min listeners segmented the words of an artificial language through listening to continuous repetitions of the words. Shortening the voiced onsets in word-initial syllables did not significantly improve segmentation. The results thus did not support the consonantality-based explanation but were in line with the phonetic richness account. Possible mechanisms underlying the findings are discussed.

1. Introduction

- Listeners exploit various cues to segment continuous speech into individual words.
  - Longer vowels are generally interpreted as word-final [1–4].
- Recent evidence shows that listeners across languages perceive **longer syllable-onset consonants as word-initial** [5–7].
  - Longer onsets with richer phonetic information are perceptually more salient and more useful for segmentation [8].
- Longer voiceless onsets (e.g., voiceless stops with longer VOTs) are also more consonant-like [9].
- Is the cross-linguistic use of longer onsets driven by greater phonetic richness or increased consonantality?
- Testing **shortening voiced onset consonants** as a word-initial position cue**:**
  - Shorter voiced onsets (e.g., nasals) have reduced sonority and are more consonant-like [9–11].
- Artificial language (AL) learning experiment with Taiwanese Southern Min (TSM) listeners

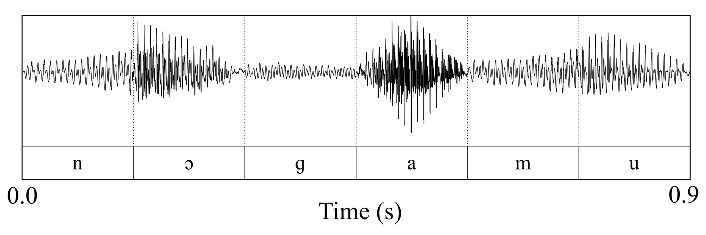
2. Experiment

Design and stimuli

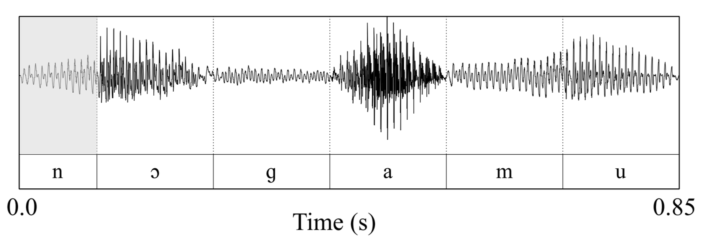


- 6 CVCVCV AL words [7]:
  - /banume/, /bimɔna/, /genigɔ/, /mimabu/, /nebɔgi/, and /nɔgamu/
  - All Cs were voiced; all Cs and Vs occurred in TSM.
  - CV syllables were recorded by a male TSM speaker and normalized for F0, duration, and amplitude.
- The AL words were randomly concatenated to form long speech streams, presented in 3 conditions:

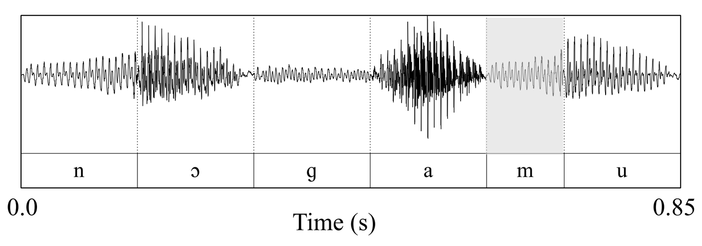
No shortening (NS)



Initial shortening (IS)



Final shortening (FS)



- The onsets were shortened by a factor of 1.5 [6, 7].
- Test: heard an AL word and a "partword." Selected the AL word.
  - Higher response accuracy indicated better segmentation

