Temporal flexibility of articulation within syllables

Mark Tiede[1] and Sam Tilsen[2]

[1]Yale Brain Function Laboratory, Yale University (USA), [2]Cornell University (USA)

Speakers are able to adjust their rate of speech to signal a wide range of linguistic and paralinguistic information, but exactly how they accomplish rate control is an open question. This study investigates how the temporal adjustability of articulation depends on syllable-internal structure. Participants in an articulography experiment were tasked with producing the same utterance across a wide range of speech rates, from very slow to very fast. In addition to standard analyses of timing, temporal variability was measured using dynamic time warping (DTW). The analyses reveal several ways in syllable structure constrains temporal flexibility.

Method. Acoustic and electromagnetic articulograph (EMA) signals were recorded while speakers produced a carrier phrase (*I will be ___ on Monday*) with one of two target words: {*pop*, *popped*}. A visual scale was used to iconically cue speech rate, with speakers being instructed that the scale bounds should correspond to their slowest/fastest intelligible rates. (The purpose of this manipulation was to elicit a wide range of rates, not a specific rate.) A location on the scale was presented at the start of each trial. Three speakers of English each produced a total of 352 trials, grouped into 8 blocks in which the target word remained constant. Rate cues increased from slowest to fastest in each block. GAM analyses of rate effects use the empirically produced rate, measured as the duration of the carrier portion of the utterance. DTW of EMA sensor velocities was conducted pairwise over all trials from the same speaker/target utterance, and local rate variability was calculated from the geometric standard deviations of the slopes of DTW warping functions for each trial.

Results. Articulatory processes associated with achieving vocalic articulatory postures were more adjustable than those associated with achieving consonantal constrictions. Fig. 1 label (i) illustrates this effect: local rate variability peaks in the vicinity of the vocalic target, in the middle of the acoustic vowel interval. Further support of this is seen in the ranges of variation predicted by GAM models in Fig 2; specifically, the ranges of predicted acoustic vowel duration (top row: a1 dur) and vocalic target to consonantal onset (middle row: a1 trg – p2 ons) were larger than those associated with other measures. Crucially, the local relative rate variabilities measured at the onset and coda constriction targets (bottom row: p1 trg, p2 trg) were lower than variabilities associated with the vocalic target landmark (a1 trg).

Furthermore, the presence of an additional lingual constriction associated with the [t] of *popped* constrains the extent of rate-driven adjustment. Fig. 1 label (ii) illustrates this: in the vicinity of the plateau and release of the first post-vocalic constriction, articulatory process rate variability is substantially lower when the additional post-vocalic constriction of [t] is present (orange line). Further support of this finding can be seen in predicted differences from GAM models shown in Fig. 2: at slow rates, acoustic vowel duration (top row: a1 dur) is greater in *pop* than *popped*; in addition, local rate variability is greater at the post-vocalic bilabial constriction target (bottom row: p2 trg) for *pop* than *popped*, especially with moderate to fast rates.

Conclusion. Overall the findings show that onset and coda articulations are less flexible than vocalic ones when speakers intentionally vary rate, and that the presence of an additional post-vocalic constriction (in *popped*) further limits this flexibility. These findings show that even within a syllable, speakers do not implement rate-related adjustments uniformly. Sub-syllable-scale influences on rate adjustment are important for our understanding of timing because they reveal the organizational structure of articulation, and because they influence where larger scale adjustments are most likely to be observed.

