ARTICULATORY AND ACOUSTIC CORRELATES OF THE MID-CENTRAL VOWEL

Caroline Menezes, Kelley Moote, Alexis Garon, Jordan Baker, Marisa Lucarelli, Kristyn Nichols and Brandy Pelfrey

University of Toledo, Ohio, U.S.A.
caroline.menezes@utoledo.edu, kelley.moote@rockets.utoledo.edu, Alexis.Garon@rockets.utoledo.edu, Jordan.Baker2@rockets.utoledo.edu, Marisa.Lucarelli@rockets.utoledo.edu, Kristyn.Nichols@rockets.utoledo.edu & Brandy.Pelfrey@rockets.utoledo.edu

ABSTRACT

This paper reports on the articulatory and acoustic correlates of the mid-central vowel (schwa) in an attempt to find out the perimeter of this vowel in both the articulatory and acoustic vowel space of female American English speakers. F1xF2 plots were used to determine the acoustic vowel space. The articulatory vowel space was determined by plotting the maximum vertical displacement of the jaw from the bite plane against the maximum horizontal displacement of the tongue blade from normal resting position. Results indicated that speakers had significantly different articulatory strategies for the production of the schwa, but no speaker differences were observed in the acoustic signal. The acoustic vowel space of this vowel when plotted against the Hillenbrand *et al.*, study appeared to be closer to the back high vowel /u/. The placement of the target word in the utterance had no effect on its acoustic or articulatory realization.

Keywords: mid-central vowel, articulatory, acoustic, EMA, vowel space

1. INTRODUCTION

The mid-central “schwa” in English is a lax neutral vowel whose quality varies greatly depending on the phonetic environment in which it exists [5, 9, & 3]. It can be an epenthetic vowel, a reduced vowel, a rhotic vowel, and an unstressed vowel [1, 2, 5, 9]. When it is emphasized, it can also be produced as the high front vowel /i/. An ESL (English as a second language) teacher refers to this vowel as the “maverick vowel,” “an impure vowel,” and a “lazy vowel”, even going further to call it “a linguistic virus which attacks any and all vulnerable vowels” [2].

This study is motivated by the speech errors of people with motor speech disorders like dysarthria. Research in dysarthria claim that tense vowels are more affected than lax vowels partly because of the reduction of the vowel space However, studies have also found that the vowel space in this population can be increased when speech rate is decreased [13]. However, little is known about the effect of a reduced vowel space on the quality of the mid-central schwa. Since the schwa is important to English rhythm, its quality bears consequence on speech intelligibility in people with speech disorders.

It is well established that the first two formants (F1 & F2) of vowels are strongly influenced by the shape of the airway between the glottis and lips. Perturbation studies have proved that this shape is dependent on the position of the tongue body and lips [4, 12]. High or low tongue body positions would result in altering the first formant frequency, while the front or back tongue body position would affectively change the second formant frequency [12]. The acoustic consequence of a uniform cross-sectional area with opening at the lips and the tongue body at neutral position both in the high-low and front-back dimension would be the central vowel schwa.

In this paper we analyse the articulatory and acoustic variability in the production of the mid-central vowel (schwa). Towards this end the utterances we use contain the English definite article “the” as the target word in sentence initial and medial positions. We will study the articulatory kinematics of healthy subjects to gain understanding of the actual position of the schwa in the oral cavity. The articulatory-acoustic relationship will also be studied by comparing the kinematic data to the first and second formants values derived from the acoustic signal.

2. METHOD

2.1. Subjects

A total of seven female subjects within an age range of 18-20 years were recorded for this study. One subject was excluded from analysis due to large rms errors for the tongue blade coil. All of the subjects were college students attending a university in mid-western United States. All subjects were not reported to have any hearing or speech disorders and spoke mid-western Standard American English.

