Articulatory-Acoustic and Kinematic Relationships in Comfortable and Loud Speech

Jason A. Whitfield[1], Ph.D., Christopher Dromey[2], Ph.D., & Panika Palmer[2], M.S.

[1]Communication Sciences & Disorders, Bowling Green State University, [2]Communication Disorders, Brigham Young University

INTRODUCTION

During speech, movements of the vocal tract systematically alter vowel acoustics. A number of investigations have examined these articulatory-acoustic relationships (e.g., Mefferd & Green, 2010). Studies have shown that systematic changes in speaking style, such as in clear or loud speech, yield greater excursions of articulatory gestures that are accompanied by an expansion of formant frequency measures in F1-F2 space (Mefferd & Green, 2010; Tasko & Greilick, 2010). A recent investigation introduced the Articulatory-Acoustic Vowel Space (AAVS), a measure of working formant space that uses continuously sampled formant data from connected speech (Whitfield & Goberman, 2014). No data, however, that describe the relationship between the AAVS and articulatory kinematics are available. The purpose of this work was to examine the relationship between the AAVS and its kinematic equivalent.

METHODS

Participant Characteristics and Data Collection Protocols

Twenty participants, 10 males (*Mean age*= 24.8 yrs; *SD*=2.4; *Range*:22-29) and 10 females (*Mean age*=25.1 yrs; *SD*=4.0; *Range*:20-34) volunteered as speakers for the current study. Sensor coils from an electromagnetic articulograph (NDI Wave) were attached near the tongue tip, on the mid-tongue, and to the lower incisors and the upper and lower lips at midline (Fig. 1). The participants produced two target utterances three times each (*It's time to shop for two new suits* and *A good AC should keep your car cool*) under several speaking conditions that included comfortable and loud speech. There was no recording from the tongue-front marker for Subject M3. Therefore, his data were excluded from the analysis. Kinematic activity was recorded at 400 Hz and the audio signal was sampled at 22050 Hz.

Acoustic Analysis: Articulatory-Acoustic Vowel Space (AAVS)

The audio samples were analyzed offline to obtain the AAVS for each sample. Formant trajectories from each recorded sample were extracted using PRAAT and exported to MATLAB for further processing. A custom MATLAB script was used to identify and remove outlying LPC data. The script then applied a Butterworth filter to the data and the formant trajectories were filtered using a 10 Hz low-pass filter. The AAVS was then calculated from these filtered formant trajectories. The AAVS is the bivariate variability of F1-F2 data sampled continuously during connected speech and is calculated as the square root of the generalized variance of the F1-F2 data (Whitfield & Goberman, 2014).

Kinematic Analysis: Articulatory-Kinematic Vowel Space (AKVS)

The x-y positions of each sensor coil, time-aligned with the audio signal, were imported into custom MATLAB routines for analysis. For the articulatory data, the AAVS calculation was applied to the x-y kinematic data from the tongue-front (TF) and tongue-middle (TM) markers, and the resulting metric is referred to here as the Articulatory-Kinematic Vowel Space (AKVS).

Statistics

First, linear mixed model (LMM) analyses were conducted to examine the effect of stimulus and loudness conditions on the variables of interest, namely the AAVS and the AKVS for the TF and TM markers. To follow, within-participant univariate correlations between the AAVS and AKVS measures were examined. To quantify these regression effects at a group level, while still taking into account the individual variability in regression equations between each participant, the relationship between the AAVS and the AKVS measures were examined using LMM regression analyses. For this analysis, the AAVS was used as a fixed covariate factor and Sex was used as a between-subjects fixed factor. The AAVS and participant were included as a random factors to allow the slope and intercept to vary between participants. In addition to the fixed effects, the *marginal* r^2 (r^2_m), which quantifies the proportion of the variance explained by the fixed factors (i.e., the AAVS and sex), and *conditional* r^2 (r^2_c), which calculates the proportion of the variance explained by both the fixed and random factors, were calculated.

