Acoustics'08 Paris
June 29-July 4, 2008
www.acoustics08-paris.org

Analysis of total vowel space areas in three regional dialects of American English

Robert Fox and Ewa Jacewicz

The Ohio State University, 110 Pressey Hall, 1070 Carmack Road, Columbus, OH 43210, USA
fox.2@osu.edu

The acoustic characteristics of vowel systems in different varieties of American English are greatly affected by regional variation. Given the significant positional differences of vowels within the acoustic space across regional dialects of English, one should expect that the size and extent of the vowel space is also affected by this type of variation. Traditionally, the size of the acoustic vowel space has been measured as the triangular area defined by the three corner vowels. An obvious weakness of this approach is that it underestimates the actual "working space" of vowel system in that the onsets and/or offsets of other vowels are often found outside this triangular area. This paper proposes a procedure to estimate the area of a complete vowel space, taking into account dynamic formant pattern of all vowels and diphthongs. Complete vowel space areas are calculated for individual speakers and compared across three distinct regional varieties (representing Northern Cities Shift, Southern Vowel Shift, and a Midland variety). The comparison also examines changes to vowel space area as a function of speaker generation (younger and older adult female speakers.)

1 Introduction

The acoustic characteristics of vowels in languages such as American English and Dutch vary significantly across geographic regions. The cross-dialectal differences have been found primarily in formant patterns [1, 2] and vowel duration [3]. Given the reported large variation in vowel formants which determine the position of vowels in the acoustic vowel space, the question arises whether the speakers of respective regional varieties utilize the same "working area" in articulating their vowels or whether the size of the acoustic vowel space (in the F1 x F2 plane) differs across dialects.

This question was initially explored in [4] where formant frequency values were used to examine the size of the vowel space in three American English dialects spoken in central Ohio, south-central Wisconsin, and western North Carolina. Significant differences were found for the 4-vowel space (the F1-F2 quadrilateral /i-æ-ɑ-u/), which persisted even when the effects of speaker gender were minimized using normalized formant values. However, these significant differences disappeared for the 5-vowel space area which included the diphthong /ɔɪ/, whose onglide (measured at its 20% temporal point) constituted the most back position of a back vowel in the F1 x F2 plane of any of the three dialects. It was tentatively concluded that significant dialectal differences obtained for the 4-vowel space stem from underestimating the size of the vowel space used by the speakers. Although the positions of the four "corner vowels" may differ, the size of the extended vowel space area remains the same, as shown for the /i-æ-ɑ-ɔɪ-u/ space.

The present study further examines the vowel space areas in the same three dialects with the objective to provide a more complete characterization of the "working vowel space" and to examine the changes in the size of the vowel space area over the course of vowels' durations. This approach is taken to understand and estimate usual variation in the vowel space area which occurs as the formant trajectories change dynamically over time. The areas are calculated for each of the five equidistant temporal locations in the vowel, which sample the trajectories of F1 and F2 at 20-35-50-65-80% temporal locations across the vowel's duration. Speaker age was also included as a factor to observe the changes to the vowel space for young and older adults. This was done not merely to examine the changes as a function of aging which can be somewhat expected based on what has been found for developmental and maturational effects on the vowel space [5]. Rather, the age-related changes to the vowel space area may reflect different stages in the patterns of vowel shifts and other changes which are currently undergoing in the respective vowel systems of the three regional varieties of American English.

The present study extends [4] to include all perimeter vowels that define a more complete vowel space. The results are first presented for the four corner vowels /i,æ,ɑ,u/ and then compared with the results for the total vowel space which is determined by dialect-specific perimeter vowels, regardless of their quality.

As a whole, the study seeks to characterize the size and nature of the variation in the vowel space area as a function of vowel dynamics, diachronic change, and speaker dialect. In terms of measurement, it proposes a way to estimate a *complete vowel space area* which takes into account all the above factors.

