Analysis of Total Vowel Space Areas in Three Regional Dialects of American English

Robert A. Fox & Ewa Jacewicz
Speech Perception and Acoustic Laboratories
Speech and Hearing Science
The Ohio State University, USA

Acoustics'08 Paris
June 29 - July 4, 2008
euronoise

SPA
speech perception & acoustics labs

Acknowledgment

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.

SPA
speech perception
& acoustics labs

Acoustic Variation in Vowel Spaces

- In many languages (such as American English and Dutch) there is considerable variation in the acoustic characteristics of vowels across speakers from different geographical regions
- The majority of studies examining these variations have been primarily interested in the positions of the vowels in the acoustic vowel space and phonetic quality.
- However, what truly constitutes and defines the "vowel space" and why might the concept be important? In our view the "vowel space" is best conceived as the "working area" of a speaker's acoustic vowel space, that is, those areas in the acoustic vowel space which a speaker actually uses in producing vowel distinctions. This might be quite different from the vowel space area in terms of the positions of the corner vowels measured at their midpoints.
- If there are cross-dialectal differences in the vowel spaces used, they may potentially provide insights into sound change in progress, such as ongoing vowel chain shifts (e.g., Northern Cities Shift or Southern Vowel Shift in contemporary American English).

SPA speech perception & acoustics labs

Specific Questions Addressed

- What is the most appropriate way to characterize the vowel space and/or measure its area (i.e., in terms of the traditional 4 corner vowels or the borders of the vowel space)?
- Placement of vowels in the F1 by F2 plane varies as a function of dialect; do the vowels space areas of these dialects also vary systematically?
- Does vowel space area vary significantly and/or systematically across different generations of speakers (reflecting language change/vowel shifts)?
- Vowels often change their formant frequency values (and position in the F1 by F2 plane) over time. How does the vowel space area change dynamically and should the vowel space be measured in terms of the formant changes made by vowels across their durations?
- If there is systematic variation in vowel space size across dialects (however measured), will formant normalization schemes (commonly used to remove speaker variation) eliminate these dialectal differences?

SPA speech perception & acoustics labs

Dialects to be Examined

- We here compare vowel space areas among the American English dialects spoken in:
  - **South-central Wisconsin** (undergoing the Northern Cities Shift)
  - **Central Ohio** (no known vowel shift although there are specific changes currently underway)
  - **Western North Carolina/Appalachian English** (affected by the Southern Vowel Shift)
- The participants of the study were 54 women who grew up and continued to live in very narrowly defined locations in these three dialectal regions (2-3 counties wide). In each regional group, there were nine young women aged 20-34 years and nine older women aged 51-65 years. Each group represented a very homogeneous speaker set.

SPA speech perception & acoustics labs

Speech Materials

- Speech materials consisted of the following single words of the /hVd/ structure produced by each speaker three times: *heed*, *hid*, *head*, *hey'd*, *had*, *hawed*, *hod*, *who'd*, *hood*, *hoed*, *heard*, *hide*, *howed*, *hoyd.*

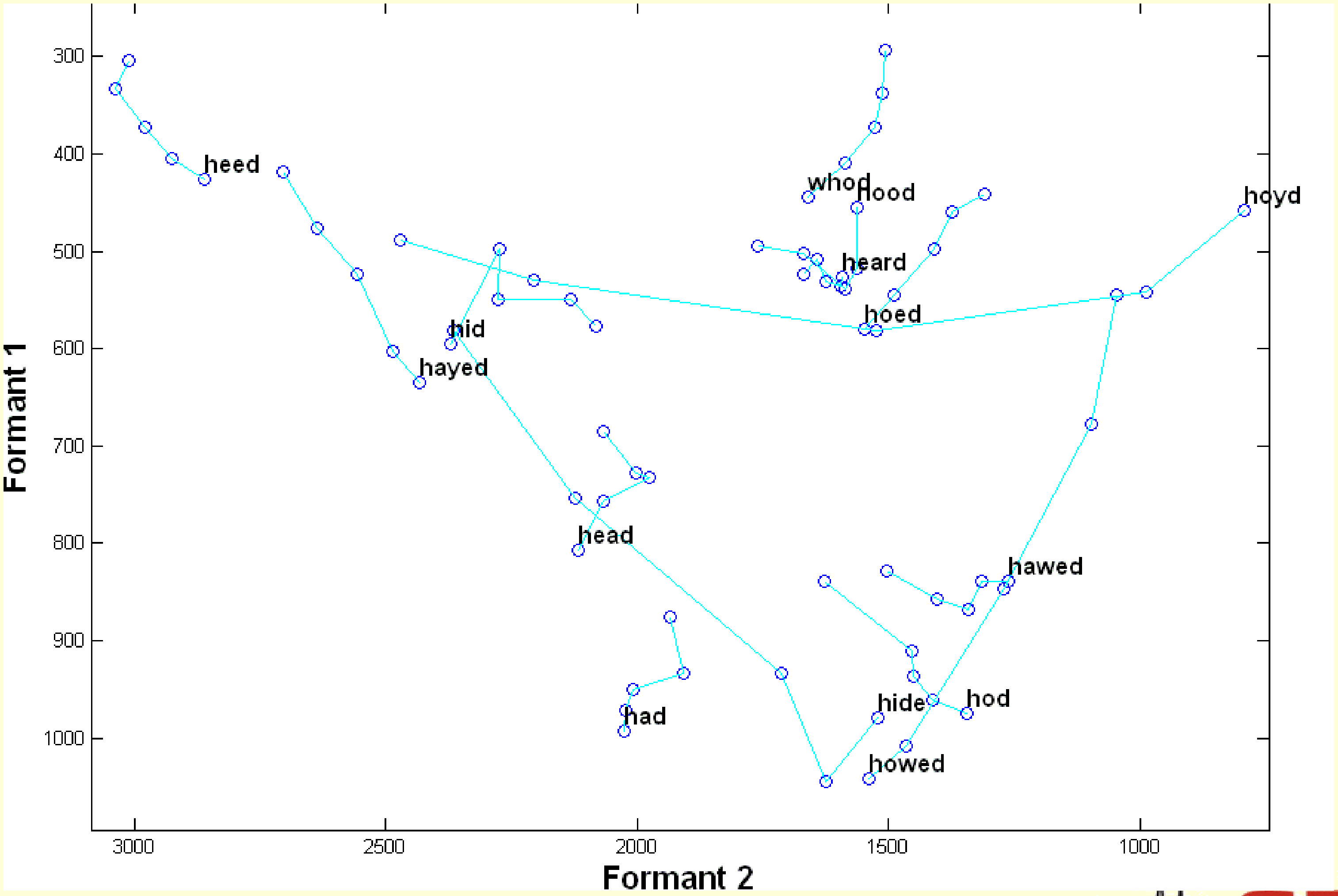
  The words contained 14 vowels and diphthongs of American English: /i, ɪ, ɛ, e, æ, ɔ, ɑ, u, ʊ, o, ɚ, aɪ, aʊ, ɔɪ/.
- Words were presented in random order on a computer screen to a subject seated in a sound-attenuating booth and were recorded directly onto a hard drive disk at a 44.1-kHz sampling rate using head-mounted microphone positioned at a distance of 2 inches from the speaker's lips.

Acoustic Measurements

- All tokens were digitally filtered and downsampled to 11.025 kHz and pre-emphasized (98%) prior to analysis.
- Onsets and offsets were located by hand from the waveform (with reference to a spectrogram) and the overall vowel duration was calculated.
- Formant frequency values were then extracted automatically using a custom program in Matlab (a 14-order LPC analysis with a 25-ms Hamming window) and hand-corrected.
- The frequencies of F1, F2, and F3 were measured at five equidistant temporal locations corresponding to the 20-35-50-65-80% points over the course of each vowel's duration—this allows an examination of formant trajectories over time, estimating the dynamic formant change for each vowel.
- Next are exemplar vowels from 3 different speakers, first Ohio.