2.2. Stimuli

The study focused on the mid-central vowel, the schwa, in the English article "the." Allophonic variations of the vowel were studied by using varying phrase positions. By placing the vowel in different positions of the phrase, we changed the quality of the mid-central vowel without changing its phonemic identity. Phrasing was also manipulated by the placement of syntactical boundaries and variation of sentence emphasis as shown in the sentences below. One of the test utterance showing the four boundary (prosodic conditions is listed here.

The sad American story. Pre-target sentence initial
The SAD American story. Post-target sentence initial
AMERICA the sad story. Pre-target sentence medial
America, the SAD story. Post-target sentence medial

A total of 185 sentences were analyzed. Subjects read the stimuli from a computer that was placed directly in front of them. All sentences were presented in a randomized order.

2.3. Procedure

The AG500 Electromagnetic Articulograph (EMA) was used to gather articulatory data from each subject at a sampling rate of 200 Hz. The acoustic data was recorded to the computer at a sampling rate of 16K Hz. For this study, three coils were used as reference; one on the bridge of the nose and two on the left and right mastoid process. Coils were also placed on the vermillion border of the upper and lower lip, on the base of the incisor teeth of the maxilla and mandible, and three were placed on the tongue corresponding to the tongue tip (1cm from the pointed tip), tongue blade, and tongue dorsum.

2.4. Analysis

The target vowel was isolated and labelled in the acoustic signal for all utterances using the PRAAT phonetic software. The articulatory and acoustic signals along with the PRAAT text grids were then imported into the articulatory analysis software Visartico (v 0.9.1 developed by Slim Ouni & Loïc Mangeonjean). The maximum jaw opening values (jaw_z/vertical trajectory), maximum lowering of the tongue blade (tongue blade_z/vertical trajectory) and maximum horizontal movement of the tongue blade were measured within the acoustic target of vowel to be representative of the articulatory point for the realization of the vowel.

PRAAT was also used to identify a stable point within the target vowel from which the first formant frequency (F1) and second formant frequency (F2) measurements were made. The articulatory and acoustic points of measurement did not necessarily coincide for all data points.

3. RESULTS

3.1. Articulatory analysis

Fig. 1 is a scatterplot of jaw_z and tongue blade_z separated by speakers. It shows a strong positive correlation between the jaw position and tongue blade position when all speakers were pooled. Pearson correlation for the pooled speakers was significant (*r(184)* =.858, *p*<.001). However, individual speakers revealed unique relationships between the tongue and the jaw. Speaker KP had a significantly weak but positive correlation (*r*(184) = .483, *p*=.01), and speakers CL and KD had no significant correlation between these two parameters (CL: *r* =.065, KD: *r* =.234). In CL we see large independence of the jaw and tongue.

Figure 1: Scatterplot of jaw_z and tongue blade_z position separated by speakers.

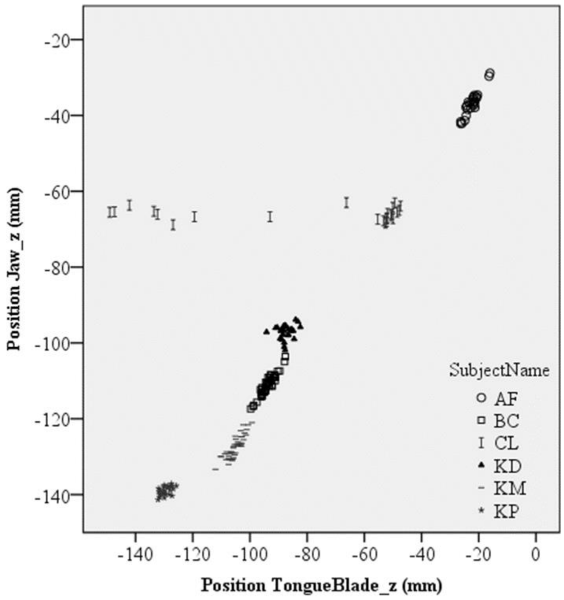
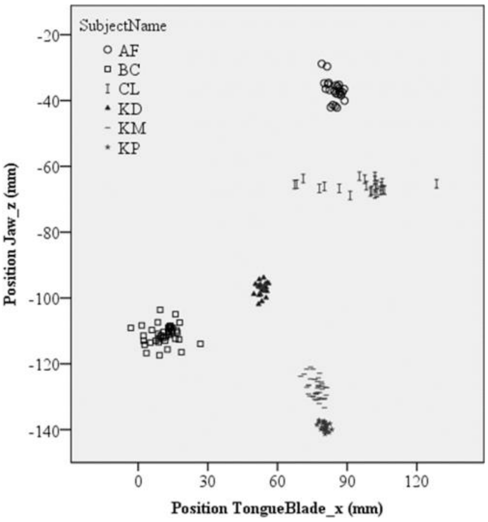


