

Evidence for timing prediction from a novel nonword segmentation task

Laurence White¹, Wei Li¹, Sarah Knight¹, Sven Mattys², Robert Lennon¹, Julie Morris¹

¹Newcastle University, ²University of York

Word onset consonants are lengthened in many languages [1], and listeners use this cue to identify spoken word boundaries [2]. Such exploitation of timing cues must entail predictive processing: thus, the interpretation of consonant lengthening, as with listeners' perception of reduced syllables and lexical stress [3, 4], must be influenced by durational expectations based on foregoing speech rate.

We investigated how preceding utterance timing affects the use of localised timing cues in a series of studies using a novel nonword segmentation task. In Studies 1-4, listeners heard diphone-synthesised 12-syllable nonsense utterances (e.g., *dumipakolibekubinudafolu*) containing trisyllabic nonword targets (e.g., *libeku*) in **Early**, **Medial**, or **Late** positions. In the **flat** baseline timing condition, all utterance segments were 120ms. In the **c1** condition, the target-initial consonant was lengthened (170ms), providing a target-congruent boundary cue; in **c2**, the second-syllable onset was lengthened, creating an incongruent cue. Target detection accuracy was analysed using generalised linear mixed-effects regression models with fixed effects of Timing Cue and Utterance Position.

In Studies 1-2 (lab-based), listeners responded "present" or "absent" to target probes heard *after* utterances. **Early** targets were detected at chance, likely due to phonological memory limitations. **Medial** target detection exceeded chance, but with no effect of timing cues. Differential timing effects emerged with **Late** targets, however: e.g., in Study 1, **c1** improved target detection relative to other conditions. The emergence of timing sensitivity for **Late** targets suggests that cues are only detected after *sufficient of the foregoing utterance* is heard to allow listeners to generate temporal predictions.

Studies 3-4 (online) presented target probes *before* utterances, with detection rates predictably higher. Timing × Utterance interactions remained, potentially complicated by between-trial target repetition. Study 5 (online) reinstated *post-utterance* probes, removed target repetitions, added three utterance-initial buffer syllables, and included trial-by-trial ("correct"/"incorrect") feedback for some participants. The feedback condition participants (Fig. 1; N = 109) showed reliable **c1** cue use (across a range of onset consonants) for **Medial** targets, where – given the buffer syllables – the lengthened word onset occurs as far from the utterance onset as in the **Late** condition in Studies 1-4. This further demonstrates that temporal predictions arise after accumulated utterance exposure. Ongoing work examines how stimulus naturalness, task demands, and individual differences modulate this effect.

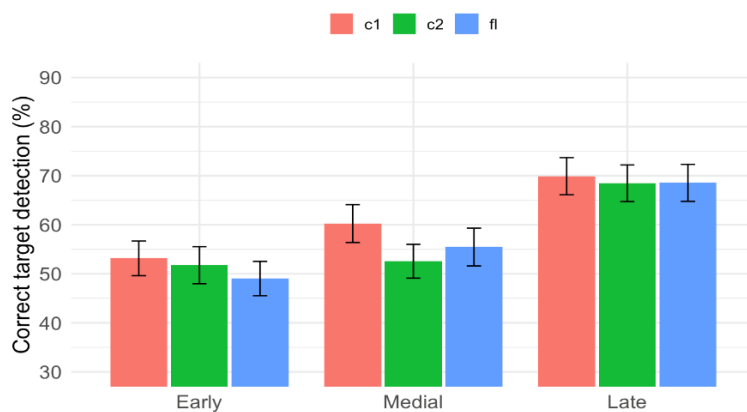


Figure 1: Study 5, feedback version: nonword segmentation accuracy according to Timing Cue and Utterance Position of targets. (Note also the unexpected lack of effect of timing cues for Late targets.)

[1] Blum, F., Paschen, L., Forkel, R., Fuchs, S., & Seifart, F. (2024). Consonant lengthening marks the beginning of words across a diverse sample of languages. *Nature Human Behaviour*, 8, 2127-2138.

[2] 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.

[3] Dilley, L.C., & Pitt, M.A. (2010). Altering context speech rate can cause words to appear or disappear. *Psychological Science*, 21, 1664-1670.

[4] Reinisch, E., Jesse, A., & McQueen, J.M. (2011). Speaking rate affects the perception of duration as a suprasegmental lexical-stress cue. *Language and Speech*, 54, 147-165.