VOWEL SPACE AREAS ACROSS DIALECTS AND GENDER

Ewa Jacewicz†*, Robert Allen Fox*†*, and Joseph Salmons*‡

†Speech Perception and Acoustics Labs, The Ohio State University
‡University of Wisconsin-Madison
jacewicz.1@osu.edu, fox.2@osu.edu, jsalmons@wisc.edu

ABSTRACT

This study compares vowel spaces in three regional varieties of American English spoken in central Ohio, south-central Wisconsin, and western North Carolina to determine whether the significant variation in the vowel systems of these dialects also affects the dialect-specific vowel space area. The gender-related differences are assessed by comparing the unnormalized (in Hz) and normalized formant frequency values. Significant effects of speaker dialect were found for the vowel space area defined by four "corner" vowels. However, there were no differences between dialects in the area of an extended 5-vowel space. The results indicate that, despite large cross-dialectal differences in the positions of the vowels in the acoustic space, the extended vowel space area encompassing a complete vowel system is unaffected by dialectal variation. The differences in the size of the vowel space due to speaker gender were eliminated by normalizing formant frequency values.

Keywords: vowel space, regional dialect, vowel normalization, speaker gender, sociophonetics.

1. INTRODUCTION

A number of recent sociophonetic studies have demonstrated that the acoustic characteristics of vowels in languages such as American English and Dutch vary significantly across geographic regions [1] [3]. The present study examines whether the significant regional variation in vowels also affects the size of the "working" vowel space (in the F1 x F2 plane) used by the speakers. Regional differences in the size of the vowel space area may provide insight into ongoing vowel chain shifts, helping to track, measure and predict such changes as push and pull chains. Two outcomes are possible. First, given the significant differences in vowel systems across regional dialects, one might find that the size of the vowel space is also affected by this type of variation and will differ cross-dialectally. Alternatively, the area of the vowel space may remain the same across dialects and only the relative positions of vowels within the vowel space contribute to the dialectal differences.

This study investigates the size of vowel spaces used by speakers of three distinct regions in the United States. In particular, we compare vowel space areas among the American English dialects spoken in central Ohio, south-central Wisconsin, and western North Carolina, each of which are currently undergoing distinct patterns of changes and shifts in their respective vowel systems (see [4]). For example, English spoken in Wisconsin is undergoing the Northern Cities Shift, English in western North Carolina is affected by the Southern Vowel Shift, and although there is no known shift operating in central Ohio there are specific changes currently underway.

In the past, the size of the acoustic vowel space was measured as a triangular area defined by the three traditional "corner" vowels /i, a, u/ (e.g., [5]) or /i, o, u/ [2]. This was done to assess the expansion of the vowel space such as in child-directed and in intelligibility-enhancing speech. However, this triangular area cannot adequately characterize the size of the vowel space across dialects as the 3-vowel area largely underestimates the actual "working space" of vowel system. Inevitably, the 3-vowel area does not take into account a number of vowels located outside this area. To avoid this problem, the present study examines the size of extended vowel spaces encompassing four and five vowels, respectively.

2. METHOD

2.1. Speakers

Recordings were obtained from 54 speakers aged 20-34 years. There were 18 speakers (9 male, 9 female) for each dialect area who were born and raised in either central Ohio, south-central Wisconsin, or western North Carolina.

2.2. Stimuli and recording procedure

Stimulus material consisted of the following single real and nonce words in /hVd/ context: *heed, hid, head, hey'd, had, heard, who'd, hood, hoed, hawed, hod, hide, howed, hoyd* which contained 14 vowels and diphthongs of American English /i, ɪ, ɛ, e, æ, ɚ, u, ʊ, o, ɔ, ɑ, aɪ, aʊ, oɪ/. Recordings were under computer control using a program in Matlab. Words were presented in random order on a computer screen to a subject seated in a sound-attenuating booth and were recorded directly onto a hard drive disk at a 44.1-kHz sampling rate. A head-mounted microphone (Shure SM10A) was used, placed at a 1-inch distance from the lips. A total of 42 words were recorded from each speaker (14 words x 3 repetitions).