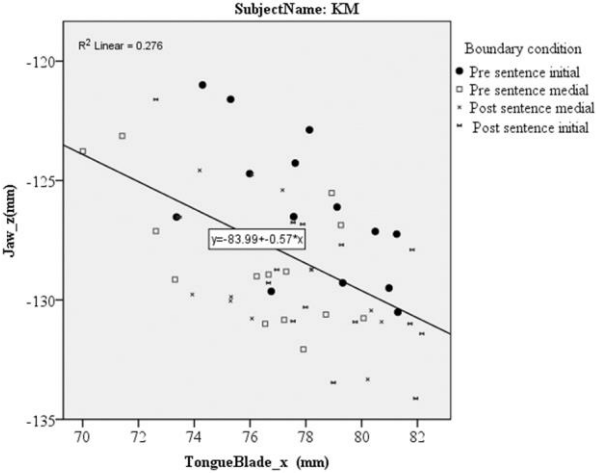
Fig. 2 plots the jaw opening against the horizontal movement of the tongue blade separately for all speakers. Here again we see a clear separation of values by speakers. ANOVA analysis indicated that both vertical position of the jaw and horizontal position of the tongue blade were significantly different for all subjects (*F*(5,160) = 6568 *p*<.001; *F*(5,160) = 860.8 p<.001, respectively). Speakers AF and CL produced this vowel with a lesser jaw opening when compared with speakers BC, KD, KM and KP, while speaker BC had a forward positon of the tongue when compared to the other speakers.

Figure 2: Scatterplot of jaw_z and tongue blade_x position separated by speakers.



A stepwise discriminant analysis revealed that vertical position of the jaw was the most relevant parameter in classifying the different speakers, followed by tongue blade_x (Jaw_z: λ=.000, p<.001; Tongue_x: λ=.020, p<.001; Tongue blade_z: λ=.726, p<.001). Fig. 2 indicates that different speakers have distinctly different articulatory kinematics for the production of this lax central vowel.

Figure 3: Scatterplot of jaw_z and tongue blade_x position separated by boundary conditions for speaker KM.

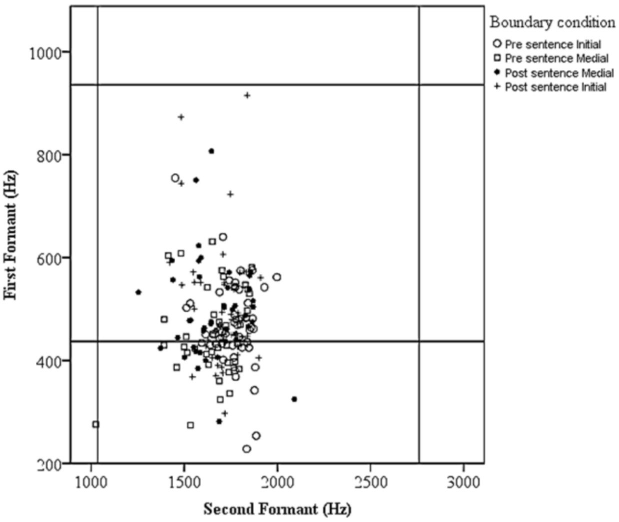


However, when we look closer at each speaker's articulation of the schwa, we see a lot of variability that is not explained by its position in the sentence or the phrasing pattern of the utterance. Fig. 3 is a scatterplot of jaw_z with tongue blade_x separated by boundary conditions for speaker KM. No significant differences were observed for the different prosodic conditions of the target vowel for both jaw opening (jaw_z) and forward position of the tongue (tongue_blade x) for any of the speakers.