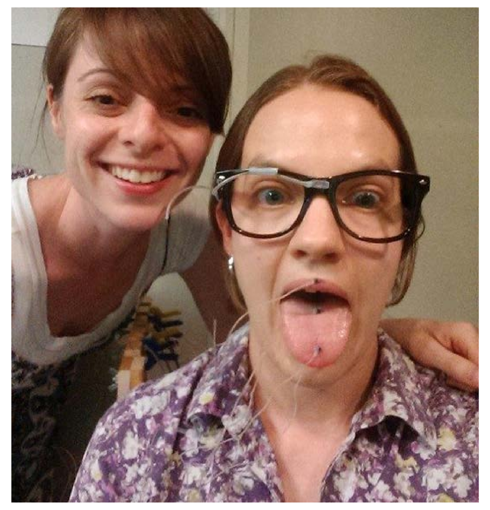

Figure 1. Kinematic recordings were made using the NDI wave system. The picture shows an example the locations of the sensor coils. This investigation examined movement records for the tongue-front marker, located near the tongue tip, and the tongue-middle marker, located near the dorsum of the tongue.

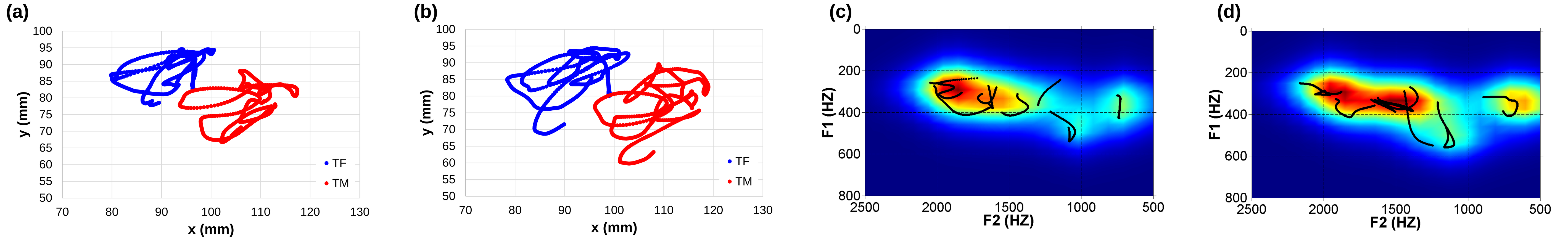

Figure 2. Kinematic record (a, b) for the tongue-front (TF) and tongue-middle (TM) markers and the corresponding Formant Trajectory Trace (FTT; c, d) for the comfortable (a, c) and loud (b, d) speaking conditions, showing the relationship between the x-y lingual kinematics and the F1-F2 trajectories for the utterance "*A good AC should keep your car cool*" spoken by a male participant.