2 Methods

2.1 Speakers

The participants of the study were 54 women who came from three distinct dialectal regions in the United States: south-central Wisconsin, central Ohio, and western North Carolina. In each regional group, there were nine young women aged 20-34 years and nine older women aged 51-65 years. The Wisconsin speakers came from the Madison area (Dane and Dodge counties). The Ohio speakers were from the Columbus area (Franklin and Delaware counties). The North Carolina speakers came from the Sylva, Cullowhee, and Waynesville areas (Jackson, Swain and Haywood counties). Defined geographically, these participants created highly homogenous samples of regional speech. The Wisconsin and Ohio speakers grew up in a suburban setting and the North Carolina speakers mostly in rural areas and small towns. Most of the young adults were students at either University of Wisconsin-Madison, The Ohio State University, or Western Carolina University. Older adults represented a variety of educational backgrounds. None of the participants reported any speech disorder.

2.2 Speech materials and procedure

Speech materials consisted of the following single words of the /hVd/ structure: *heed, hid, head, hey'd, had, heard, who'd, hood, hoed, hawed, hod, hide, howed, hoyd.* The words contained 14 vowels and diphthongs of American English which were used for calculation of vowel space areas: /i, ɪ, ɛ, e, æ, ɚ, u, ʊ, o, ɔ, ɑ, aɪ, aʊ, ɔɪ/. Recordings were controlled by a custom program written 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 Shure SM10A dynamic microphone was used, positioned at a distance of 2 inches from the speaker's lips. A total of 42 words were recorded from each speaker (14 words x 3 repetitions).

2.3 Acoustic measurements

Prior to acoustic analysis, the tokens were digitally filtered and downsampled to 11.025 kHz and pre-emphasized (98%). First, vowel onsets and offsets were located by hand from the waveform (with reference to a spectrogram) and the overall vowel 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 custom 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 equidistant temporal locations corresponding to the 20-35-50-65-80%-point over the course of each vowel's duration. These measurements allowed an examination of formant trajectories over time, estimating the dynamic formant change for each vowel.

2.4 Calculation of the vowel space area

Vowel space areas in the F1 x F2 plane at each temporal location were calculated on the basis of each speaker's mean formant values in two ways using a dedicated MATLAB program.

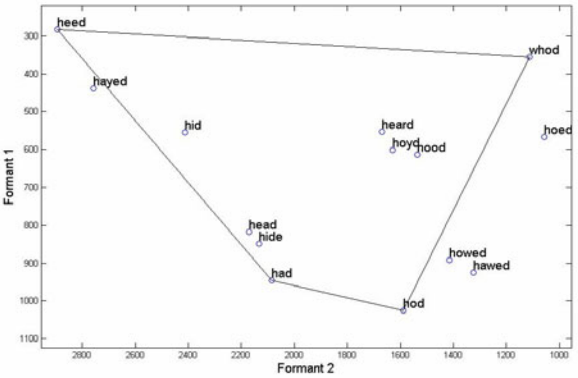


Fig. 1 Mean values of F1 and F2 for young Wisconsin female speakers for 14 stimulus vowels at vowel midpoint (the 50% temporal position). Outline of vowel quadrilateral is superimposed.

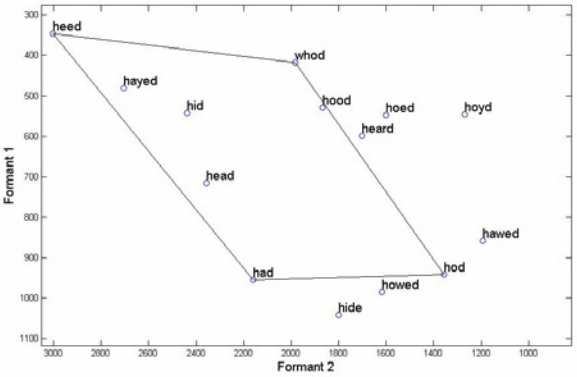


Fig. 2 Mean values of F1 and F2 for young North Carolina female speakers for 14 stimulus vowels at vowel midpoint (the 50% temporal position). Outline of vowel quadrilateral is superimposed.

First, the vowel space area was calculated on the basis of the four "corner" vowels /i, u, ɑ, æ/. This constitutes "vowel quadrilateral." The vowel areas of the /i-u-ɑ/ and /i-æ-ɑ/ triangles (of which the quadrilateral consists) were calculated using Heron's method:

$$\text{area} = \text{sqrt}\,(s(s-a)(s-b)(s-c) \qquad (1)$$

where $s = (a + b + c)$;

a, b, c, represent the lengths of the three sides of a vowel triangle.