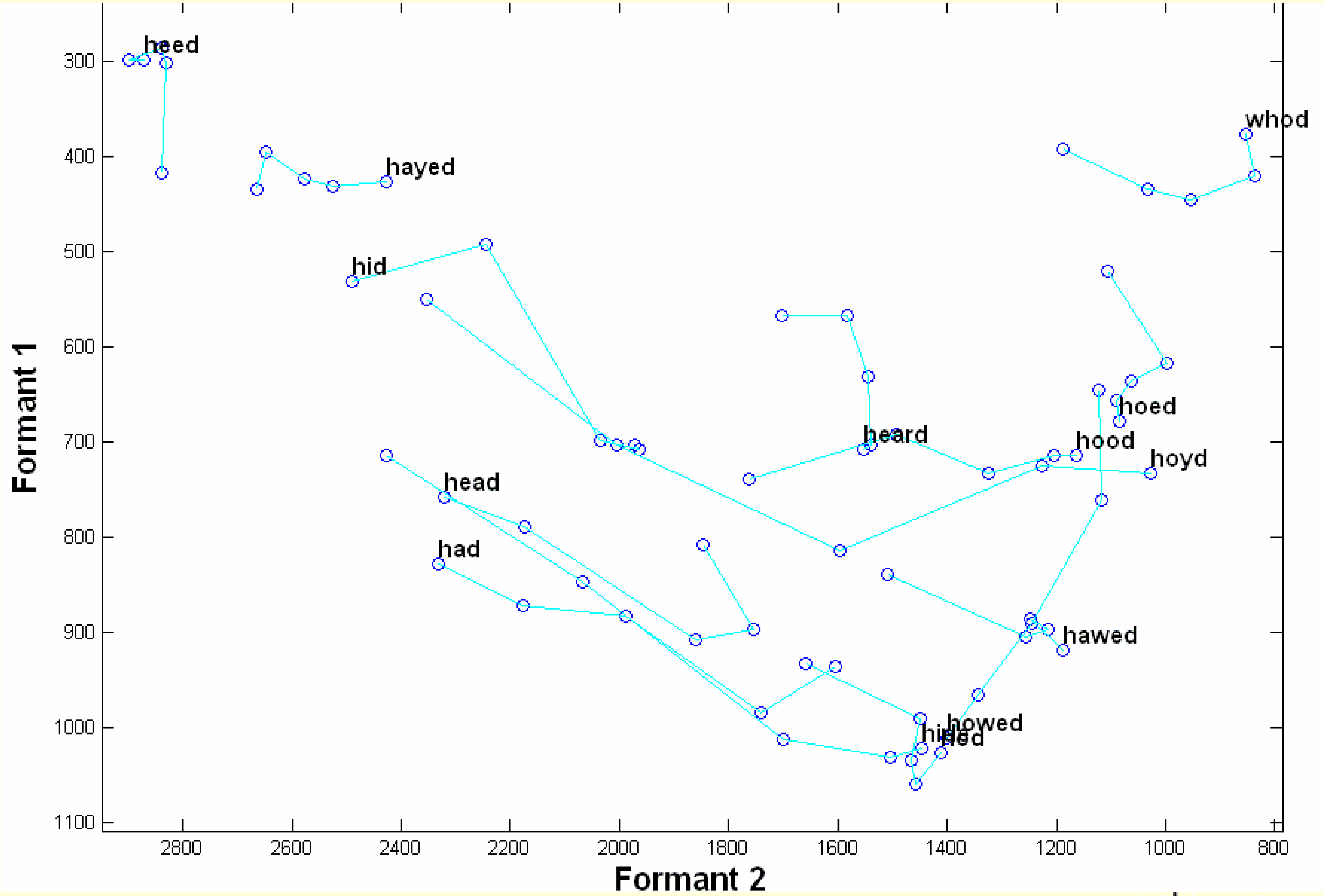
SPA
speech perception
& acoustics labs

Exemplar Vowel Space for Ohio Dialect

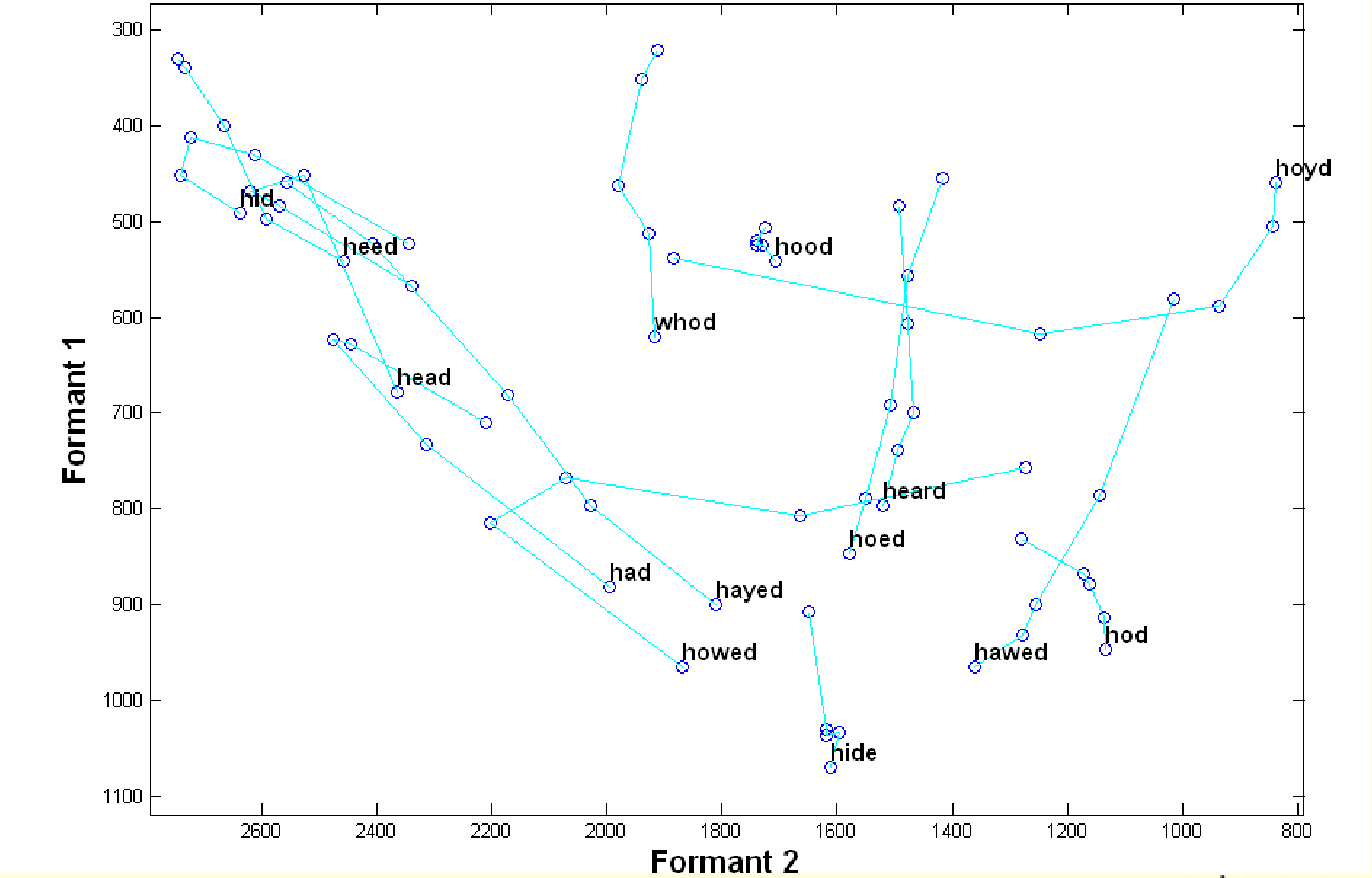


SPA speech perception & acoustics labs

Exemplar Vowel Space for Wisconsin Dialect



Exemplar Vowel Space for North Carolina Dialect



speech perception & acoustics labs

Calculation of Vowel Quadrilateral Area

- Vowel space areas in the F1 by F2 plane at each of the five temporal locations were calculated on the basis of each speaker's *mean* formant values in two ways.
- First, the vowel space area in the F1 by F2 plane was calculated on the basis of the four "corner" vowels /i, u, ɑ, æ/. This polygon 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:

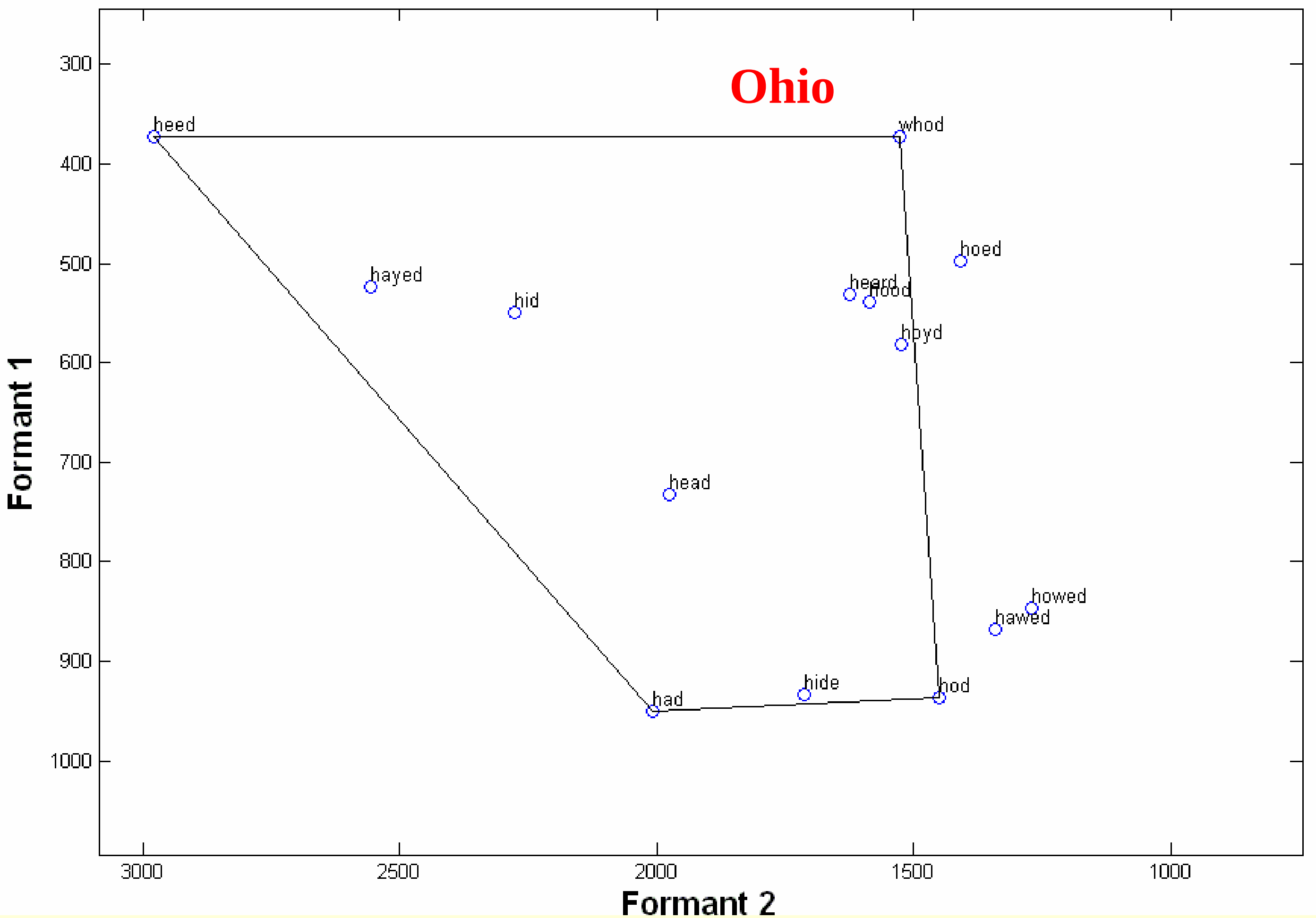
area of triangle = sqrt (s(s-a)(s-b)(s-c))

where *s* = the vowel perimeter *(a + b + c)*

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

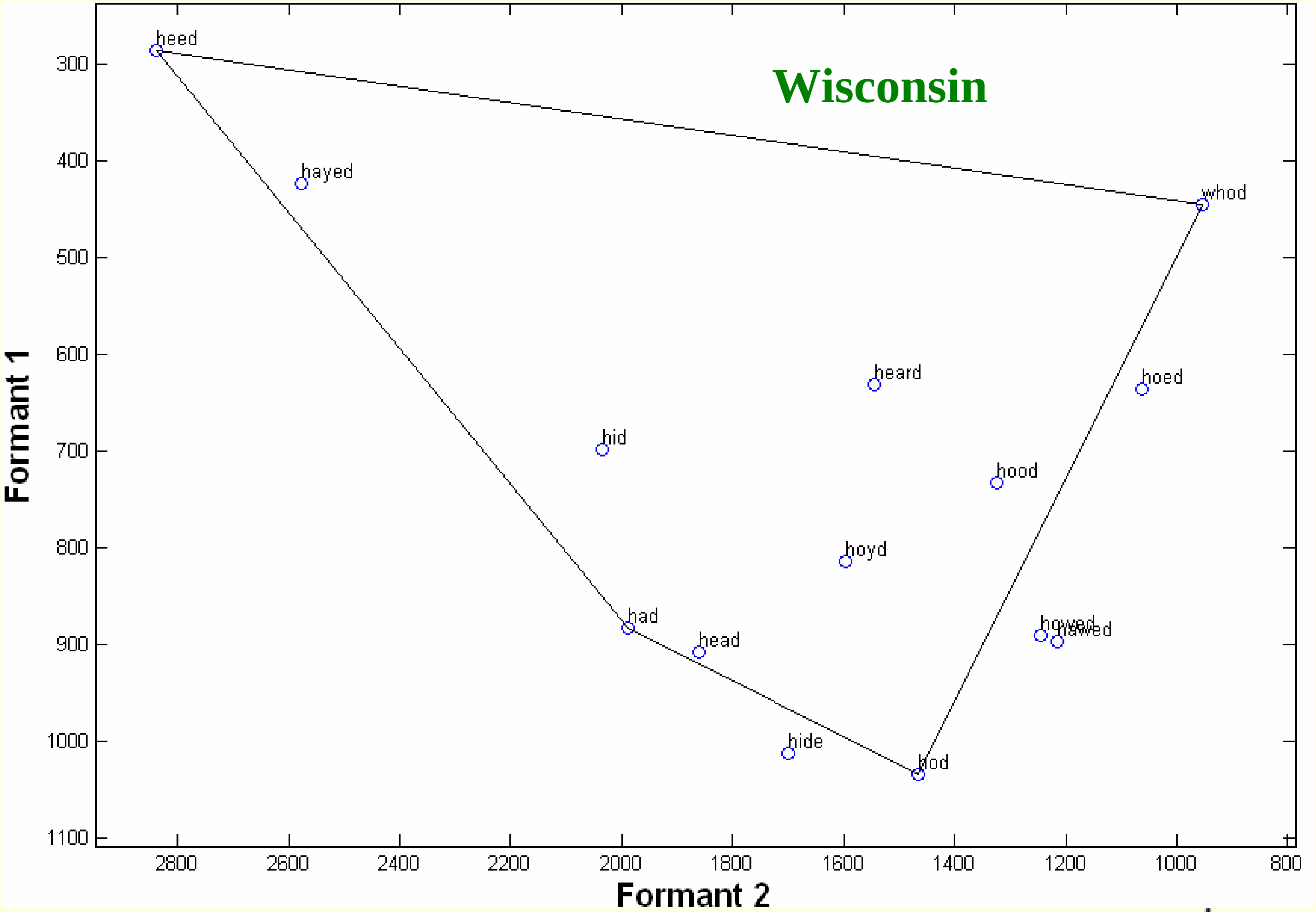
The vowel quadrilateral area was then determined by adding the areas of these two triangles.