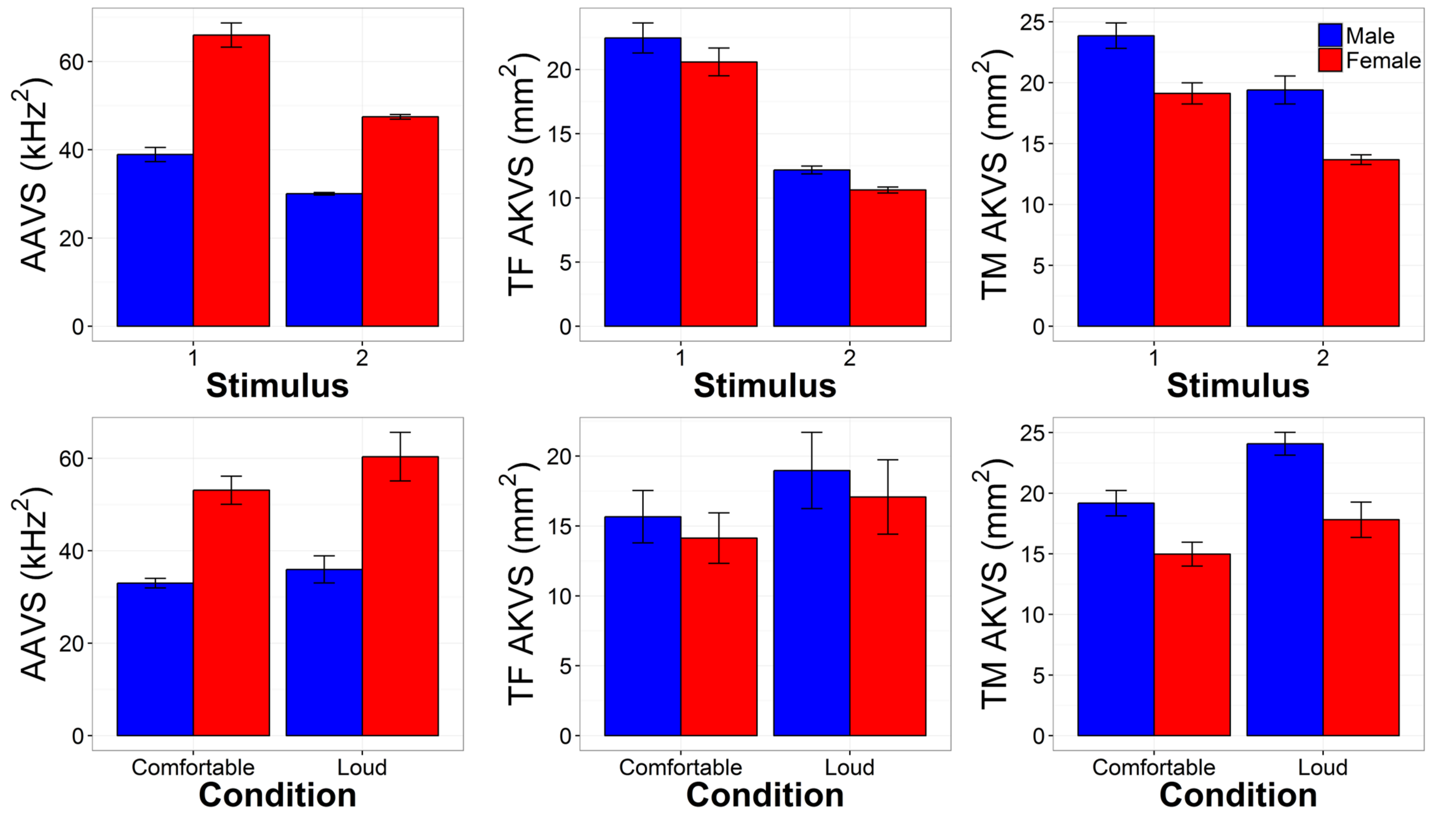

Figure 2. Univariate effects of stimulus (top panes) and condition (bottom panes) for the Articulatory-Acoustic Vowel Space (AAVS; left panes), Articulatory-Kinematic Vowel Space (AKVS) for the tongue-front (TF) marker (middle panes) and tongue-mid (TM) marker (right panes). Note: Stimulus 1 was "*A good AC should keep your car cool*" and Stimulus 2 was "*It's time to shop for two new suits*".