The vowel quadrilateral area was then determined by adding the areas of these two triangles. Examples of the quadrilaterals are shown in Fig. 1 and Fig. 2, based on the mean values of vowels from Wisconsin and North Carolina speakers, respectively.

Next, the overall vowel space area at each temporal location was calculated for each speaker. This vowel space was defined by the perimeter vowels in the F1 x F2 referred here as the "vowel multilateral." The MATLAB program allowed the experimenter to select additional vowels that created non-overlapping triangles within the overall vowel space. These areas were then summed to create the overall vowel space area (as exemplified in Fig. 3 and Fig. 4 for the mean Wisconsin and North Carolina acoustic vowel space, respectively). Later versions of this program calculated the overall vowel space area using MATLAB's polyarea function.

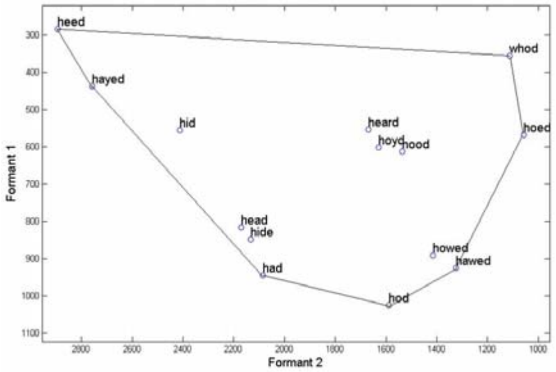


Fig. 3 Mean values of F1 and F2 for young Wisconsin female speakers for 14 stimulus vowels at vowel midpoint (the 50% temporal position). Outline of vowel multilateral is superimposed.

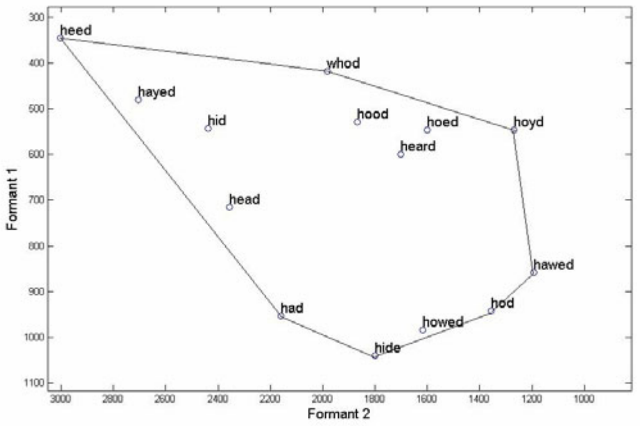


Fig. 4 Mean values of F1 and F2 for young North Carolina female speakers for 14 stimulus vowels at vowel midpoint (the 50% temporal position). Outline of vowel multilateral is superimposed.

3 Results

3.1 Quadrilateral vowel space

The areas of the vowel quadrilaterals at each of the five temporal locations for each of the speakers within the three dialects (OH, WI and NC) and two age groups (young adults and older adults) were analyzed using ANOVA with the within-subject factor location and the between-subject factors dialect and age.

There was a significant difference in the size of the quadrilateral areas as a function of dialect (F(2,48)=18.9, p<.001, η^2=0.440). Wisconsin had the largest area (.670 kHz^2) followed by Ohio (.513 kHz^2) and North Carolina (.401 kHz^2). The primary reason for these significant differences is evident in Fig. 1 and Fig. 2. The /u/ vowel in the OH and NC is fronted while the /u/ in the Wisconsin vowel space is retracted.

There was also a significant effect of temporal location (F(4,192)=53.6, p<.001, η^2=0.527). The mean areas of the 20%, 35%, 50% and 65% locations were relatively close in size (.541, .568, .557 and .537 kHz^2, respectively) but the area of the final location (80%) was significantly smaller (.436 kHz^2, respectively). However, the change in the size of the quadrilateral vowel area remained relatively stable for the OH and NC vowels. The quadrilateral area for the WI speakers dropped significantly from the 35% to the final 80% location. This produced a significant location x dialect interaction (F(8,192)=15.3, p<.001, η^2=0.405) which is illustrated in Fig. 5.