Exemplar Ohio Vowel Quadrilateral at Vowel Midpoint

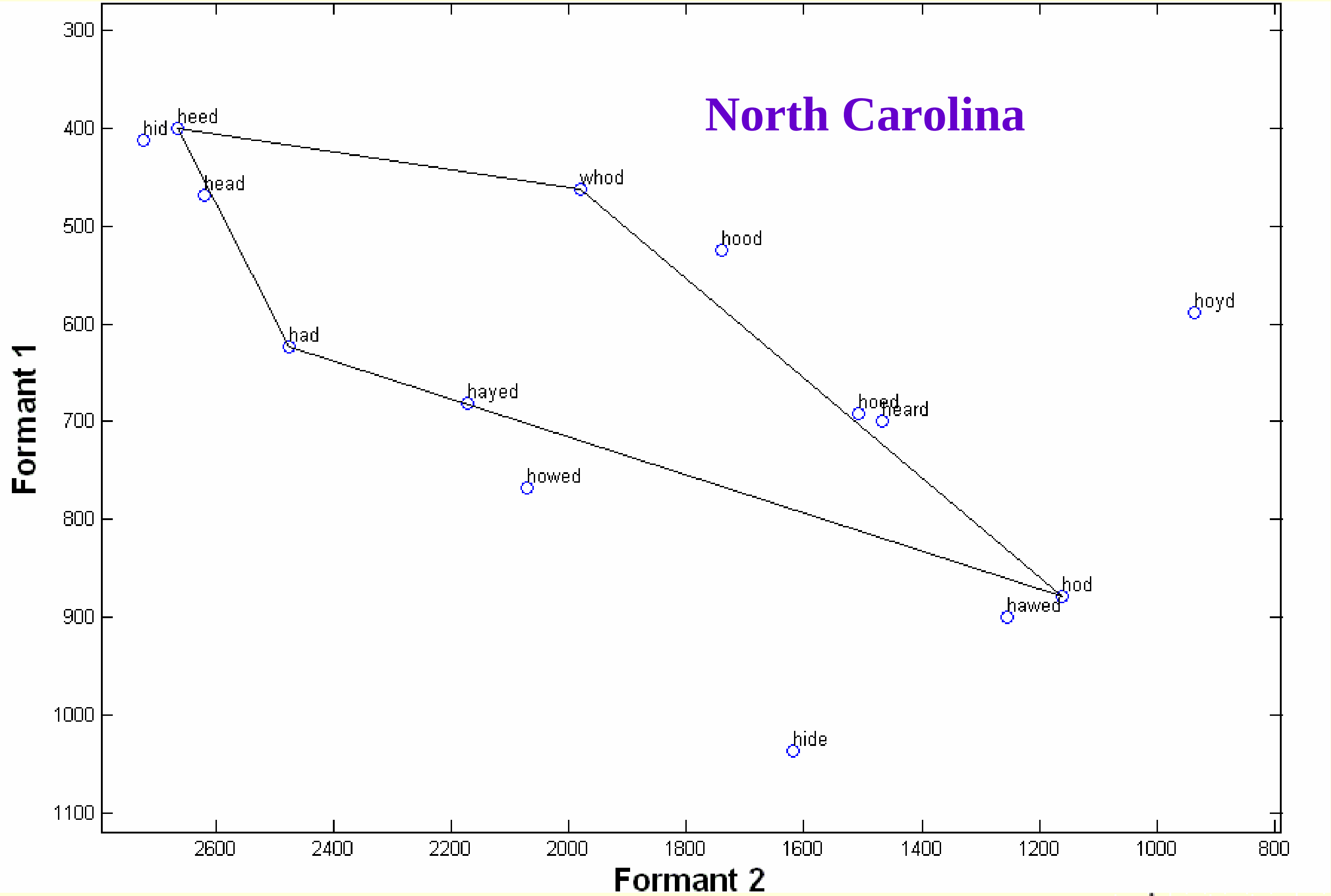


SPA
speech perception
& acoustics labs

Exemplar Wisconsin Vowel Quadrilateral at Vowel Midpoint



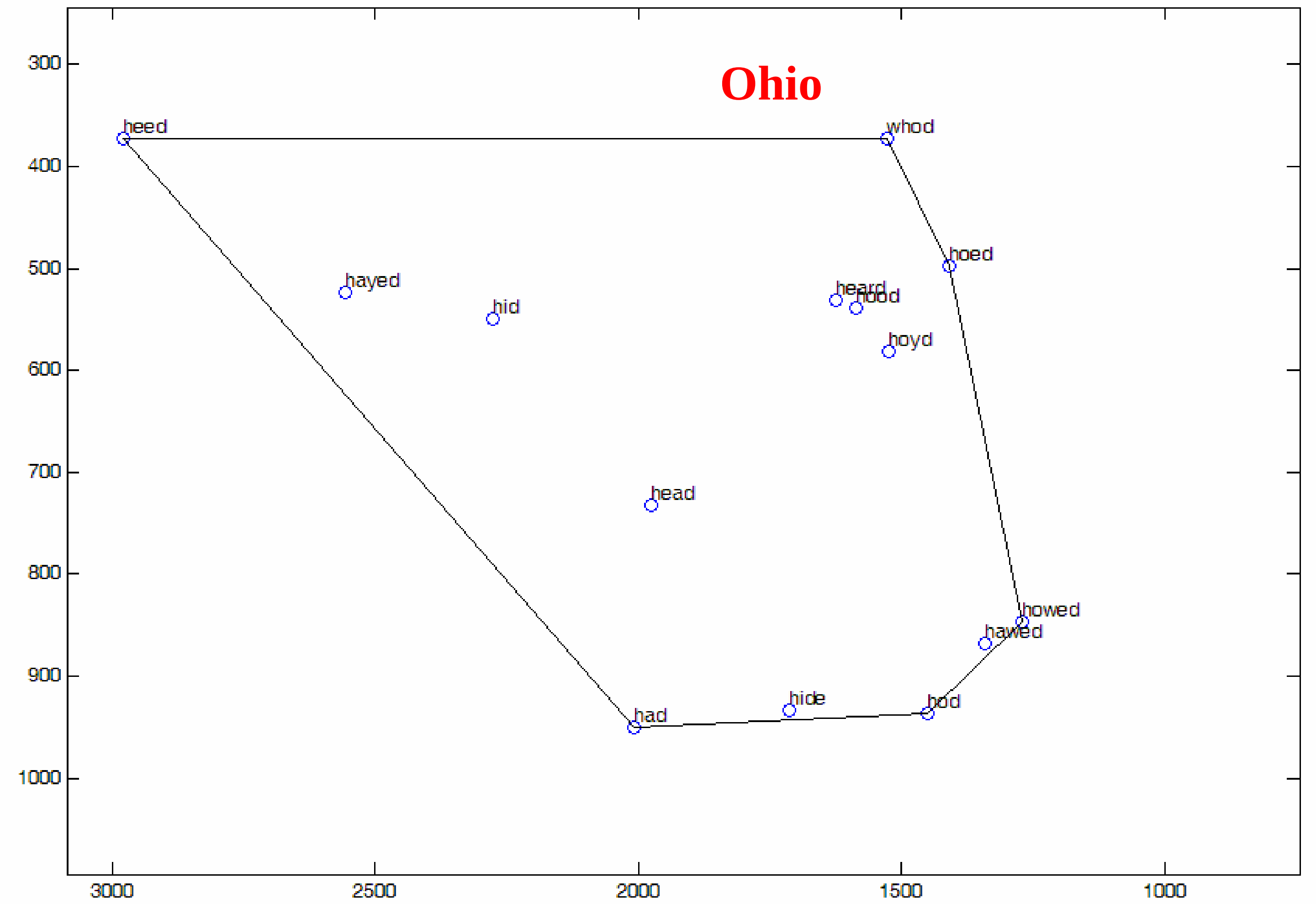
Exemplar North Carolina Vowel Quadrilateral at Vowel Midpoint



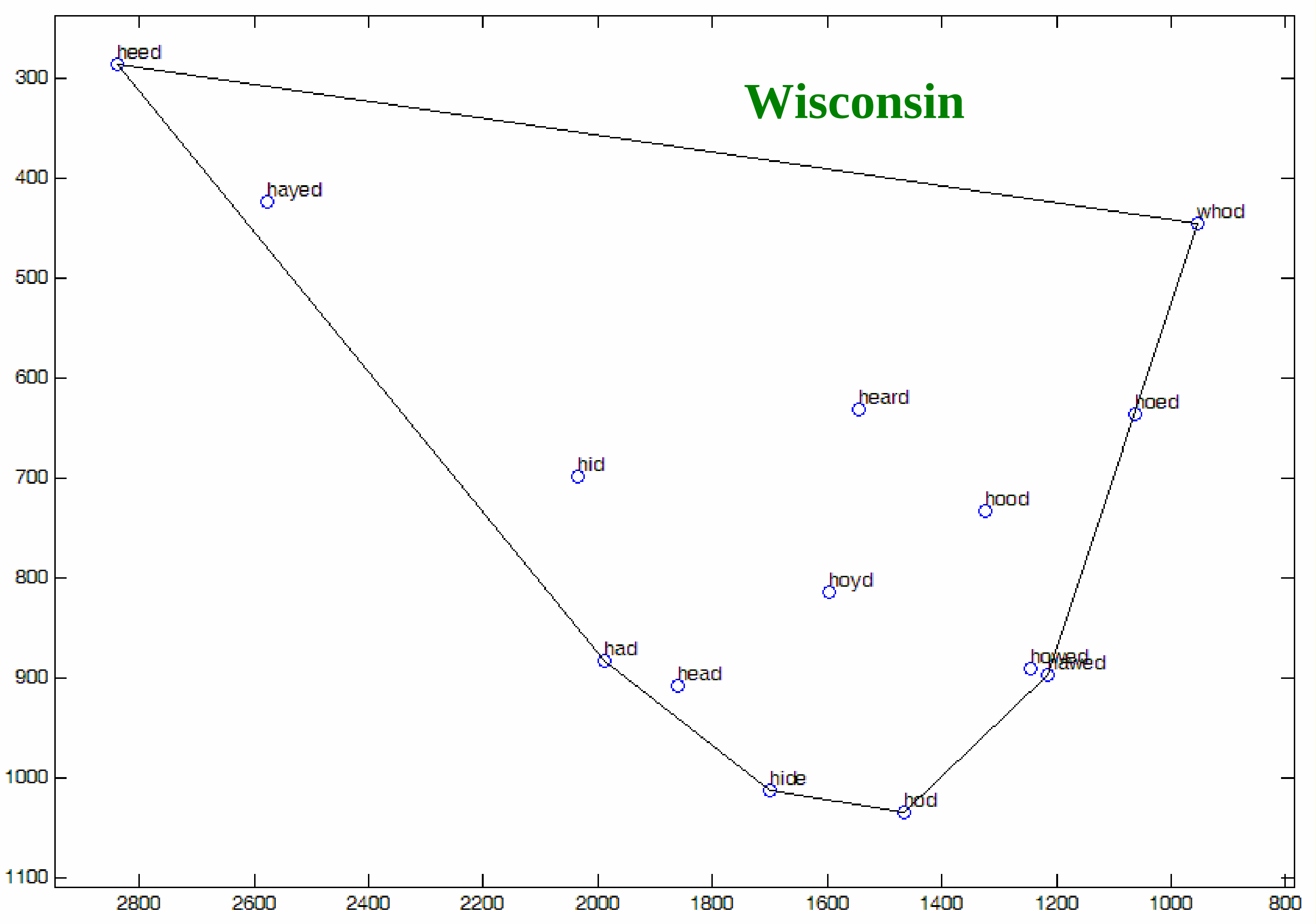
Calculation of Vowel Multilateral (Representing Maximum Vowel Space Area)