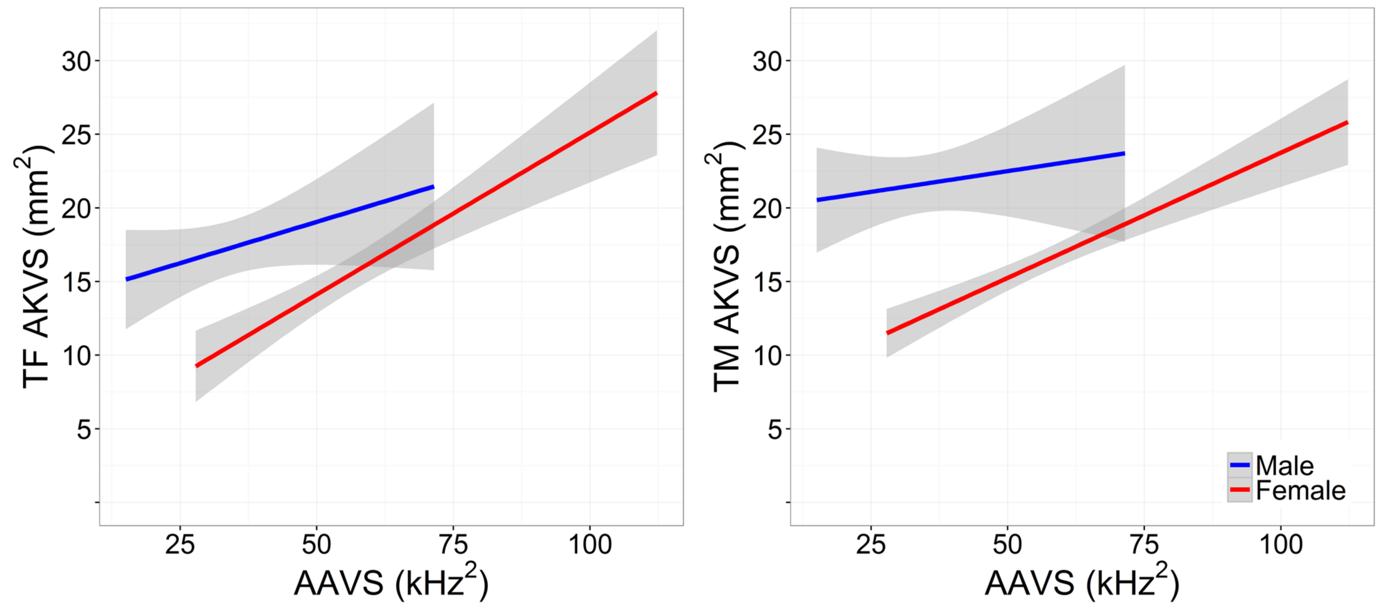

Figure 3. Linear trends showing the relationship for the male and female participants between the Articulatory-Acoustic Vowel Space (AAVS) and the Articulatory-Kinematic Vowel Space (AKVS) for the tongue-front (TF) marker (left) and tongue-mid (TM) marker (right). Grey shaded region represents 95% confidence interval.

