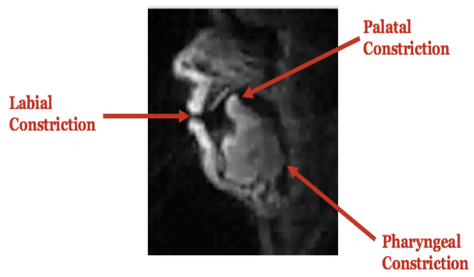
Stylistic Effects on the Acoustic and Articulatory Properties of English Rhotics

Sarah Harper, Louis Goldstein, Shrikanth Narayanan
University of Southern California

TASK EFFECTS IN SPEECH PRODUCTION

- Acoustic and articulatory aspects of speech production have been observed to vary systematically across different speech styles, with "undershoot" effects frequently observed in more spontaneous or connected speech.
  - Reduced amplitude of tongue raising and lowering gestures in connected speech [1].
  - Increase in acoustic vowel space, dynamic formant characteristics, and vowel duration in clear speech [2].
- For speech sounds formed through the interaction of multiple supralaryngeal constriction gestures, such as American English /ɹ/, unclear how the magnitude of each gesture is affected by speech task.
- Complexity of articulatory-acoustic relations in /ɹ/ production also introduces possibility of non-linear relationships between articulatory and acoustic task effects.



RESEARCH QUESTIONS

1. What articulatory effect(s) of speech task are observed in the production of the palatal and pharyngeal constriction gestures for American English /ɹ/?
2. How do other factors that may affect gestural realization, such as a segment's position in the syllable, interact with the observed task effects?
3. How do task effects on articulation affect the acoustic outcome in /ɹ/ production, specifically its characteristically low F3 value?

PARTICIPANTS AND METHODS

- Articulatory data taken from real-time MRI recordings of three native speakers of American English (2 male, 1 female)
  - Acoustic data from simultaneous audio recordings of each MRI session
- Participants recorded reading single words in isolation (***Citation Condition***) and sentences (***Connected Condition***)
  - A total of 501 tokens of /ɹ/ analyzed (180 from Citation Condition, 321 from Connected Condition
- Syllabic position and surrounding vowel context of the target consonant controlled for all words in the Citation Condition.

Syllabic Positions	Simple Onset (#ɹ_)	Complex Onset (#Bɹ_)	Simple Coda (_ɹ#)	Complex Coda (_ɹB#)
Vowel Contexts	**/i/**	**/ɑ/**	**/o/**	**/u/**	**/ə/**

Table 1: Positions and Vowel Contexts from which tokens were taken in both the Citation and Connected Conditions

- Connected Condition tokens taken from words in the USC-TIMIT corpus [3] that contained /ɹ/ in the same contexts found in the Citation Condition.

ARTICULATORY AND ACOUSTIC ANALYSIS

Articulatory Analysis

- Time of maximum constriction for each gesture found using a Region of Interest technique [4].
  - Four pseudo-circular regions (radius of 3 pixels) manually defined along vocal tract midline (*Figure 1*).
    - Regions used in analysis: Labial region, Anterior and Posterior Palatal regions, and Pharyngeal region.
  - Average pixel intensity inside each region calculated for each frame (higher pixel intensity = more tissue within the region).
  - Time of maximum constriction found using the find_gest algorithm [5], and defined as the movement velocity minimum closest to a manually identified pixel intensity maximum.

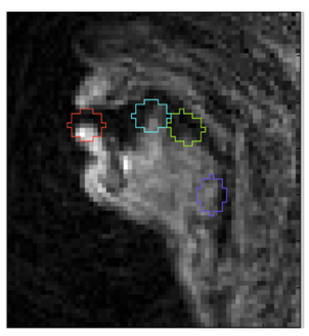


Figure 1: Vocal tract with regions overlaid (Red = Labial, Blue = Anterior Palate, Green = Posterior Palate, Purple = Pharynx)

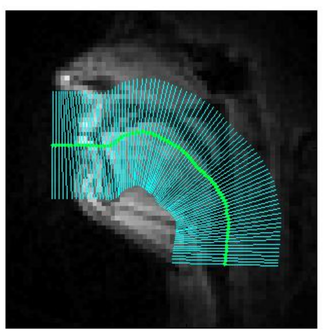


Figure 2: Example of the grid lines used for the segmentation analysis

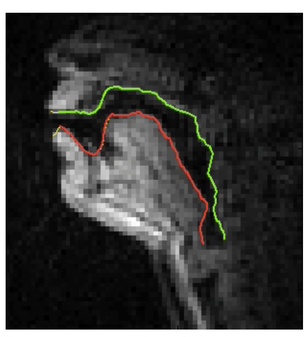


Figure 3: Example of the upper and lower air-tissue boundaries constructed by the segmentation algorithm

- Degree of articulator aperture measured at the time of maximal constriction for each constriction location using air-tissue boundary segmentation [5].
- Aperture measured as the smallest distance between upper and lower air-tissue boundaries within a pre-determined span of grid lines (*Figures 2 and 3*).