- Next, the overall vowel space area at each of the five temporal locations was calculated for each speaker. This vowel space was defined by the perimeter vowels in the F1 by F2 plane referred here as the **"vowel multilateral."** These areas were calculated using Matlab's polyarea function.

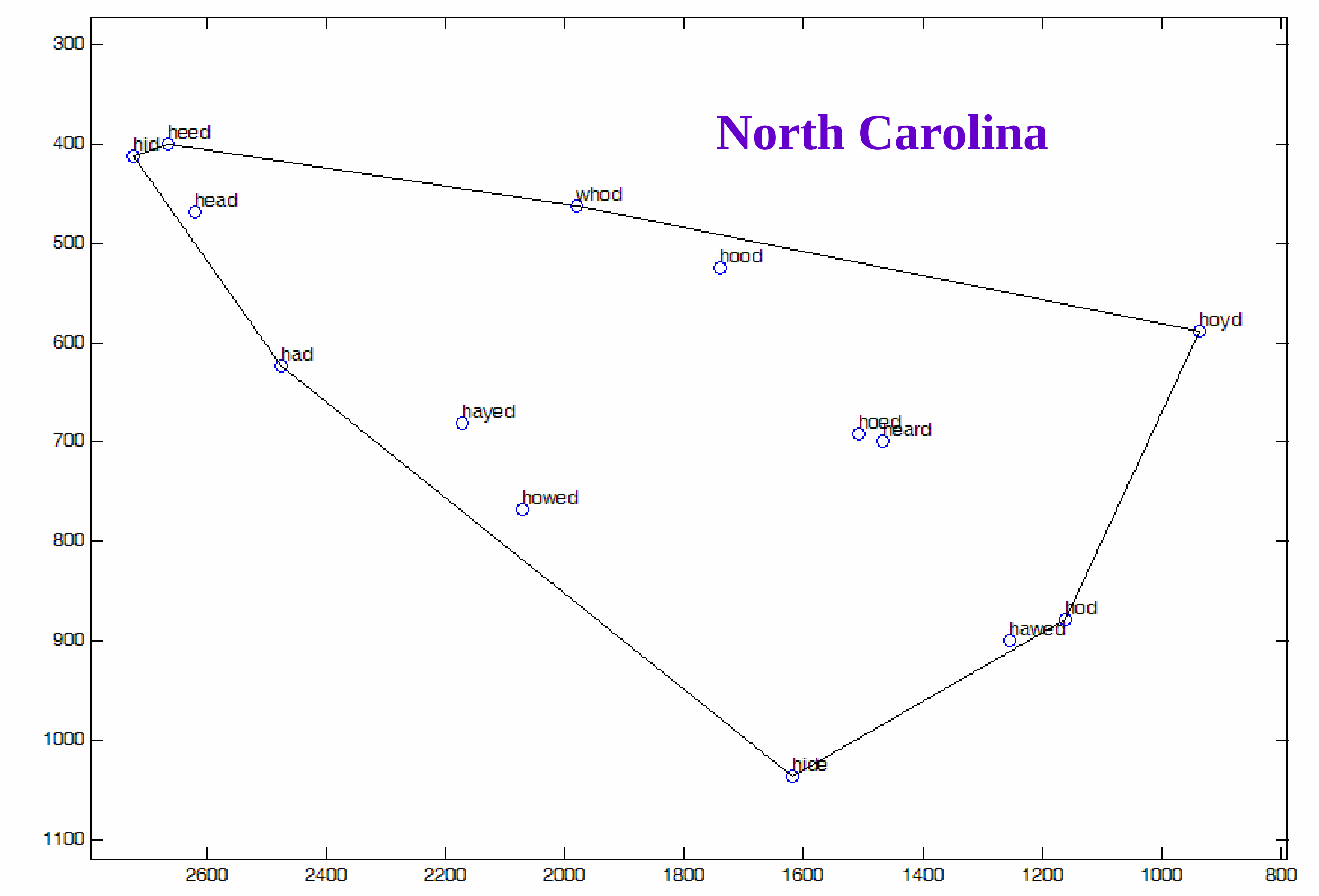
Exemplar Ohio Vowel Multilateral at Vowel Midpoint



Exemplar Wisconsin Vowel Multilateral at Vowel Midpoint



Exemplar North Carolina Vowel Multilateral at Vowel Midpoint



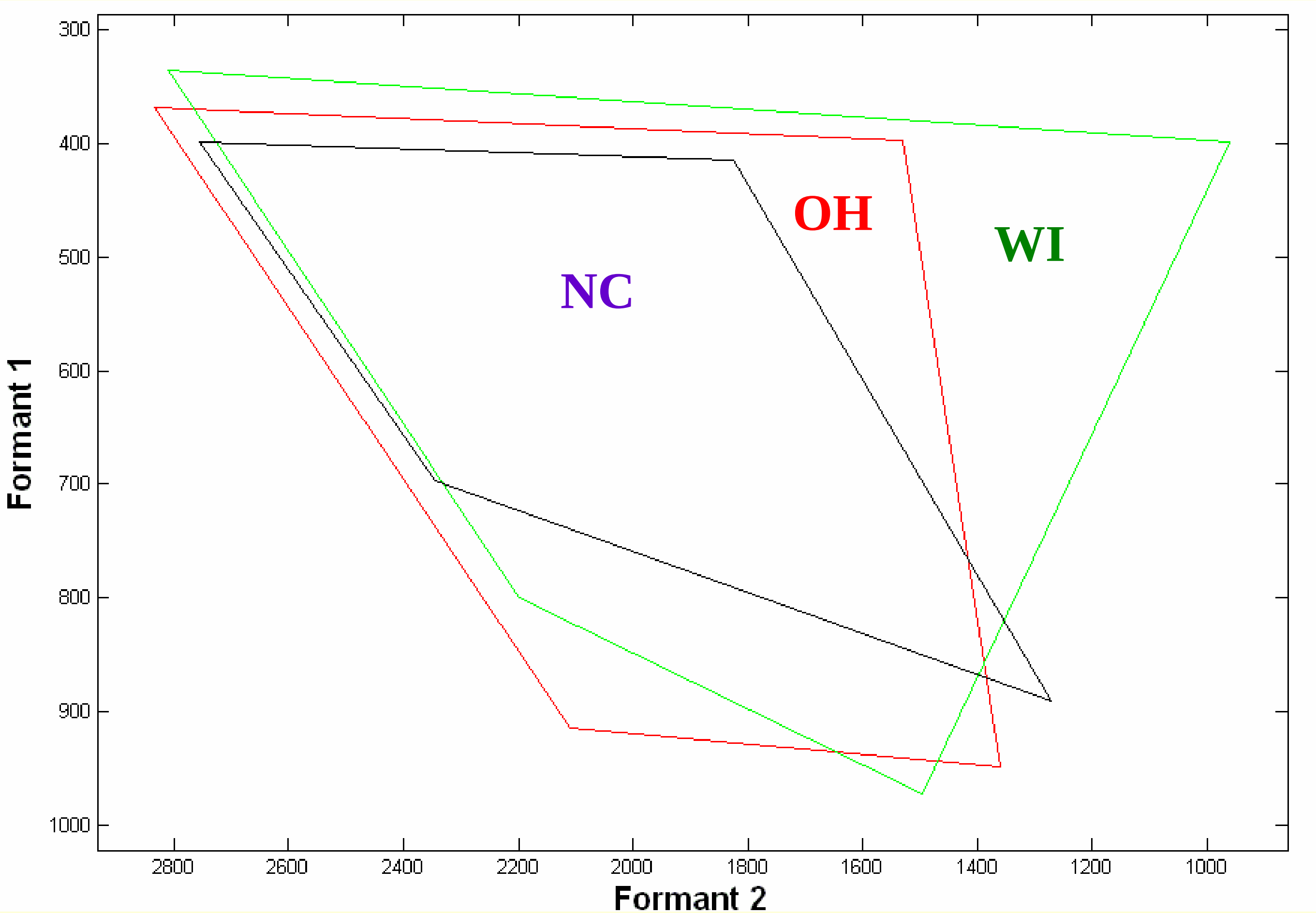
speech perception & acoustics labs

Analyses of Quadrilateral and Multilateral Vowel Space Areas

- Separate analyses were conducted for both the quadrilateral and multilateral areas. These analyses included mixed-design ANOVAs with the within-subject factor ***temporal location*** (20%, 35%, 50%, 65% and 80% points) and the between-subject factors ***dialect*** (OH, WI and NC) and ***age*** (young adults and older adults).
- First the quadrilateral analysis.

