Intrasegmental Gestural Timing for American English /ɹ/ in Isolated and Connected Speech

Sarah Harper[1]
[1]*University of Southern California*

Recent research on intraspeaker variability in speech production has observed that many acoustic and articulatory attributes of speech vary systematically across different speech tasks, presumably as a consequence of greater temporal constraints on articulatory movement in faster, more casual speech [1]. One domain in which the effects of these temporal constraints on articulation are observed is in the degree of temporal overlap between discrete speech units (such as articulatory gestures), with the amount of overlap exhibited by adjacent gestures increasing at faster speech rates [e.g., 2]. However, as most existing research examining task and rate effects on gestural timing has focused on intersegmental timing relationships, the extent to which similar contextual effects are observed in intrasegmental gestural timing for multi-gesture consonants is unknown.

This study expands on previous work in both intergestural timing and contextual variability in speech production by examining the effect of speech task on the relative timing of two lingual gestures involved in the production of American English /ɹ/. American English /ɹ/ is widely characterized as involving the coordination of three supralaryngeal gestures, two of which are lingual: a palatal constriction made with either the tongue tip or body, and a pharyngeal constriction made with the tongue root [3]. The sequencing of these gestures has been shown to differ depending on the position of /ɹ/ in the word: the palatal gesture precedes the pharyngeal gesture in syllable-initial position and follows the pharyngeal gesture in syllable-final position. This positional difference in gestural sequencing has been shown to parallel observed differences in gestural magnitude for these constrictions across syllabic positions, leading to proposals that these magnitude asymmetries may explain the similarly asymmetric timing pattern [4]. As recent findings demonstrate systematic variation in the magnitude of both the palatal and pharyngeal gestures in /ɹ/ across speech tasks [5], the examination of speech task effects on timing enables a more thorough investigation of the relationship between gestural timing and magnitude, and enhances our understanding of the scope of gradience and variability in intrasegmental timing, an issue of importance in due to the role of articulatory variation in sound change [6].

Articulatory data for this study was taken from recordings of 40 speakers in the Wisconsin x-ray Microbeam (XRMB) database [7]. A total of 3,468 tokens of word-initial and word-final /ɹ/ were taken from three of the experimental tasks included in the corpus and separated into two conditions for analysis: (1) an **Isolated Speech** condition containing data from a Citation Word task (990 tokens) and (2) a **Connected Speech** condition containing data from Sentence and Prose Passage reading tasks (2,478 tokens). For each token of /ɹ/, maximum constriction time was measured for both the palatal and pharyngeal constrictions, with the gestural timing lag (**MLag**) calculated as the difference between the two maximum constriction times for a given token. The maximum palatal constriction degree and pharyngeal retraction were also measured to gauge the magnitude of each gesture.

The analysis of the MLag measurement indicates that although gestural sequencing remains constant between the Isolated and Connected Speech conditions, with the expected positional timing asymmetry observed in both conditions, the timing of these two gestures differs significantly across conditions. Specifically, less extreme timing differences are observed in the Connected Speech condition than in the Isolated Speech condition, mirroring the pattern of generally reduced gestural magnitude in the Connected Speech condition and reflecting an observed relationship between gestural magnitude and timing in this data.

[1] Moon, S.-J., & Lindblom, B. 1994. Interaction between duration, context and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, 96, 40-55.
[2] Browman, C.P., & Goldstein, L. 1984. Tiers in Articulatory Phonology, with Some Implications for Casual Speech.
[3] Alwan, A., Narayanan, S., & Haker, K. 1997. Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II: The rhotics. *Journal of the Acoustical Society of America*, 101, 1078-1089.
[4] Campbell, F., Gick, B., Wilson, I., & Vatikiotis-Bateson, E. 2010. Spatial and Temporal Properties of Gestures in North American English /R/. *Language and Speech*, 53, 49-69.
[5] Harper, S., Goldstein, L., & Narayanan, S. 2017. Stylistic Effects on the Acoustic and Articulatory Properties of English Rhotics. Poster presented at PaPE 2017, Cologne, Germany.
[6] Lawson, E., Stuart-Smith, J., and Scobbie, J.M. 2018. The role of gesture delay in coda /r/ weakening: An articulatory, auditory and acoustic study. *Journal of the Acoustical Society of America*, 143(3), 1646-1657.
[7] Westbury, J. 1994. *X-ray Microbeam Speech Production Database User's Handbook*. University of Wisconsin, Madison, WI.