

Timing variability at different parts of movement: Tongue body articulation of *yeah* in the Doubletalk corpus

Alice Turk¹, Yuqi Qin¹, Sonja Schaeffler², James M. Scobbie² & Benjamin Elie¹

¹University of Edinburgh (UK), ²Queen Margaret University (UK)

Findings in the non-speech motor control literature suggest that parts of movements most closely related to goal-achievement have higher temporal precision relative to a reference time point (e.g. Shaffer, 1982; Semjen, 1992; Billon et al., 1996; Bootsma & van Wieringen, 1990; Craig et al., 2005, Gentner et al., 1980). Perkell and Matthies (1992) and Leonard and Cummins (2010) provide similar findings for speech and speech-accompanying gesture, respectively. As discussed in Turk and Shattuck-Hufnagel (2020), findings such as these suggest that particular parts of movement are prioritised for temporal accuracy (Todorov & Jordan 2002, 2003), and findings for speech can be explained in models where abstract symbolic phonological representations map onto parts of movement most closely related to them. However, these findings challenge models of speech production with spatiotemporal representations and/or goals because these models lack an explanation for reduced temporal variability at a particular part of movement. In other words, in spatiotemporal models, when a movement starts determines the timing of all other parts of movement; any variability in onset time should be propagated to the rest of movement. In contrast, in models where a particular part of movement can be prioritised for temporal accuracy, lower temporal variability can occur at the prioritised point, e.g. at a goal-related part of movement.

In this paper, we investigated temporal variability at different parts of tongue body movement for productions of spontaneous post-pausal productions of *yeah* from the Edinburgh Speech Production Facility *Doubletalk* corpus. These data consisted of temporal measurements from 12 speakers' tongue body raising movements for /j/, relative to the onset of voicing for *yeah*, shown in Figure 1. Relative timings of 1) onset tangential velocity minimum (Onset), 2) acceleration maximum (AccMax), 3) peak tangential velocity (TVMax), 4) deceleration minimum (DecMin), and 5) offset tangential velocity minimum (DecMin) were measured from each trajectory.

Preliminary results of the 300+ tokens in the dataset without long (quasi)-steady states show that temporal variability reduces from the onset of movement towards the offset (Figure 2). These results support the view that the mental representation for /j/ maps onto the part of tongue body movement most closely related to constriction formation, and that this part of movement is prioritised for accurate temporal coordination with the onset of voicing. These results provide support for symbolic phonological representations mapped onto goal-related part(s) of movement.

References

- Bootsma, R. J., & van Wieringen, P. C. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance*, 16(1), 21-29.
- Craig, C., Pepping, G.-J., & Greal, M. (2005). Intercepting beats in predesignated target zones. *Experimental Brain Research*, 165(4), 490-504.
- Gentner, D. R., Grudin, J., & Conway, E. (1980). Finger movements in transcription typing. *Center for Human Information Processing, University of California, San Diego, Technical Report 8001*, 1-8.
- Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, 26(10), 1457-1471.
- Perkell, J. S., & Matthies, M. L. (1992). Temporal measures of anticipatory labial coarticulation for the vowel /u/ - within-subject and cross-subject variability. *Journal of the Acoustical Society of America*, 91(5), 2911-2925.
- Scobbie, J. M., Turk, A., Geng, C., King, S., Lickley, R., & Richmond, K. (2013). The Edinburgh Speech Production Facility Doubletalk corpus. *Interspeech 2013*, 764-766.
- Semjen, A. (1992). Determinants of timing in serial movements. In F. Macar, V. Pouthas, & W. J. Friedman (Eds.), *Time, Action and Cognition* (Vol. 66). Springer.
- Shaffer, L. H. (1982). Rhythm and timing in skill. *Psychological Review*, 89(2), 109-122.

Todorov, E., & Jordan, M. I. (2003). A minimal intervention principle for coordinated movement. In *Advances in neural information processing systems* (pp. 27-34).

Turk, A., & Shattuck-Hufnagel, S. (2020). *Speech Timing: Implications for Theories of Phonology, Phonetics, and Speech Motor Control*. Oxford University Press.