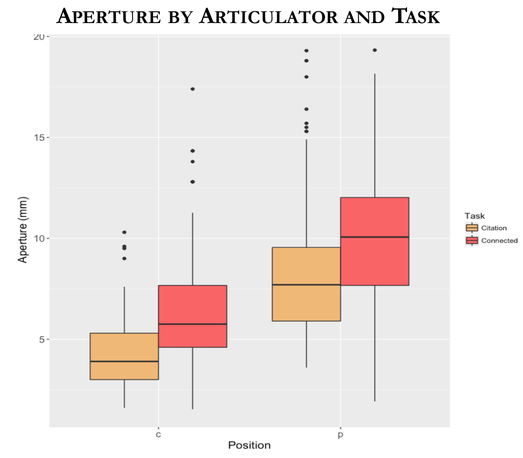
Acoustic Analysis

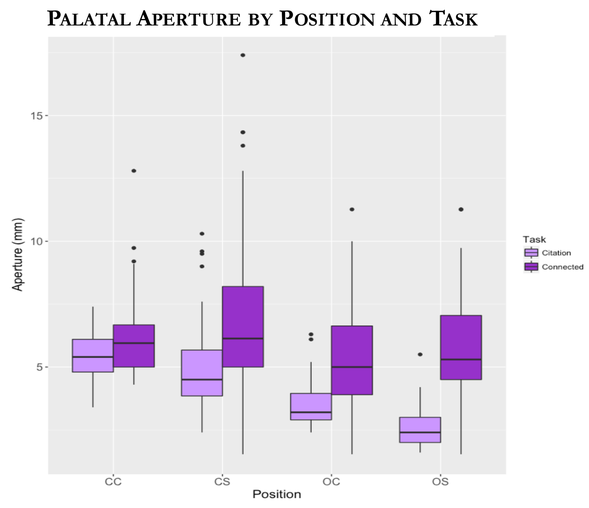
- First five formants automatically extracted in Praat at the time of maximal constriction of each constriction gesture.
  - F3 selected for analysis due to the importance of low F3 as an acoustic feature of /ɹ/.

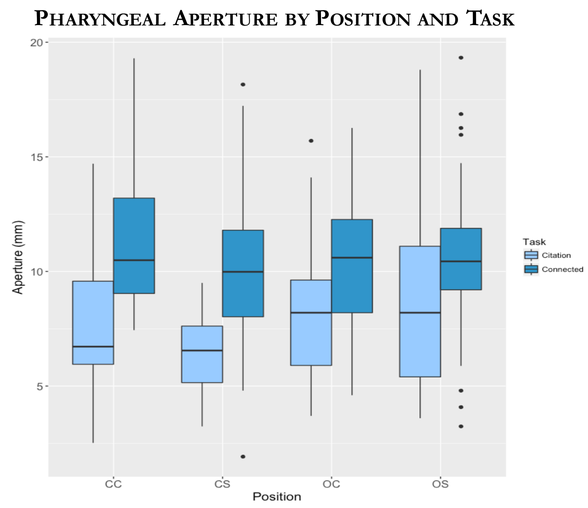
Supported by NIH grant DC007124

RESULTS: TASK EFFECTS ON ARTICULATION

- Linear Mixed Effects Models were fit to the data in R [7] to examine both task effects on articulation and articulatory-acoustic relations within the two tasks.
- Palatal aperture values significantly smaller than Pharyngeal aperture values overall ($p < 0.0001$).
- Significant main effect of Task ($p < 0.0001$) indicates that smaller average aperture values are observed in the Citation Condition than in the Connected Condition.
  - Interaction between Articulator and Task not significant at the .05 level ($p = 0.09$).