3.2. Acoustic analysis

Fig. 4 is a F1-F2 scatterplot for all speakers. The grid lines indicate the vowel space for female speakers according to the Hillenbrand *et al* study [4]. The Hillenbrand study recorded speakers from mid-western U.S.A. similar to this study. As seen in this figure, the mid-central vowel falls within the vowel space however, there is a large spread (n=184, F1: Mean = 480, SD = 99; F2: Mean = 1699, SD = 141). Both F1 and F2 were not significantly different for speakers; however, a small effect was seen for prosodic condition. Table 1 lists the four parameters reported in this paper along with their ANOVA results and partial Eta squared values. Large partial Eta values indicate the importance of that parameter to the ANOVA model. Notice that though F1 and F2 were significantly different for boundary conditions (prosodic condition), the low partial Eta values indicate that they add no value to the analysis.

Figure 4: Scatterplot showing first formant and second formant frequencies separated by speakers.



6. CONCLUSIONS

The motivation to study the mid-central vowel came from analysis of disordered speech. Several studies on apraxia reveal that lesions to the left frontal lobe results in a reduced vowel space [6, 10]. However, the question of how a reduction in the vowel space would affect the central vowel requires a detailed analysis of the vowel space for both clinical and healthy subjects. This paper is a part of a larger exploratory study intended to understand the articulatory deviations in clinical speech.

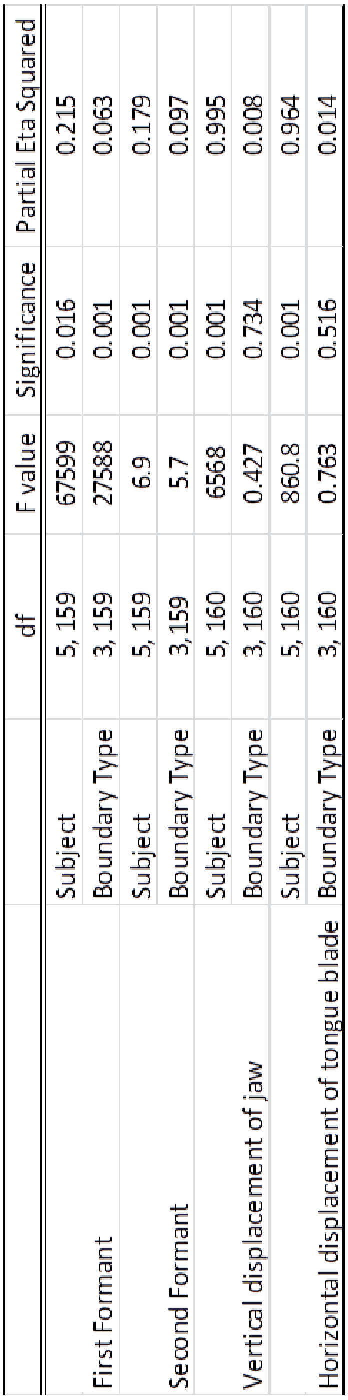
The results reported here reveal some interesting findings. First, it shows that speakers have different articulatory settings when they produce the schwa whose acoustic result was thought to be the result of an undifferentiated vocal tract [12]. To compare the articulatory patterns to the acoustic output of this vowel, we correlated the jaw opening values to tongue blade horizontal position values. This was done because the jaw opening was proven to be a better predictor than the tongue blade in the articulation of this vowel. However, a strong correlation was found between tongue lowering and jaw opening. This is not surprising given that the tongue movement is not completely independent of the jaw.