There was no significant age effect (F(1,48)=1.99, n.s.) but there was a significant dialect x age interaction (F(1,48)=5.05, p<.03, η^2=0.095). Young adult speakers, especially in the WI and NC dialect groups, had larger quadrilateral areas than did the older adult speakers in the 20%-65% temporal locations. These effects are shown in Fig. 6.

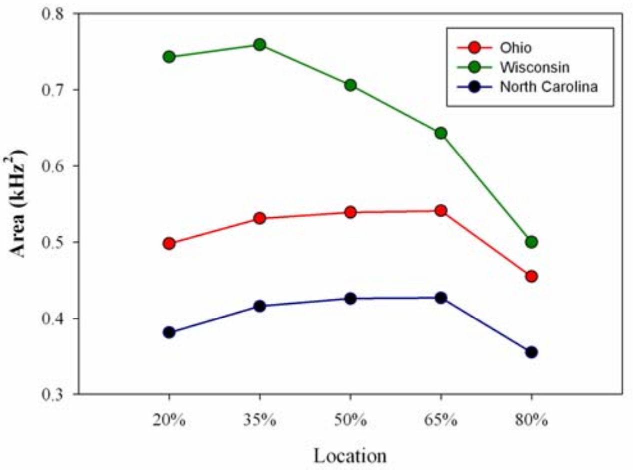


Fig. 5 Change in quadrilateral area as a function of dialect and temporal position.

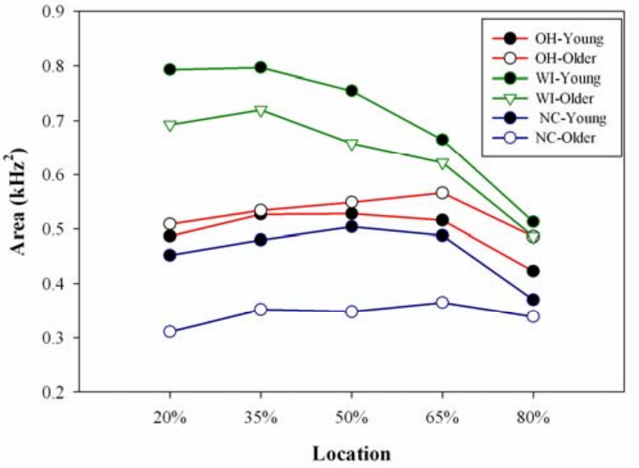


Fig. 6 Change to the quadrilateral area as a function of speaker age, dialect and temporal position.

3.2 Multilateral vowel space

It is clear that the formant frequencies of many vowels displayed in the graphs here fall, at least at some point in their production, outside the vowel quadrilateral /i-u-ɑ-æ/. If we are interested in determining whether speakers of different dialects and/or ages use vowel spaces of different sizes then we need to expand our view of the total vowel space area. We thus now turn to an analysis of the vowel space determined by the perimeter vowels, regardless of their vowel quality (i.e., not limited to the area defined by the positions of the 4 corner vowels /i-u-ɑ-æ/).

The mean total vowel space areas (i.e., the areas of the multilateral polygon) at each of the five temporal locations for the three dialects (OH, WI and NC) and two age groups are shown in Fig. 7. It is evident that there are both similarities and differences between these multilateral areas and the quadrilateral areas. These data were again analyzed using a repeated-measures ANOVA with the within-subject factor temporal location and the between-subject factors dialect and age.

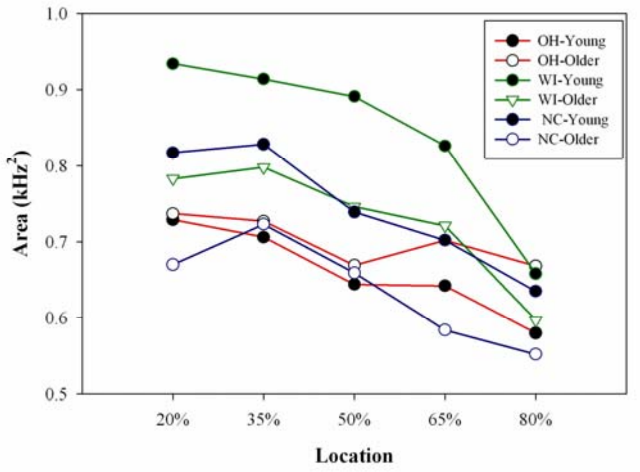


Fig. 7 Change to the multilateral space as a function of speaker age and dialect.