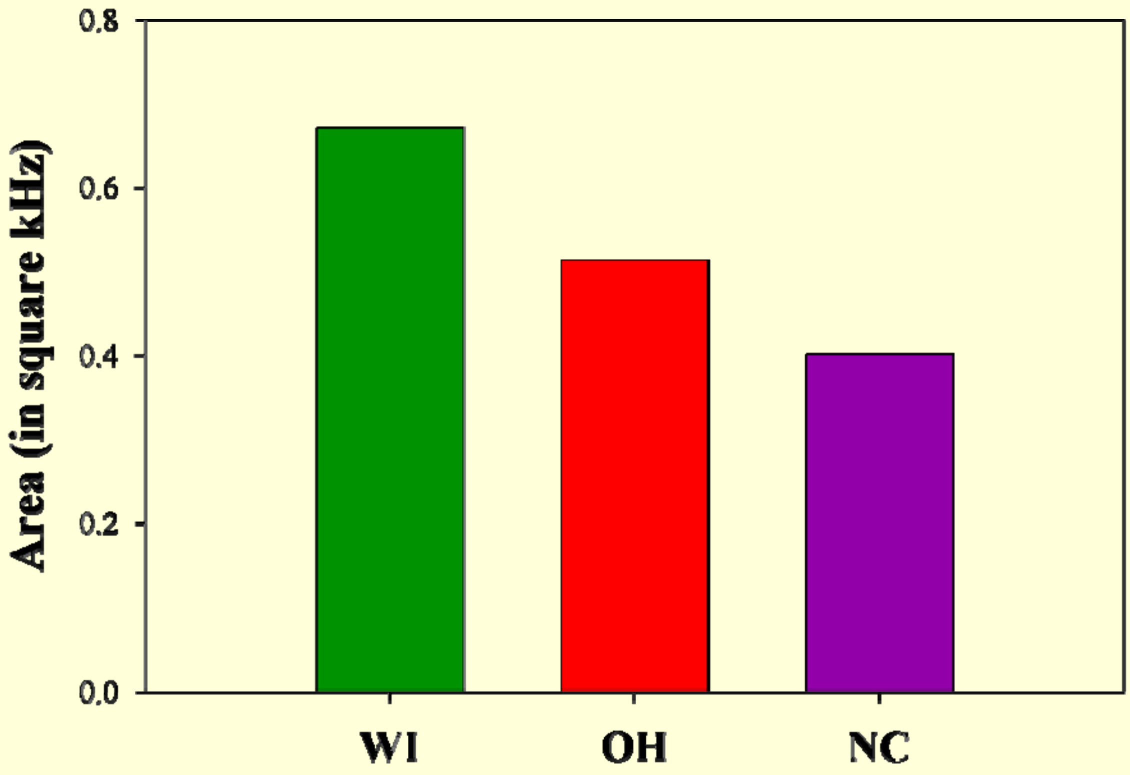
speech perception & acoustics labs

Differences in Mean Vowel Quadrilaterals (at Vowel Midpoint) Among Dialects



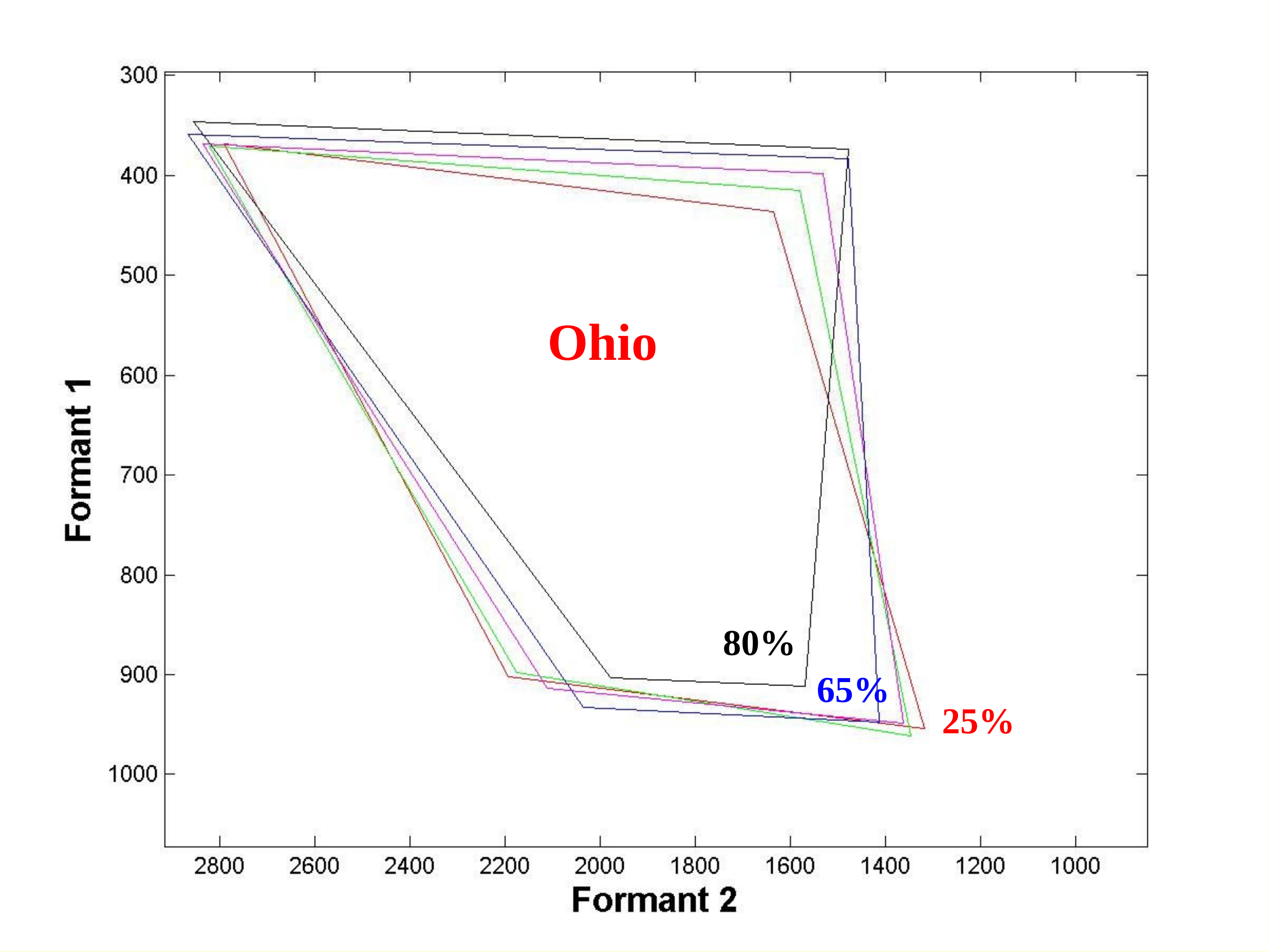
Dialect and Quadrilateral Area

There was a significant difference in the size of the quadrilateral areas as a function of dialect ($F(2,51)=18.9$, $p<.001$, $\eta^2=0.440$). Wisconsin had the largest area (.670 kHz^2) followed by Ohio (.513 kHz^2) and North Carolina (.401 kHz^2). Tukey tests indicated each pair were significant different.

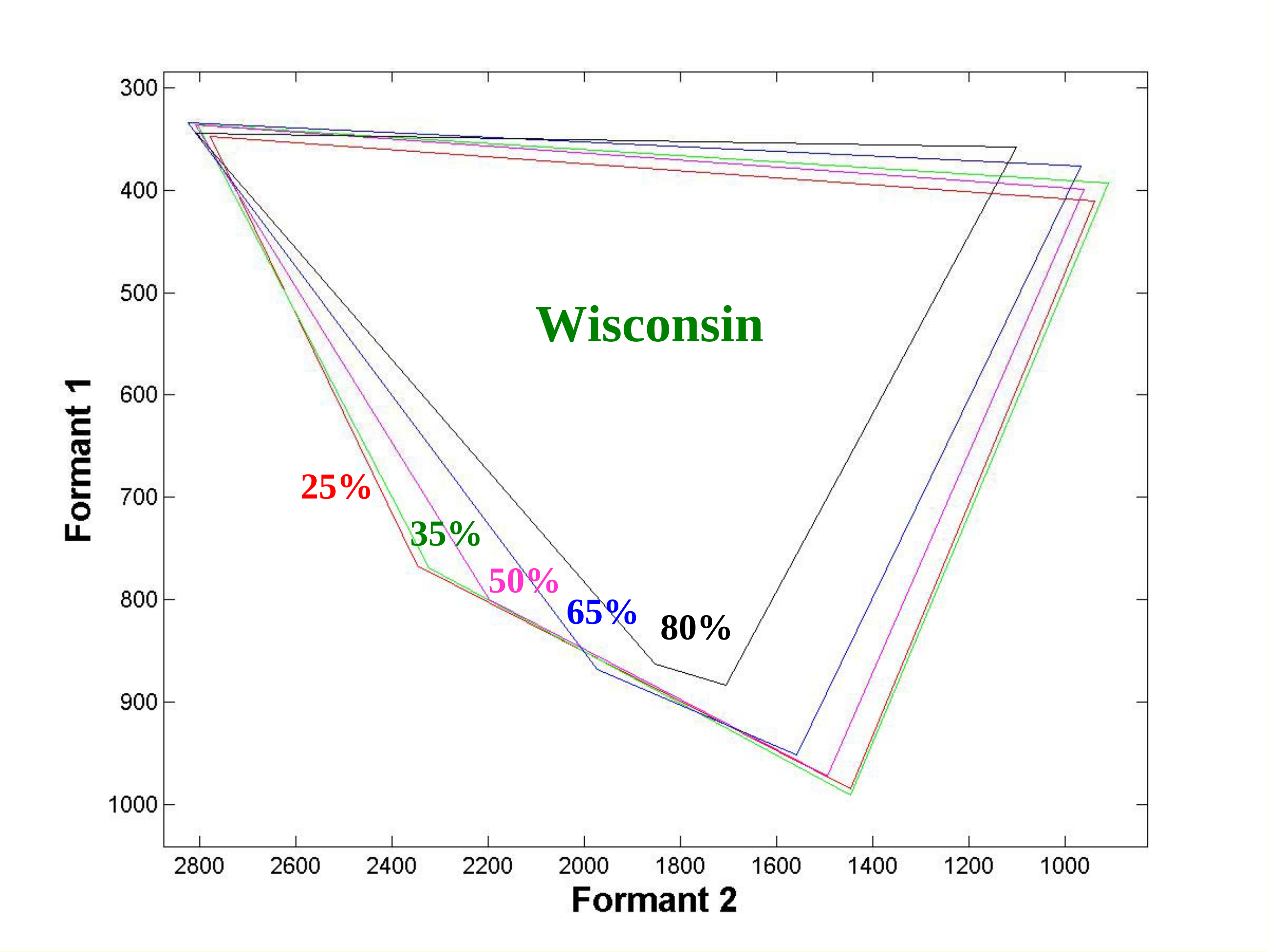


The primary reason for these significant differences is clear from a review of the acoustic spaces (displayed earlier). For example, the /u/ vowel in the OH and NC is fronted while the /u/ in the Wisconsin vowel space is retracted.