2.3. Acoustic analysis

First, vowel onsets and offsets were located by hand from the waveform (with reference to a spectrogram) and the overall duration was calculated. Vowel onset was located at the zero-crossing before the first positive peak in the periodic waveform and vowel offset was defined as the beginning of the stop closure (location of abrupt decrement in the amplitude of the waveform). Formant frequency values were then extracted automatically using a program in Matlab (a 14-order LPC analysis with a 25-ms Hamming window). The frequencies of F1, F2, and F3 were measured at five temporal locations corresponding to the 20-35-50-65-80%-point over the course of each vowel's duration, which allows an examination of formant trajectories over time. Prior to spectral analysis, the tokens were downsampled to 11.025 kHz and pre-emphasized (98%).

2.4. Calculation of the vowel space area

The vowels selected for calculation of extended areas of the vowel space were /i, æ, u, ɑ, oɪ/. These vowels encompassed most of the vowel space used by the speakers across the three dialects. This was evident from a F1 x F2 display of dynamic formant pattern sampled at five equidistant locations in the vowel.

First, the averaged F1 and F2 values of two sets of three vowels /i, æ, u/ and /u, æ, ɑ/ of each speaker were used to calculate the vowel areas of the /i-æ-u/ and /u-æ-ɑ/ triangles using Heron's method:

(1) Area = SQRT(s(s-a)(s-b)(s-c))
where s = (a+b+c)/2 or perimeter/2

To estimate the area of the 4-vowel /i-æ-u-ɑ/ quadrangle, the areas of the two vowel triangles were then combined. To calculate the 5-vowel space, the area of a third contiguous vowel triangle /u-ɑ-oɪ/ was calculated and the areas of all three triangles were combined. These areas were calculated at both the 20% and 35% - temporal point in the vowel. These two measurement locations present the most expanded characterization of the working vowel space. The measurements at later points in time such as 65 and 80% (and, to some extent, 50% point) tend to portray relatively centralized vowels which would tend to reduce the vowel space area. For all speakers, the position of /oɪ/ at both the 20% and 35% points were outside the /i-æ-u-ɑ/ quadrangle.

3. RESULTS

3.1. The 4-vowel space

Figure 1: Mean vowel space areas for male speakers.

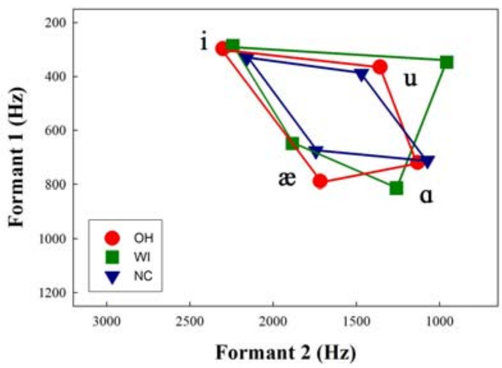
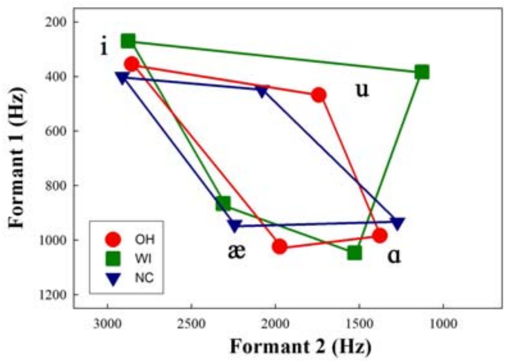


Figure 2: Mean vowel space areas for female speakers.



The 4-vowel spaces for each dialect calculated at the 20% temporal point in the vowel are shown in Fig.1 for males and in Fig. 2 for females. The larger size of the female vowel space is most likely due to the differences in the vocal tract length between male and female speakers. Despite this general difference in size, it can be seen that the shapes of the vowel spaces vary across dialects. Also, the Wisconsin (WI) space is greater than either Ohio (OH) or North Carolina (NC), most likely due to the far back position of WI /u/. Positional differences across dialects can also be found for the vowels /æ/ and /ɑ/, each reflecting regional variation in the phonetic form of these vowels.