Similar to the pattern seen for the quadrilateral areas, the mean area of the total WI vowel space (.787 kHz^2) was larger than that of either the OH (.614 kHz^2) or NC (0.691 kHz^2) vowel spaces, although the difference was just be below significance (F(2,48)=2.99, p=.06). There was a significant dialect x location interaction (F(8,192)=7.1, p<.001, η^2=0.228). The total vowel area for the WI vowel space was significantly larger than for the OH and NC vowel spaces at locations 20%, 35%, 50% and 65%, but there were no significant differences at the 80% location (see Fig. 8).

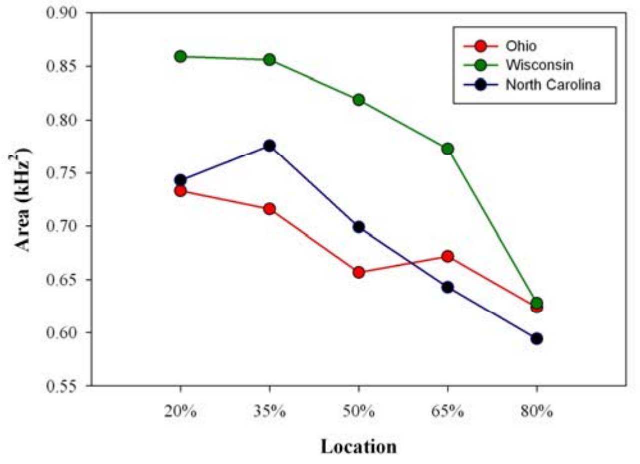
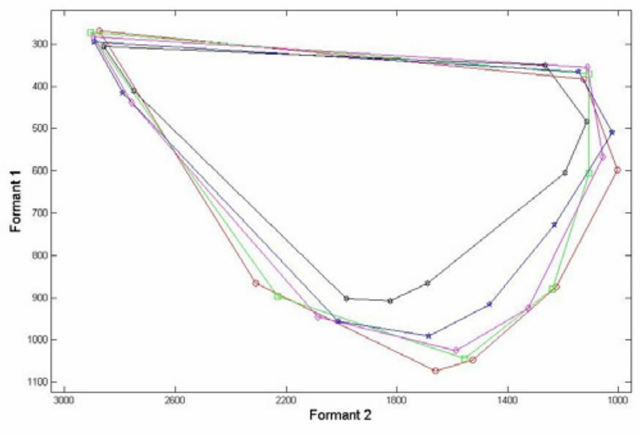


Fig. 8 Change in multilateral area as a function of dialect and temporal position.

There was a significant main effect of temporal location. In general, the total vowel areas decreased in size as a function of temporal location (the total mean vowel area was smallest at the 80% location (.615 kHz^2) and largest at the 20% (.778 kHz^2) and 35% (.783 kHz^2) locations. These variations in size are illustrated in Fig. 9 for Wisconsin and North Carolina young speakers. As can be seen for either dialect, the total vowel space area is gradually reduced over the course of vowel's duration, being largest at 20-35% and smallest at 80%.

Wisconsin



North Carolina

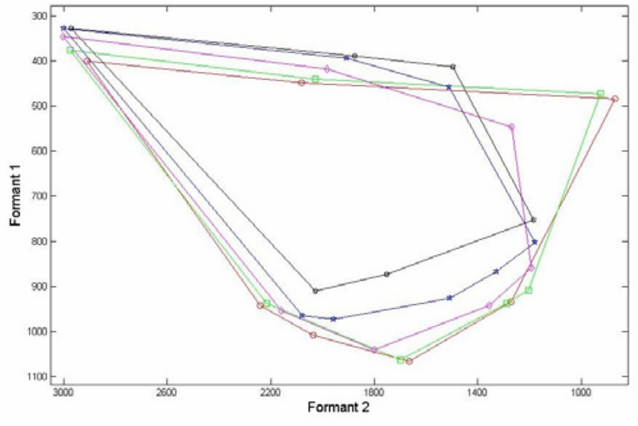


Fig. 9 Variation in size of the total vowel space area as a function of temporal location. Temporal location marks: 20% - red circles, 35% - green squares, 50% - magenta diamonds, 65% - blue pentagrams, 80% - black hexagrams.