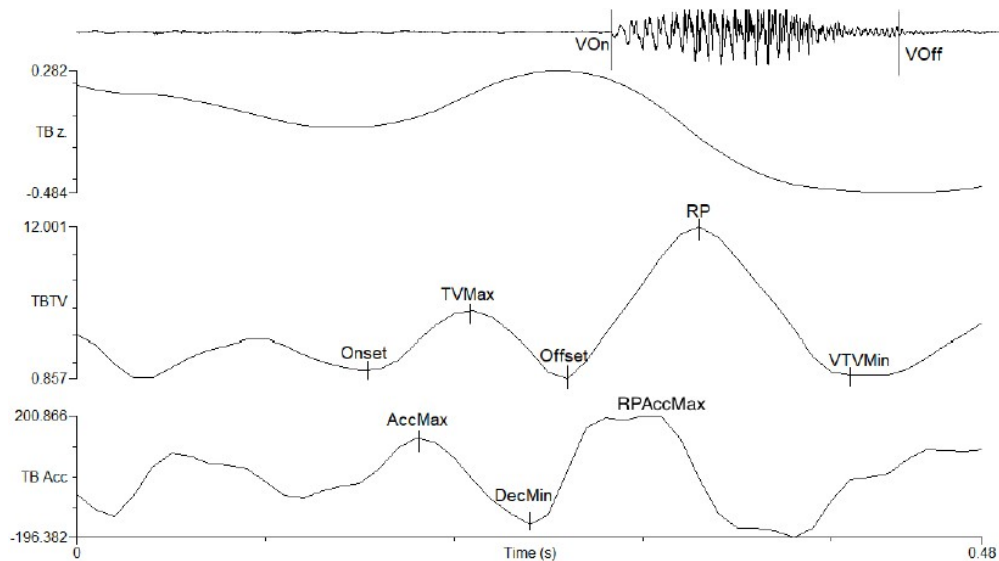


Figure 1. Token of *yeah* illustrating the measurement timepoints used for the study, as described in the text. The first trace shows the acoustic waveform of *yeah* from Voicing Onset (VOn) to Voicing Offset (VOC), time-aligned with articulatory movements that occur before, during, and after the acoustic signature of the word in the traces underneath the top waveform. The second trace (TBz) represents the vertical position of the tongue body sensor (cm) above and below the occlusal plane. The third trace (TBTv) represents the tangential velocity of the TB sensor (cm/s). The fourth trace (TB Acc) represents its acceleration (positive values) and deceleration (cm/s²).

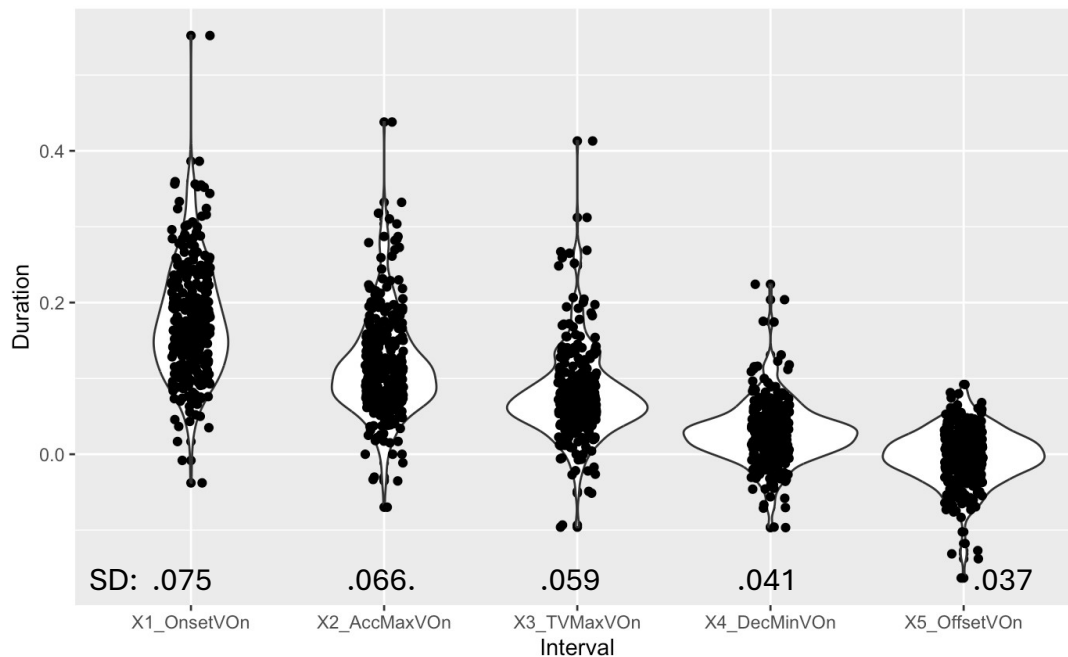


Figure 2. Violin plots with overlaid scatterplots of intervals defined from a measured timepoint in the post-pausal tongue body raising movement for /j/, relative to the time of Voicing Onset.