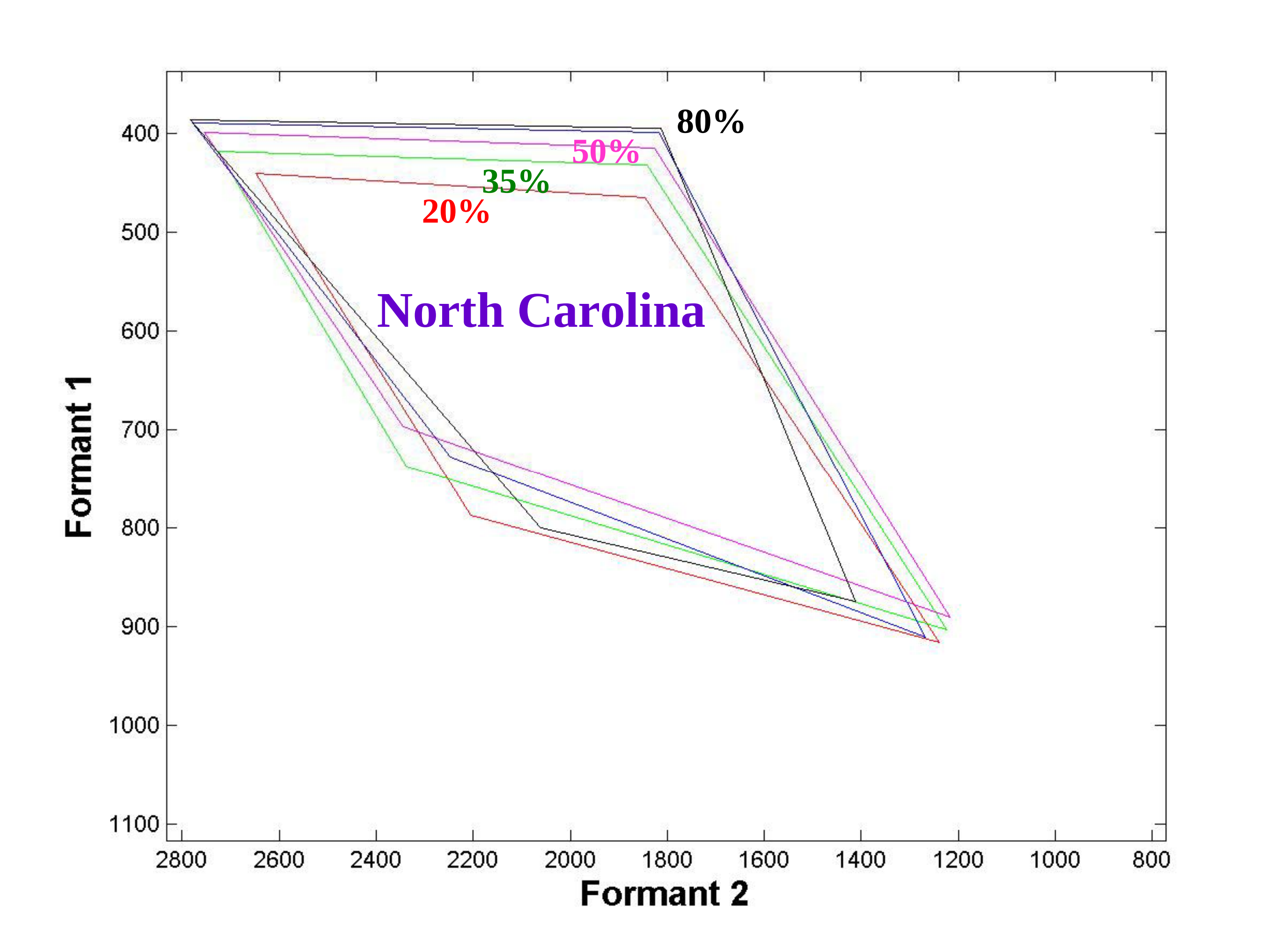
Differences in Mean Vowel Quadrilaterals for Ohio Speakers across Temporal Locations



Differences in Mean Vowel Quadrilaterals for Wisconsin Speakers across Temporal Locations

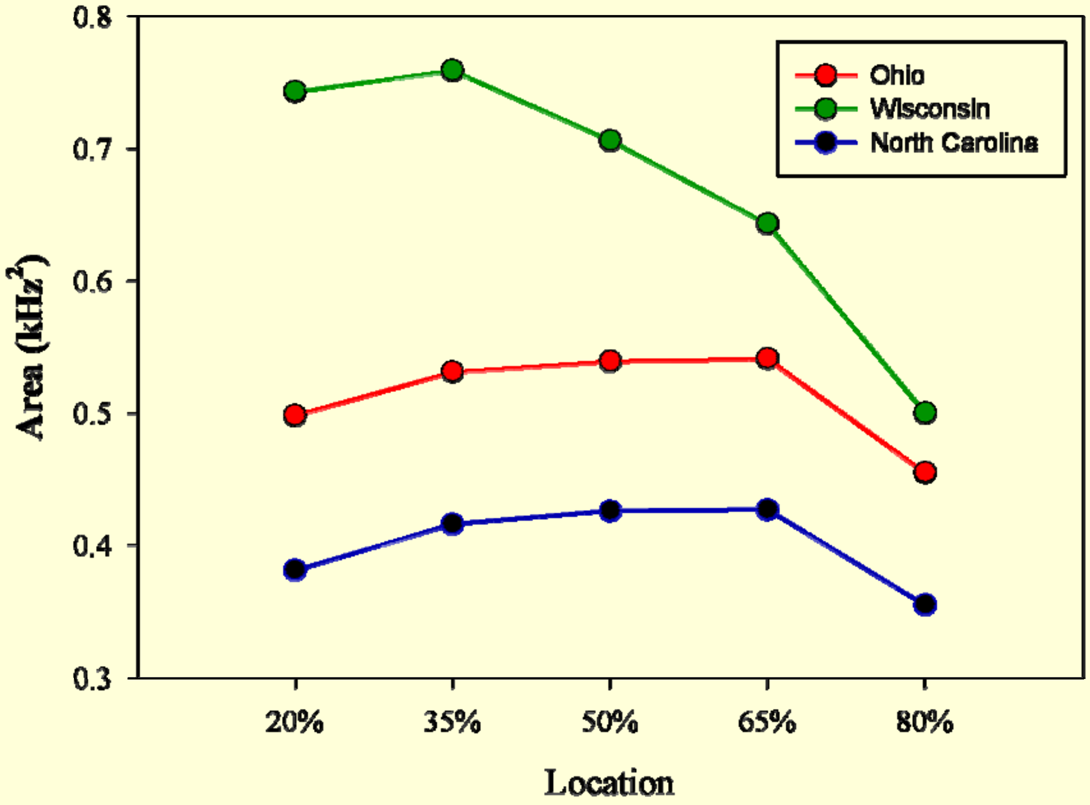


Differences in Mean Vowel Quadrilaterals for NC Speakers across Temporal Locations



Temporal Location and Quadrilateral Area

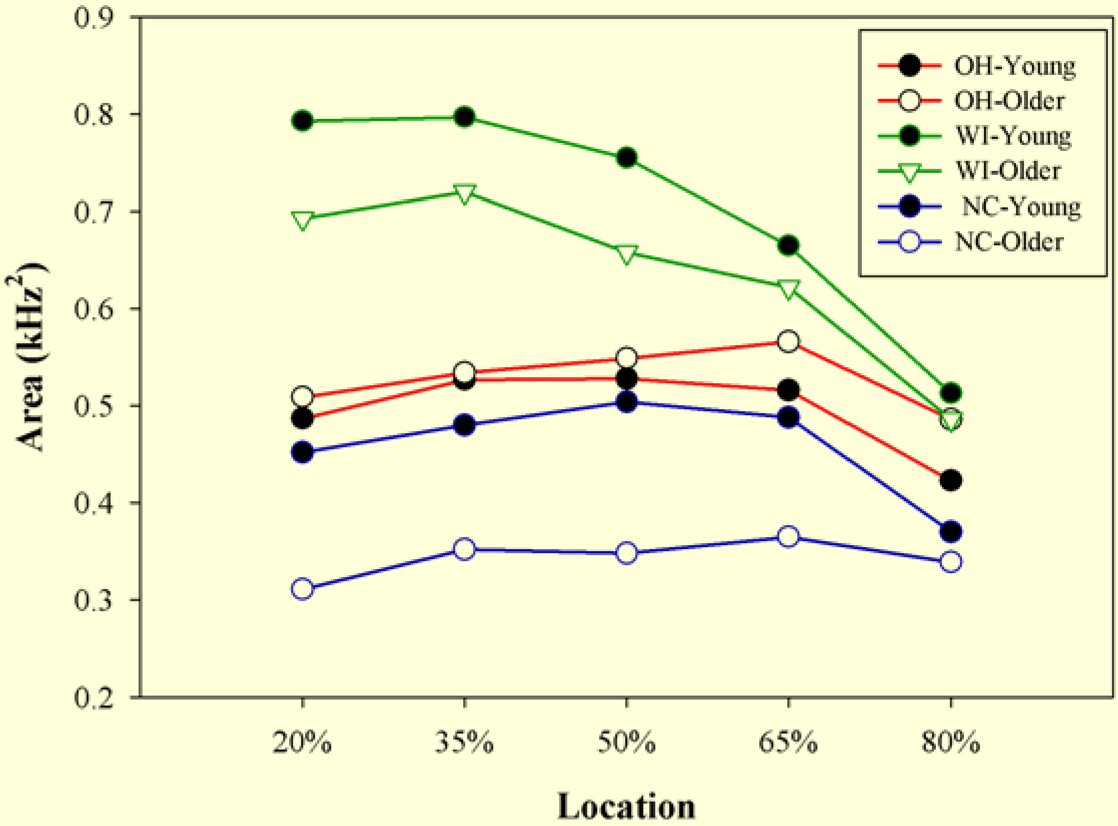
There was also a significant effect of temporal location ($F(4,192)=53.6$, $p<.001$, $\eta^2=0.527$). The area at the 80% point was significantly smaller. Changes in the areas for the OH and NC vowels were more stable across temporal locations than for WI vowels. The



quadrilateral area for the WI speakers dropped significantly from the 35% to the final 80% location which produced a significant location by dialect interaction ($F(8,192)=15.3$, $p<.001$, $\eta^2=0.405$).

Age and Quadrilateral Area

There was no significant age effect ($F(1,48)=1.99$, n.s.) but there was a significant dialect x age interaction ($F(1,48)=5.05$, $p<.03$, $\eta^2=0.095$).