Allophonic variations of the schwa however did not result in significant changes in the articulatory strategies. We have to note here that the allophonic variation of the schwa was achieved by prosodic manipulation in this study, but it is not clear if the speakers systematically manipulated prosody as the experimenters intended.

The first and second formant frequencies, unlike the articulatory parameters, did not clearly differentiate the speakers. Therefore these results support the idea of an acoustic target instead of an articulatory target. A significantly weak difference was observed for both F1 and F2 between subjects. A significantly weak difference was also observed for both F1 and F2 for the different prosodic conditions, but more data needs to be tested to see if these differences indicate an allophonic variation due to phrasing pattern of the utterances.

Further studies are planned to include male speakers. As mentioned earlier the data from our healthy subjects need to be compared to data from our clinical subjects particularly those with motor speech disorders like apraxia and Parkinson's.

Table 1: ANOVA analysis for F1, F2 and Jaw_z and Tongue blade_x with Subject and Boundary as fixed factors.



		df	F value	Significance	Partial Eta Squared
First Formant	Subject	5, 159	67599	0.016	0.215
	Boundary Type	3, 159	27588	0.001	0.063
Second Formant	Subject	5, 159	6.9	0.001	0.179
	Boundary Type	3,159	5.7	0.001	0.097
Vertical displacement of jaw	Subject	5, 160	6568	0.001	0.995
	Boundary Type	3, 160	0.427	0.734	0.008
Horizontal displacement of tongue blade	Subject	5, 160	860.8	0.001	0.964
	Boundary Type	3, 160	0.763	0.516	0.014

7. REFERENCES

[1] Alfonso, P., & Baer, T. (1982). Dynamics of vowel articulation. *Language and Speech*, 25, 151-173.

[2] Brown, R. L. (1990). A maverick vowel. *Hispania American Association of Teachers of Spanish and Portuguese* 73, 1158-1161.

[3] Browman, C. and Goldstein, L. (1990). "Targetless" schwa: An articulatory analysis. Haskings

Laboratories status report on Speech Research, SR-101/102, 194-219

[4] Chiba T., and Kajiyama M. (1941). *The Vowel: Its Nature and Structure*. Tokyo-Kaiseikan, Tokyo.

[5]Fowler, C. A. (1981). Perception and production of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research*, 24, 127-139.

[6] Haley, K. L., Ohde, R. N., Wertz, R. T. (2001). Vowel quality in aphasia and apraxia speech: Phonetic transcription and formant analyses. *Aphasiology*, 15, 1107–1123.

[7] Hillenbrand, J., Getty, L. A., Clark, M. J., Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. A.* 97(5), 3099-3111.

[8] Liss, L.M., Spitzer, S.M., Caviness, J.N., Adler, C., & Edwards, B.W. (2000). Lexical boundary error analysis in hypokinetic and ataxic dysarthria. *J. of Acoust. Soc. of America*, 107, 3415-3423.

[9] Magen, H. (1989). An acoustic study of vowel-to-vowel coarticulation in English. Doctoral dissertation, Yale University, Department of Linguistics

[10] Menezes, C. M. (2013). Acoustic and articulatory variation in the mid-central vowel in apraxic and normal speech. *Proc. of the 8th International workshop of Models and Analysis of Vocal Emissions for Biomedical Applications*, Firenze, Italy.

[11] Sapir, S., Spielman, J., Ramig, L.O., Hinds, S.L., Countrymand, S., Fox, C., & Story, B. (2003) Effects of intensive voice treatment (the Lee Silverman voice treatment [LSVT] on ataxic dysarthria: A case study. *American Journal of Speech-Language Pathology*, 12, 387-399.

[12] Stevens, K. 1999. Articulatory-acoustic-auditory relationships. In: Hardcastle, W., Laver, J. (eds), *The Handbook of Phonetic Science*. Oxford: Blackwell, 462-506.

[13]Tjaden, K., Rivera, D. & Wilding, G. (2005) Characteristics of the lax vowel in space in Dysarthria. J. of Speech, Language, and Hearing Research, 48, 554-566.