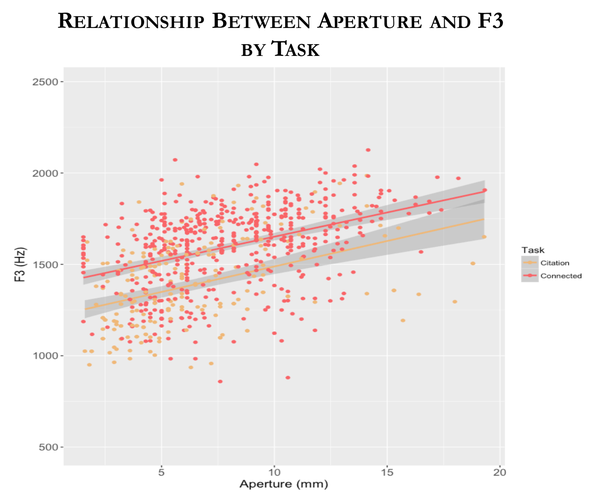




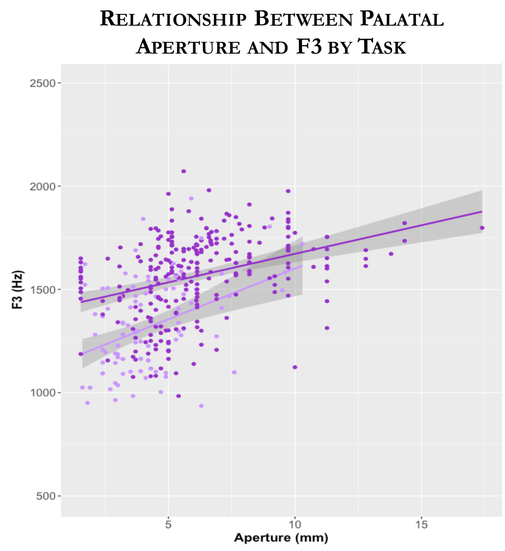
- Significant interaction of Task and Position was found for palatal aperture in the Simple Onset and Complex Onset conditions (OS: $p < 0.001$, OC: $p < 0.01$), while significant interactions were observed in the Simple and Complex Coda positions for pharyngeal aperture (CS: $p < 0.001$, CC: $p < 0.05$).
  - For palatal aperture, the effect of Task was significantly greater for the Simple Onset condition than the Complex Onset condition ($b = 0.55$, $p = 0.04$).
  - No significant difference in the effect of Task on pharyngeal aperture for the two Coda conditions ($b = 0.61$, $p = 0.09$)

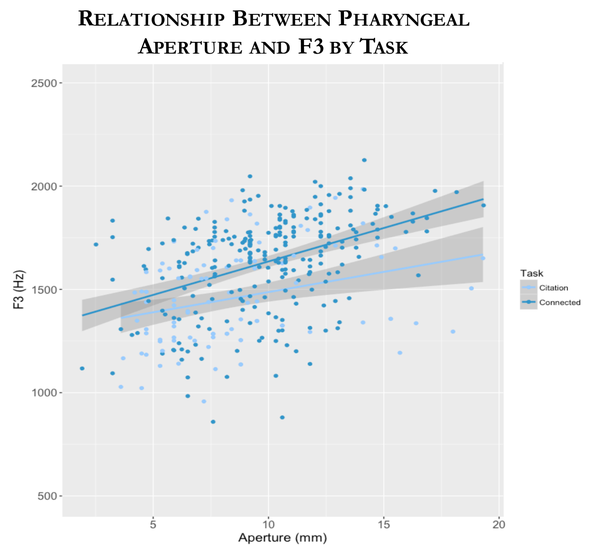
RESULTS: ARTICULATORY-ACOUSTIC RELATIONS

- Statistically significant main effect of Task on F3 ($p < 0.0001$) indicates average F3 values were consistently larger in the Connected Condition than the Citation Condition for both palatal and pharyngeal gestures.
- Significant effects of Aperture ($p < 0.0001$) and of Articulator ($p < 0.05$) also observed.
  - Lack of an interaction between Aperture and Articulator suggests that the significant effect of Articulator is due to overall smaller aperture values for palatal gesture.



- Observed interaction between Articulator and Task ($b = -75.2$, $p < 0.05$) suggests that the effect of Task on F3 is smaller for the pharyngeal gesture than for the palatal gesture.





SUMMARY AND IMPLICATIONS

- Effects of speech task observed for both Palatal and Pharyngeal aperture measurements, with a greater degree of constriction (smaller aperture values) observed in the Citation Condition for both constriction gestures.
- Nature of the interaction between position and task effects differs between the palatal and pharyngeal gestures, possibly reflecting differences between primary and secondary articulations.
  - Palatal aperture demonstrates a greater effect of task in onset position, with the greatest effect observed for simple onsets.
  - Task effects only seen for pharyngeal aperture values in coda positions.
- Difference in average F3 values across the two speech tasks reflects observed articulatory differences.
  - Lower F3 values observed in Citation Condition, which follows from the greater degree of constriction observed for both constriction gestures in this condition.