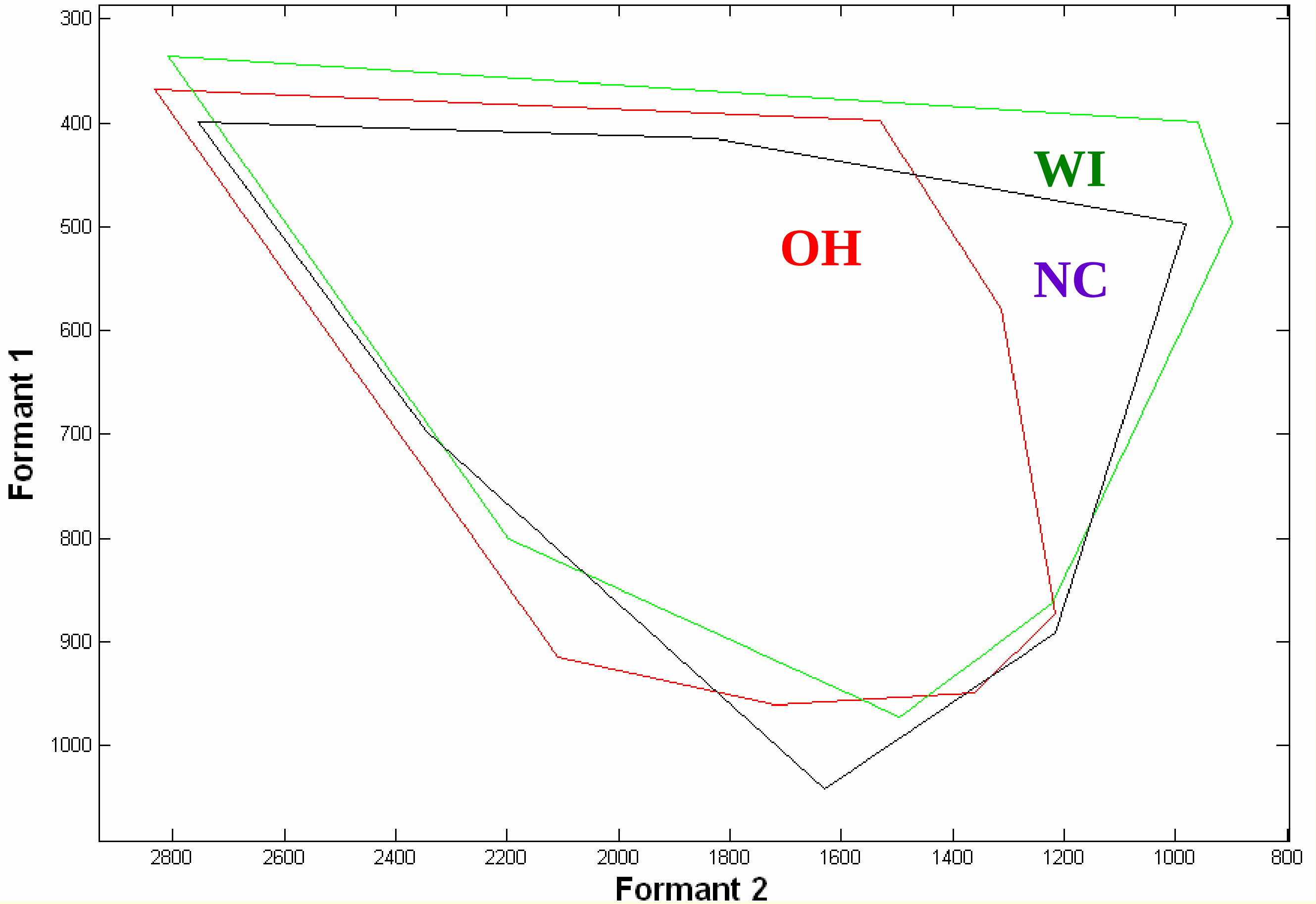


Young adult speakers, especially in the WI and NC dialect groups, had larger quadrilateral areas than did the older adult speakers in the 20%-65% temporal locations.

Analysis of Multilateral Vowel Space Area

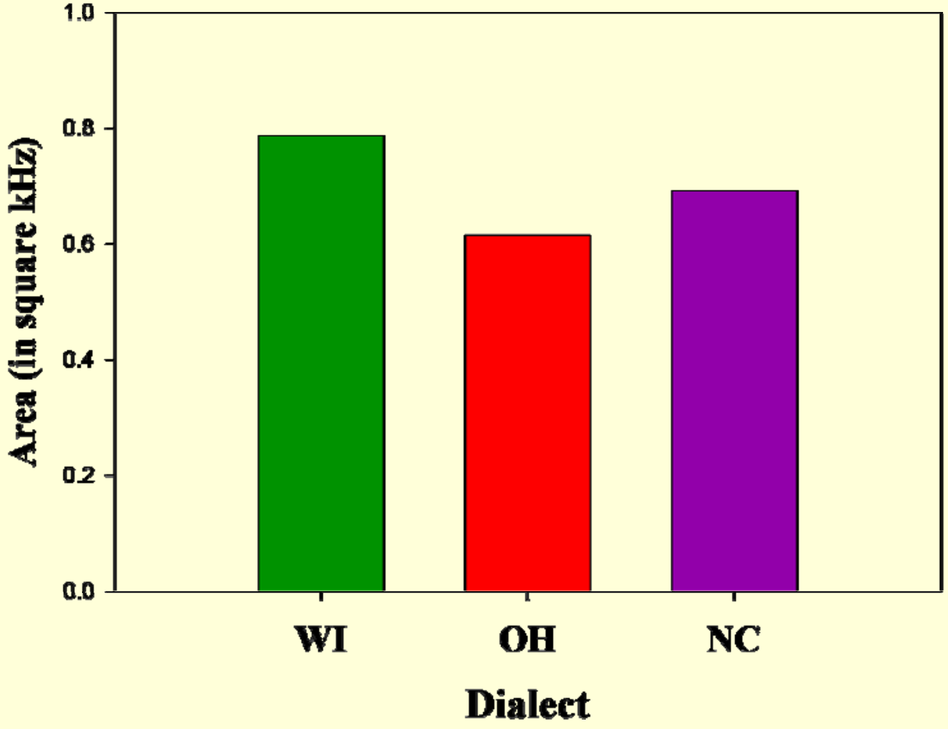
- Many vowels displayed in the graphs here fall, at least at some point in their production, outside the vowel quadrilateral defined by /i-u-ɑ-æ/. The quadrilateral is a drastic **underestimation** of the size of the working vowel space. We thus now turn to an analysis of the vowel multilaterals.

Differences in Mean Vowel Multilaterals (at Vowel Midpoint) Among Dialects

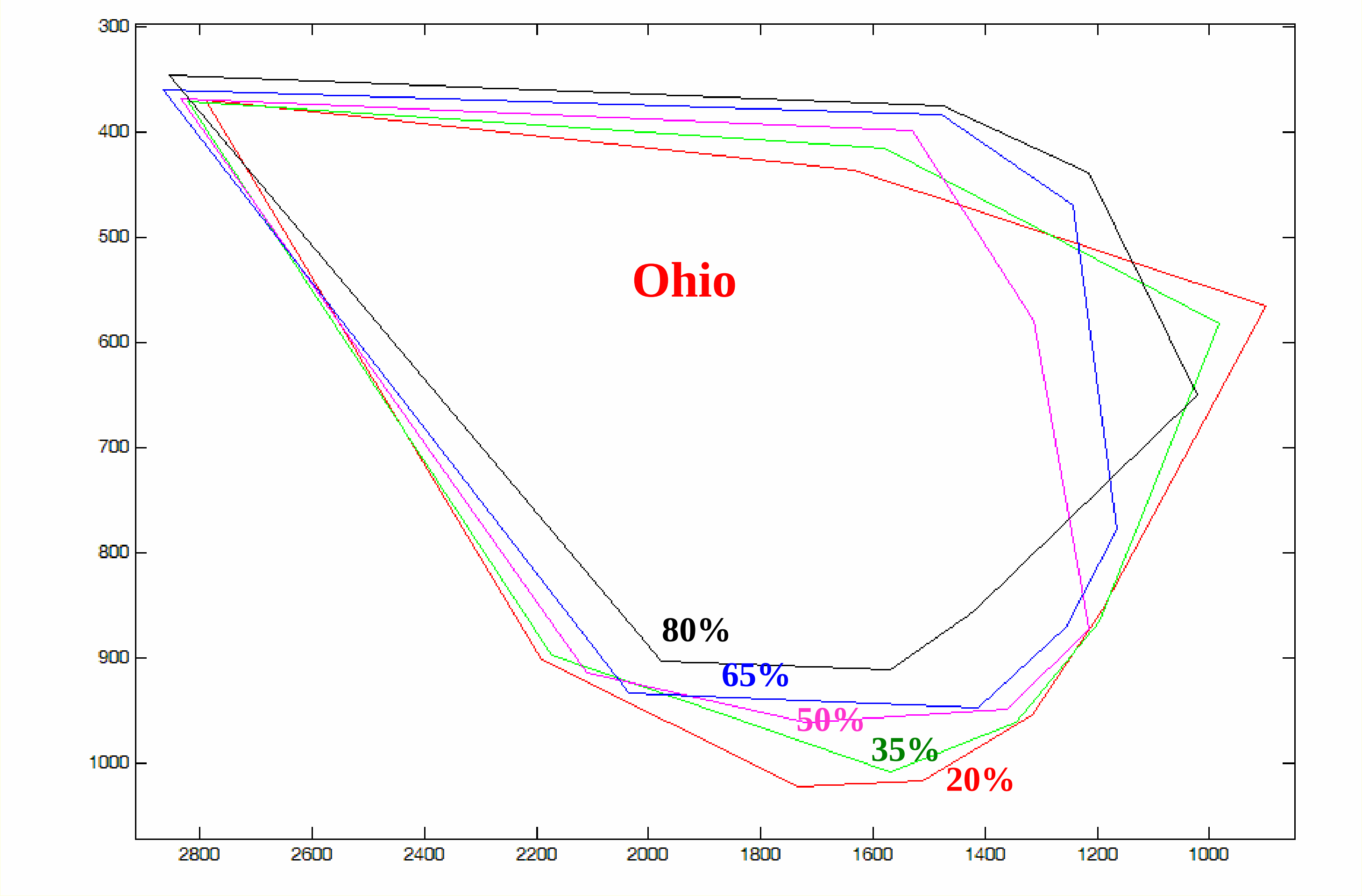


Dialect and Multilateral Area

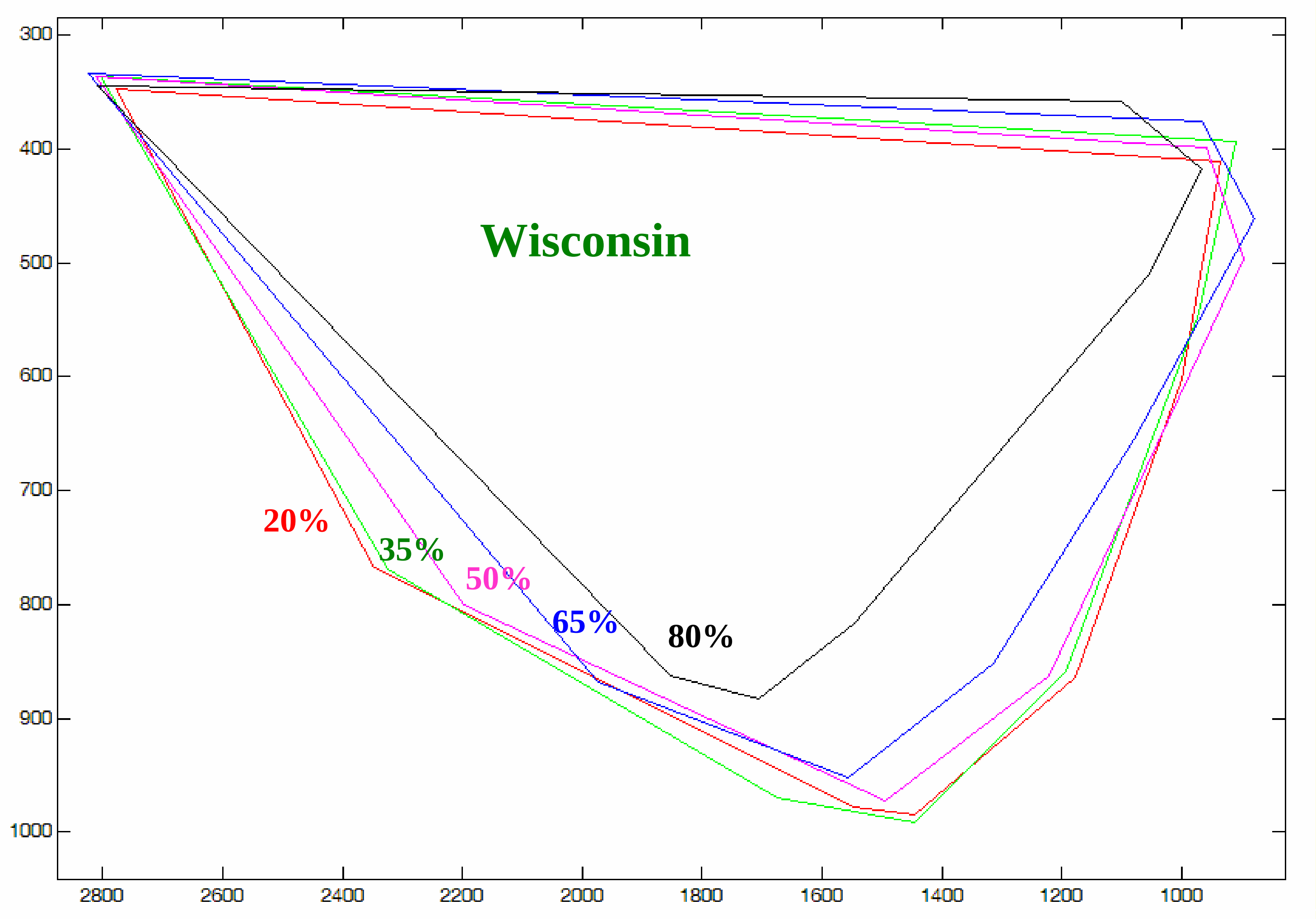
Similar to the pattern seen for the quadrilateral areas, the mean area of the total WI vowel space (.787 kHz^2) was again larger than that of either the OH (.614 kHz^2) or NC (0.691 kHz^2) vowel spaces, although the difference was just below significance ($F(2,48)=2.99$, $p=.06$).