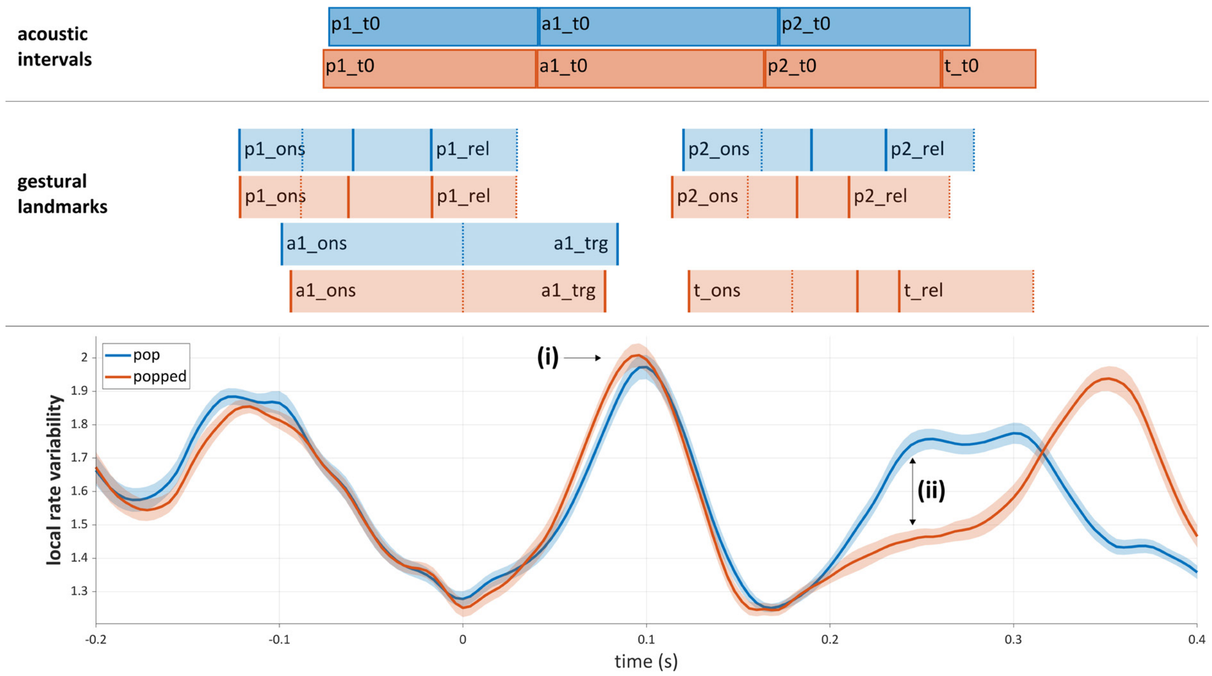

Fig. 1. Summary of data from one participant. Top: average acoustic intervals and gestural landmark intervals for *pop* and *popped*. Bottom: geometric standard deviation of local relative rate; larger values indicate greater adjustability of articulatory process rates. All signals, intervals, and landmarks are aligned to the time of maximum tongue body speed associated with formation of the vowel.

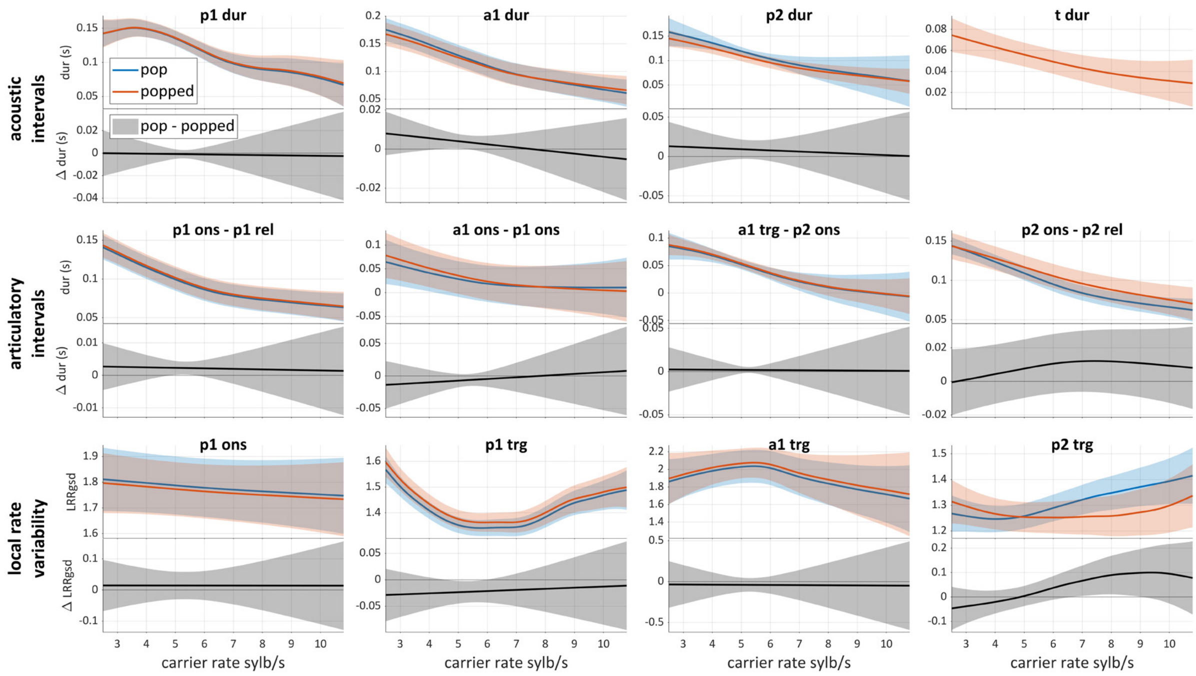

Fig. 2. Effects of speech rate on selected variables. Each pair of panels shows GAM model predictions and confidence intervals (top) and predicted differences (bottom). Top row: acoustic intervals. Middle row: articulatory intervals. Bottom row: local relative rate variability (geometric standard deviation) at articulatory landmarks.