The results of a three-way ANOVA showed no significant main effect of measurement location in a vowel (either 20% or 35%) but there were significant main effects of gender [$F(1, 47) = 51.5$, $p<0.001$] and dialect [$F(2, 47)=15.3$, $p<0.001$]. Post-hoc Tukey tests revealed that the WI vowel space area was significantly greater than either OH or NC, but the latter two did not differ significantly from each other.

To examine whether the significance of gender-related differences due to anatomical and/or physiological factors can be eliminated, the formant frequency values were converted to *z*-scores using the vowel-extrinsic normalization procedure developed by Lobanov [6].

Figure 3: Mean normalized vowel space areas for male speakers.

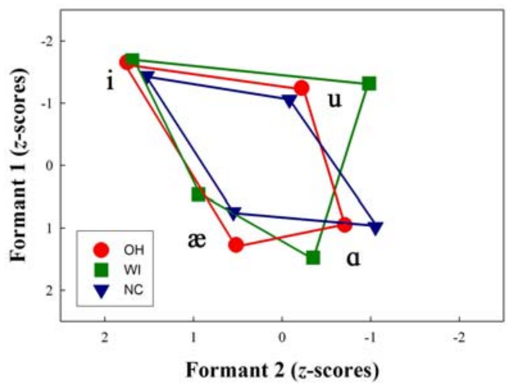
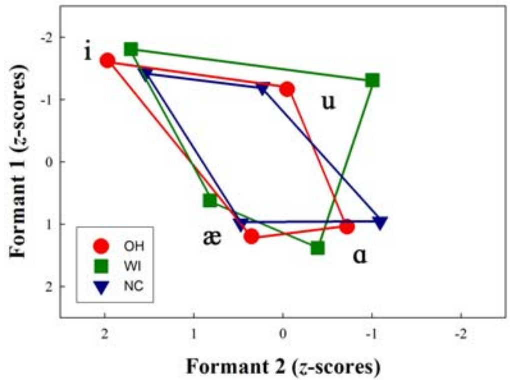


Fig. 3 and 4 show the normalized 4-vowel space areas for male and female speakers, respectively. Clearly, the sizes of the male and female spaces are now comparable and the dialect-specific differences in the shape of the vowel spaces are still preserved. A three-way ANOVA on the normalized 4-vowel area values with measurement point (20%, 35%) as a repeated-measure showed no significant main effect of measurement location. Most importantly, the effect of gender was not significant while there remained a significant main effect of dialect [$F(2,47)=13.56$, $p<0.001$]. Post-hoc analyses showed that the WI vowel space was significantly greater than either OH or NC, but the latter two did not differ significantly from each other.

Figure 4: Mean normalized vowel space areas for female speakers.



Overall, the results for the 4-vowel space showed significant dialectal differences in the vowel space area. These differences were present even when the effects of speaker gender were minimized. Although some inter-speaker variability was found within each dialect, the statistical analyses still demonstrated significant differences.

3.2. The 5-vowel space

The cross-dialectal differences in the degree of /u/-fronting may affect estimates of the extent of the vowel space in the back region. Although the 4-vowel space accounts for the cross-dialectal variation in the low front vowel /æ/, it still underestimates the region toward the far back of the vowel space. Thus, the 4-vowel space may not characterize the *total extent* of the vowel space (wherein all of the vowels—monophthongs and diphthongs—are produced).

The cross-dialectal comparison of the position of the diphthong /oɪ/ in the vowel space indicated that the beginning of the diphthong measured at its 20% temporal point constituted the farthest back location of a back vowel in vowel systems of each dialect. To better estimate a complete vowel space used by the present speakers, we thus included a third triangular area in our calculations, encompassing the vowels /u, ɑ, oɪ/.

Fig. 5, 6, and 7 show the normalized 5-vowel spaces for OH, WI, and NC, respectively. As evident from the displays, the size of the third triangular area encompassing the back vowels /u, ɑ, oɪ/ contributed appreciably to the extension of the vowel space toward the back, particularly for NC and OH vowel systems. The cross-dialectal differences in the shape of the vowel space are also noteworthy. Consistent with what was found for the 4-vowel space, the differences due to speaker gender were eliminated by normalizing formant frequency values.