Hypotheses and predictions

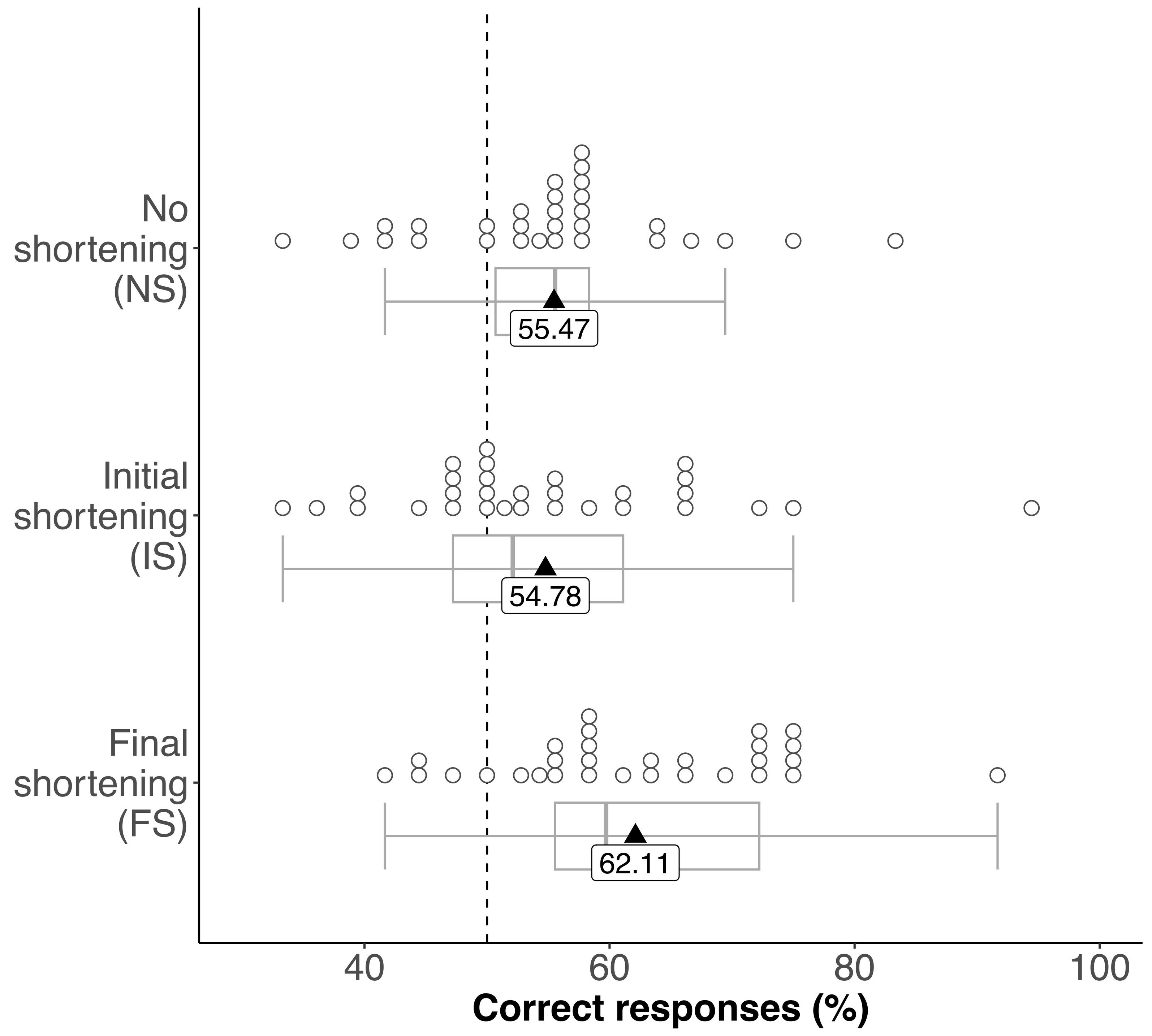
- Increased consonantality: IS (and only IS) would show significantly higher accuracy than NS.
- Phonetic richness: Neither IS or FS would show significantly higher accuracy than NS.

Participants

- 90 native speakers of TSM (*N* = 30 for each condition)

3. Results

Response accuracy of individual participants (circles) and mean accuracy of each condition (triangles). The gray bar in each box marks the median and the dashed vertical line represents the chance accuracy (50%).



- Compared response accuracy across conditions using Bayesian hierarchical logistic regression from the *brms* [12] R package.
- Neither IS (95% HDI: [–0.517, 0.350]) nor FS (95% HDI: [–0.226, 0.780]) differed significantly from NS.
- Accuracy of FS had a higher posterior probability of being above chance (96%) than that of NS (90%) and IS (77%).

4. Discussion

- No evidence that increased consonantality is perceived as word-initial, at least for TSM listeners
- The lack of shortening effects aligned with the phonetic richness hypothesis and previous work on onset lengthening [5–7].
- Longer, perceptually more salient onsets are interpreted as articulatory effort to signal important information (cf. the Effort Code Hypothesis: [13]).
- Lengthening word-initial onsets creates a delay in the timing of syllable p-centers that is interpreted as a prosodic boundary [5].
  - Shortening these onsets reduces the boundary percept.
- Highest accuracy in FS: shortening word-final onsets might cause the following vowels to be perceived as longer and word-final.
- Could onset shortening be interpreted as word-initial when accompanied by other coarticulatory cues?
  - Domain-initial nasal onsets are produced with a shorter duration and reduced coarticulatory vowel nasalization [10].

References

[1] Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. [2] Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America*, *126*(1), 367–376. [3] Ordin, M., Polyanskaya, L., Laka, I., & Nespor, M. (2017). Cross-linguistic differences in the use of durational cues for the segmentation of a novel language. *Memory and Cognition*, *45*(5), 863–876. [4] Hay, J. S. F., & Diehl, R. L. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception and Psychophysics*, *69*(1), 113–122. [5] White, L., Mattys, S. L., Stefansdottir, L., & Jones, V. (2015). Beating the bounds: Localized timing cues to word segmentation. *The Journal of the Acoustical Society of America*, *138*(2), 1214–1220. [6] White, L., Benavides-Varela, S., & Mády, K. (2020). Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues? *Journal of Phonetics*, *81*, 100982. [7] Ou, S.-C., & Guo, Z.-C. (2021). The differential effects of vowel and onset consonant lengthening on speech segmentation: Evidence from Taiwanese Southern Min. *The Journal of the Acoustical Society of America*, *149*(3), 1866–1877. [8] Kim, S., Cho, T., & McQueen, J. M. (2012). Phonetic richness can outweigh prosodically-driven phonological knowledge when learning words in an artificial language. *Journal of Phonetics*, *40*(3), 443–452. [9] Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*(4), 466–485. [10] Cho, T., Kim, D., & Kim, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics*, *64*, 71–89. [11] Hsu, C.-S. K., & Jun, S.-A. (1998). Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics*, *96*, 69–89. [12] Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. [13] Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. *Proceedings of Speech Prosody 2002*, 47–57.

Acknowledgment

We would like to thank Kai-ting Chuang and Yun-jen Wu for assistance with data collection. The research is supported by a grant to the first author from National Science and Technology Council, Taiwan (MOST109-2410-H-110-069-MY3).