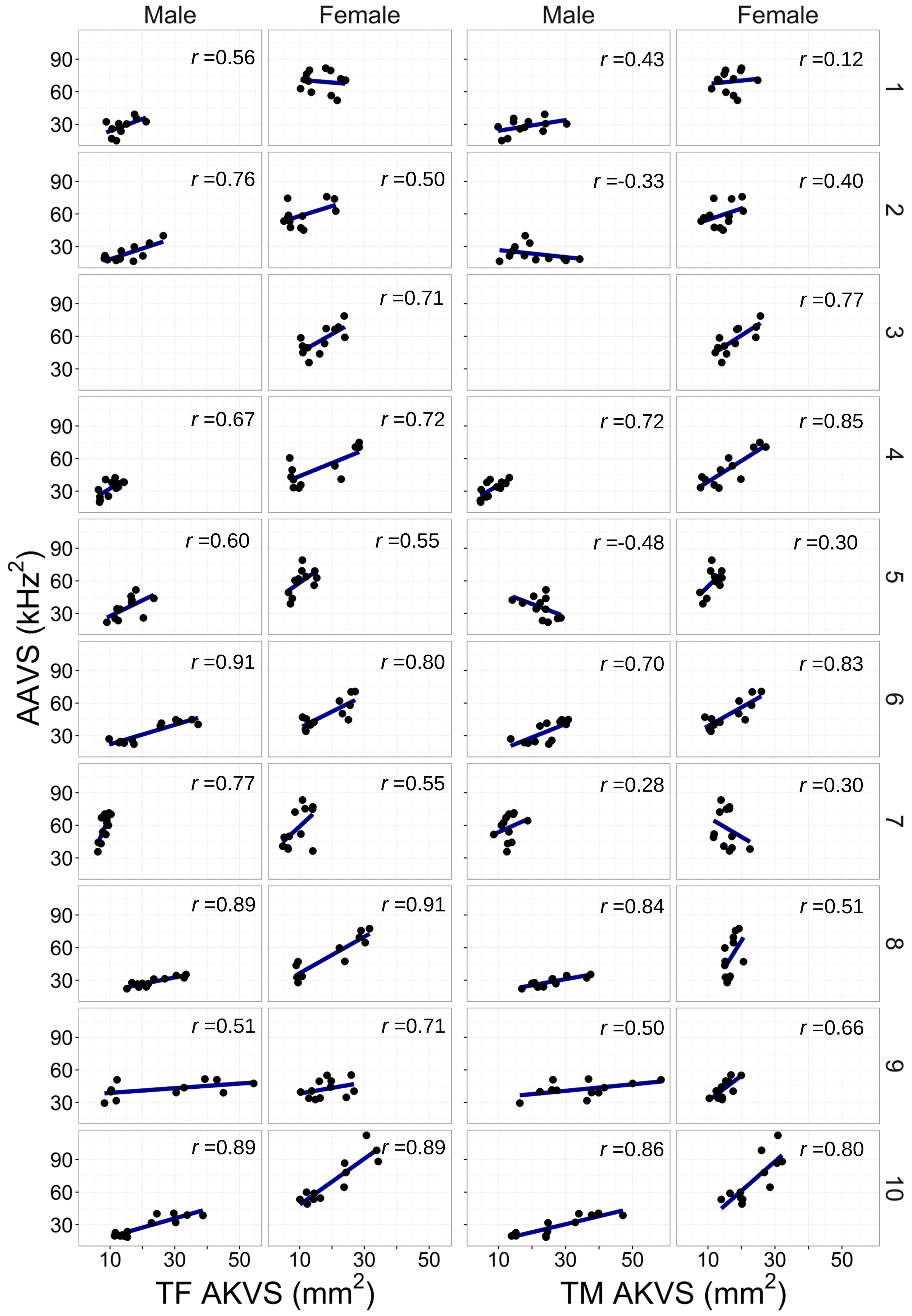

Figure 4. Scatterplots and linear fits showing the relationship between the Articulatory-Acoustic Vowel Space (AAVS) and the Articulatory-Kinematic Vowel Space (AKVS) for the tongue-front (TF) marker (left panes) and tongue-mid (TM) marker (right panes) for each participant.

RESULTS

Articulatory-Acoustic Vowel Space (AAVS)

- Main effect of Sex, F(1,17)=25.61, p<0.001
- Main effect of Condition, F(1,207)=19.38, p=0.001
- Main effect of Stimulus, F(1,207)=138.51, p<0.001

Articulatory-Kinematic Vowel Space (AKVS)

Tongue-Front Marker

- No Main effect of Sex, p>0.05
- Main effect of Condition, F(1,200)=24.26, p<0.001
- Main effect of Stimulus, F(1,198)=187.16, p<0.001

Tongue-Middle Marker

- No Main effect of Sex, p>0.05
- Main effect of Condition, F(1,181)=37.53, p<0.001
- Main effect of Stimulus, F(1,187)=37.34, p<0.001

Summary

- All vowel space measures were larger in the loud condition.
- All vowel space measures were significantly larger for the first stimulus.
- The AAVS was significantly larger for female speakers than for males.
  - No significant sex-related differences were observed for the AKVS measures, though on average females exhibited lower AKVS values.

Mixed Model Regression

The AAVS significantly predicted the AKVS for the tongue-front marker

- Model fit: r^2_m=0.25; r^2_c=0.81
- Fixed Effects:
  - AAVS: F(1,48)=21.411, p<0.001
  - Sex X AAVS: F(1,48)=5.23, p=0.027

The AAVS significantly predicted the AKVS for the tongue-middle marker

- Model fit: r^2_m=0.15; r^2_c=0.86
- Fixed Effects:
  - AAVS: F(1,1844)=21.41, p<0.001

Summary:

- The AAVS predicted the AKVS for the front and central tongue markers.
- Individual differences in the intercept and acoustic-kinematic slopes between participants accounted for a large portion of the variance.

CONCLUSION

Overall, these data show a strong relationship between articulatory range of motion measured from lingual kinematics and the AAVS, within individuals. Comparable changes were observed in the acoustic and kinematic vowel space measures between the speaking and stimulus conditions. Across all participants, the AAVS alone accounted for one quarter of the variance in the AKVS for the front tongue marker data, with additional within-participant effects accounting for over 80% of the total variance. Therefore, the AAVS is likely a strong predictor of the gross articulatory movement during connected speech.

References

Mefferd, A. S., & Green, J. R. (2010). Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations. *Journal of Speech, Language, and Hearing Research, 53(5),* 1206-1219.

Tasko, S. M., & Greilick, K. (2010). Acoustic and articulatory features of diphthong production: A speech clarity study. *Journal of Speech, Language, and Hearing Research, 53(1),* 84-99.

Whitfield, J. A., & Goberman, A. M. (2014). Articulatory–acoustic vowel space: Application to clear speech in individuals with Parkinson's disease. *Journal of Communication Disorders, 51,* 19-28.

2016 Motor Speech Conference, Newport Beach, CA.

jawhitf@bgsu.edu; dromey@byu.edu