Figure 5: Mean normalized vowel space area for male and female Ohio speakers.

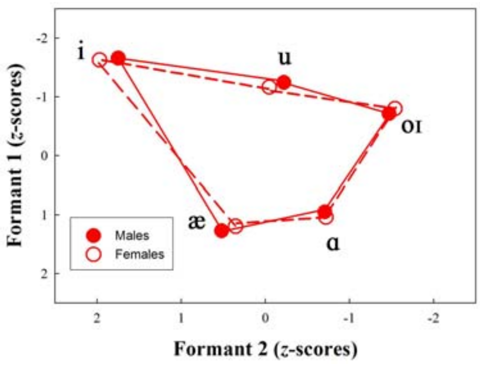


Figure 6: Mean normalized vowel space area for male and female Wisconsin speakers.

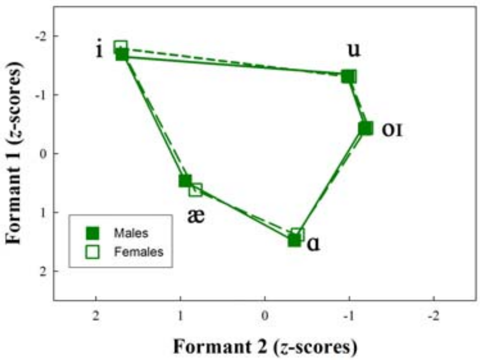
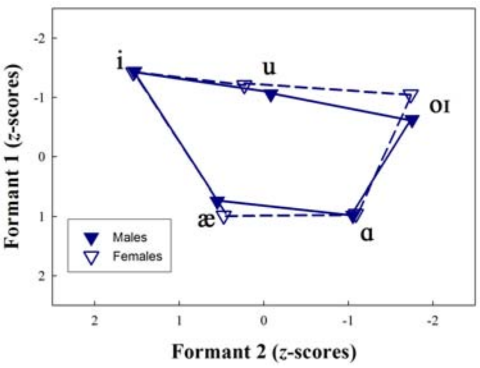


Figure 7: Mean normalized vowel space area for male and female North Carolina speakers.



Results of an ANOVA on normalized values for the 5-vowel space revealed no significant main effect of either gender or dialect. Although the lack of differences due to speaker gender is not surprising, the results for dialect indicate that the general vowel space encompassing a complete vowel system remains the same across dialects.

4. CONCLUSION

The present results indicate that, despite differences in the formant frequency values which influence the phonetic quality of particular vowels, the extended vowel space area encompassing a complete vowel system is unaffected by dialectal variation. Although the positions of "corner vowels" (and thus, the shape of the vowel space) may differ, the size of the entire vowel space area used by male or female speakers of the three distinct regional varieties of American English remains the same. This indicates that significant differences as a function of dialect obtained for the 4-vowel space stem from underestimating the size of the vowel space used by the speakers. However, this conclusion can be reached if we consider normalized formant values only. It is an empirical question of whether the normalization process eliminates actual and significant dialect variation. Of course, it is always the case that other vowel characteristics such as formant frequency change over time, rate and magnitude of formant frequency change as well as vowel duration that contribute significantly to cross-dialectal differences.

ACKNOWLEDGMENTS

This study was funded in part by NIH grant number blahbitty blah blah.

REFERENCES

[1] Adank, P., van Hout, R., van de Velde, H. 2007. An acoustic descritption of the vowels of northern and southern standard Dutch II: Regional varieties. *J. Acoust. Soc. Am.* 121, 1130-1141.

[2] Bradlow, A., Toretta, G., Pisoni, D. 1996. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20, 255-272.

[3] Clopper, C., Pisoni, D., de Jong, K. 2005. Acoustic characteristics of the vowel systems of six regional varities of American English. *J. Acoust. Soc. Am.* 118, 1661-1676.

[4] Labov, W., Ash, S., Boberg, C. 2006. *Atlas of North American English* (Mouton de Gruyter, New York).

[5] Liu, H-M, Kuhl, P., Tsao, F-M. 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6:3, F1-F10.

[6] Lobanov, B. 1971. Classification of Russian vowels spoken by different speakers. *J. Acoust. Soc. Am.* 49, 606-608.