The main effect of age was not significant (F(1,48)=1.7, n.s.), but there was a significant age x location interaction (F(4,192)=3.6, p=.007, η^2=0.070). Young adults had significantly larger total vowel areas across the first 4 temporal locations than did the older adults.

4 Discussion

The results for the quadrilateral vowel space show significant dialectal differences which are primarily related to the position of the back corner vowel /u/ within each dialect. Measured over the course of vowel's duration, the size of the vowel space did not change during the initial and medial portions of the vowels for the Ohio and North Carolina speakers. However, the Wisconsin speakers showed a different pattern in that their vowel space area became progressively smaller beginning from the 35% point in time.

For all three dialects, the respective vowel space areas were significantly reduced at the final temporal location close to the vowel offsets. This pattern of results suggests that the size of the vowel space corresponds to the dynamic changes in formant trajectories and, thus, changes in the positions of vowels (especially diphthongs) in the acoustic space as they were measured at five different temporal locations.

The results for the multilateral vowel space indicate that some differences in the vowel space areas observed for the quadrilateral vowel space persist even when all perimeter vowels are considered. Most importantly, the vowel space area for Wisconsin speakers was again largest, which indicates that it is not only the location of the back corner vowel /u/ that produces the differences in the size of the working vowel space area. Rather, the positions of other vowels contribute to the dialectal differences in the complete vowel space area. Consistent with the results for the quadrilateral vowel space, the total vowel space areas for each dialect were gradually reduced over the course of vowel's duration being smallest at 80% point in time.

The effects of age were particularly interesting for the quadrilateral vowel space. The main effect of age was not significant, indicating that there was no difference in the size of the vowel space as a function of aging. However, the significant dialect x age interaction suggests dialect-specific diachronic changes in the vowel systems. For Wisconsin speakers, these changes included lowering of the low vowels /ɑ, æ/ in the speech of the younger generation represented here by the young adults. The lowering of these two vowels contributed to the larger vowel space area for the young Wisconsin speakers as compared to the older speakers.

Although a somewhat larger quadrilateral vowel space was also found for the young North Carolina speakers as compared to the older speakers, this effect may be attributed to a different pattern of dialectal sound change. In particular, the front vowels (including /i/) are more fronted in the production of the young North Carolina speakers as compared to the older adults. In addition, the low vowels /ɑ, æ/ are lower as compared to the older adults. The diachronic changes to the corner vowels /i, æ, ɑ/ most likely contributed to the expansion of the quadrilateral vowel space for young adults in North Carolina. No diachronic changes to the position of the corner vowels can be observed for Ohio speakers and, consequently, there were no differences in the vowel space area for the young and older adults in Ohio.

5 Conclusion

This study presented a measurement technique which estimates a working vowel space area for both the simplified space (such as based on vowel quadrilateral /i-u-ɑ-æ/) or a complete vowel space defined by a larger set of perimeter vowels. This technique allowed us to characterize vowel space areas over the course of vowel's duration, taking into account dynamic changes in formant trajectories. Many of these area differences – especially for the quadrilateral vowel space – can be explained by well-known phonetic differences between dialects. As shown in this study, speaker dialect has an effect on the size of the vowel space area. The effects of speaker age were also present which can be attributed to a dialect-specific diachronic change in the vowel system rather than to the effects of aging per se.

Acknowledgments

This study was supported by the research grant No. R01 DC006871 from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health. The authors thank Leigh Smitley for help with data analysis.

References

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

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

[3] E. Jacewicz, R. A. Fox, J. Salmons, "Vowel duration in three American English dialects", *American Speech* 82, 367-385 (2007)

[4] E. Jacewicz, R. A. Fox, J. Salmons, "Vowel space areas across dialects and gender", *Proceedings of the XVIth Int. Congress of Phonetic* Sciences, J. Trouvain and W. J. Barry (Eds.), pp. 1465-1468, Saarbrücken, Germany (2007)

[5] H. K. Vorperian, R. D. Kent, "Vowel acoustic space development in children: a synthesis of acoustic and anatomic data", *J. Speech, Lang., and Hearing Res.* 50, 1510-1545 (2007)