Differences in Mean Vowel Multilaterals for Ohio Speakers across Temporal Locations

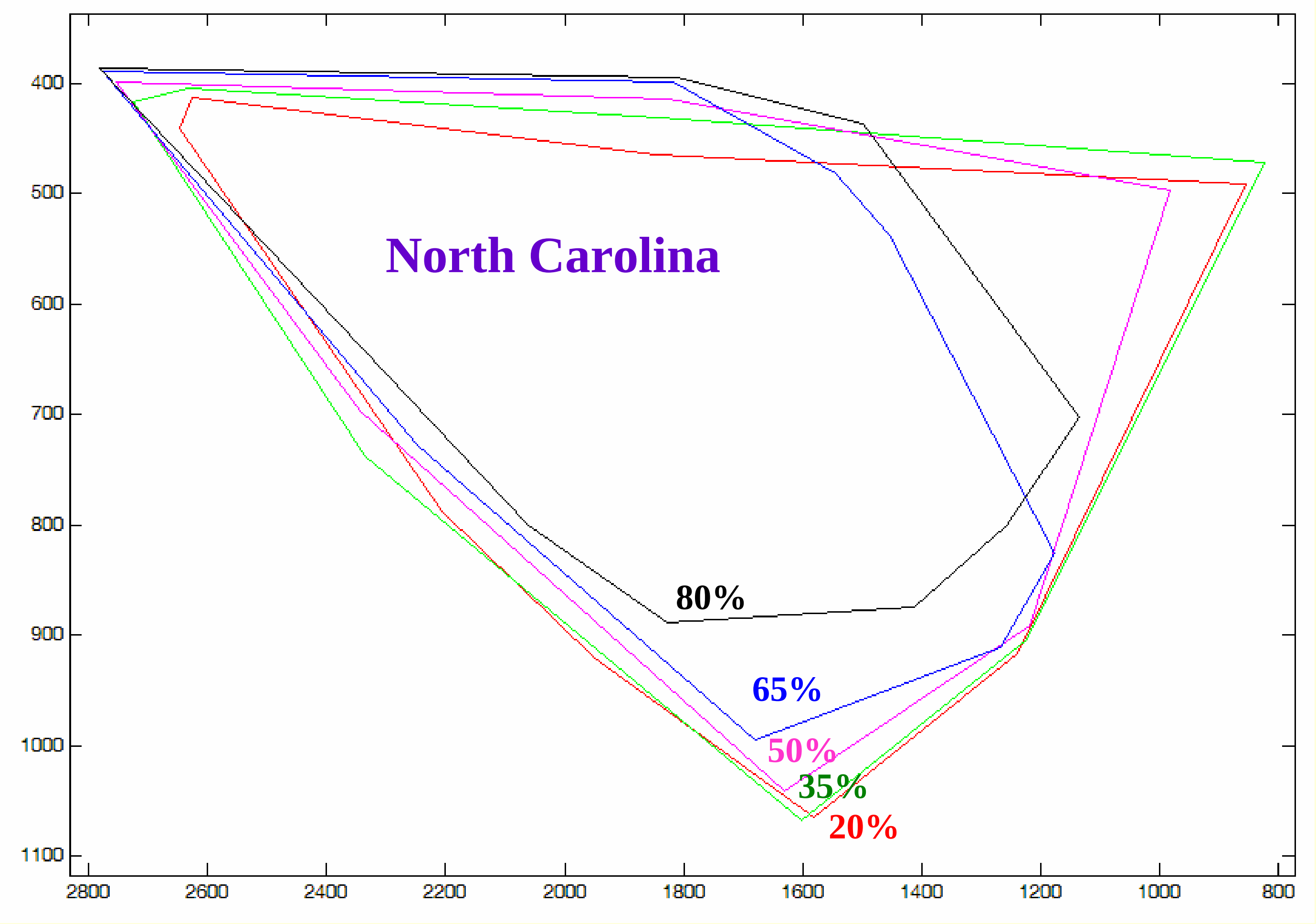


Differences in Mean Vowel Multilaterals for Wisconsin Speakers across Temporal Locations



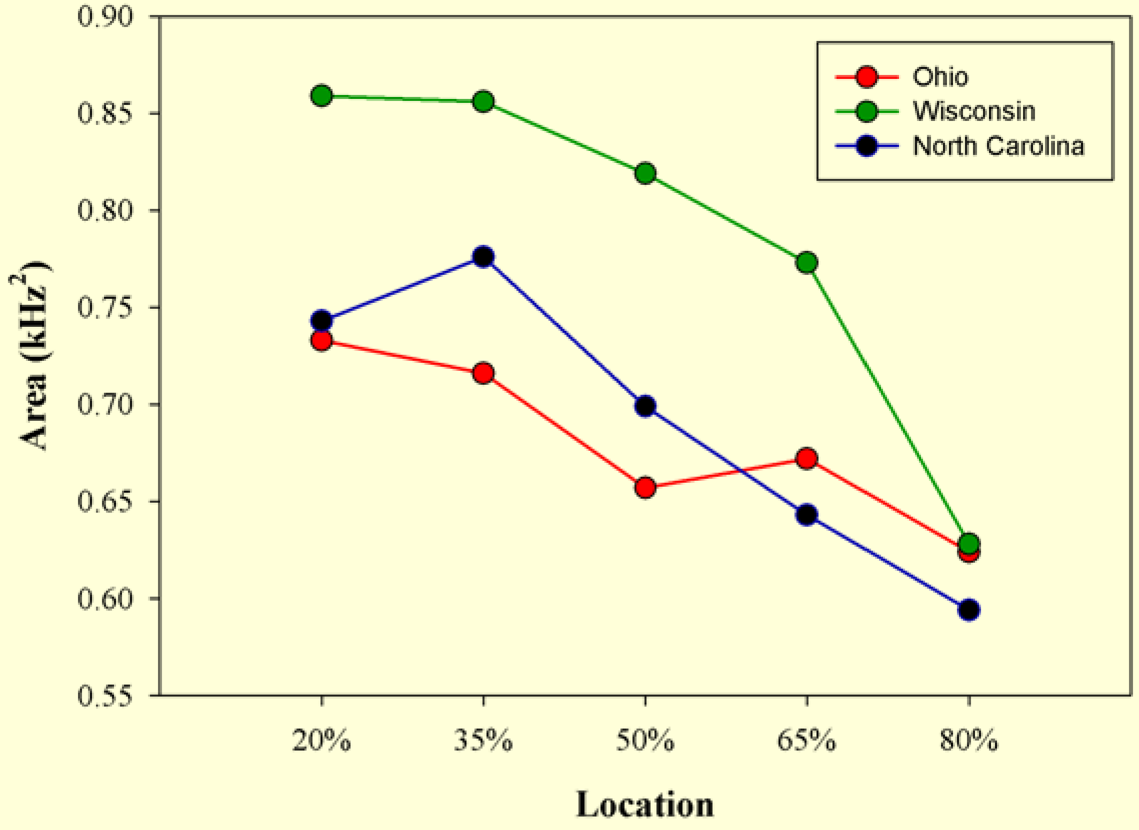
speech perception & acoustics labs

Differences in Mean Vowel Multilaterals for NC Speakers across Temporal Locations



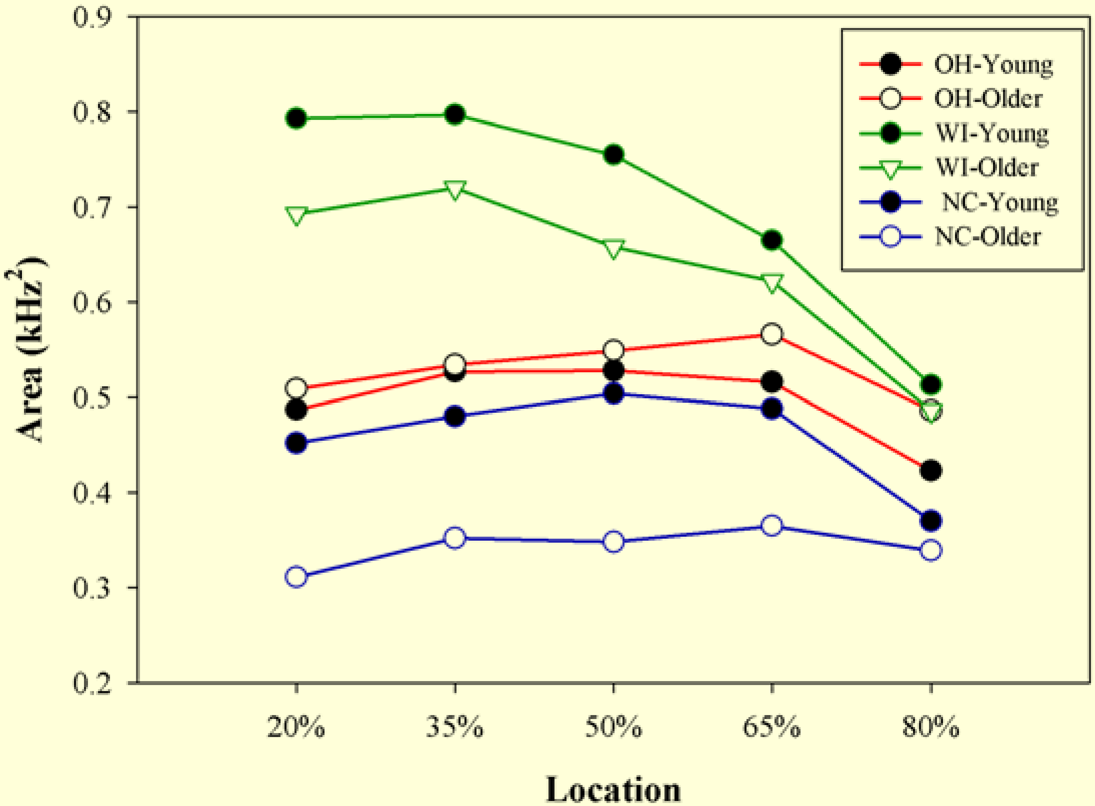
Temporal Location and Multilateral Area

There was a significant dialect x location interaction ($F(8,192)=7.1$, $p<.001$, $\eta^2=0.228$). 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.



Age and Multilateral Area

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

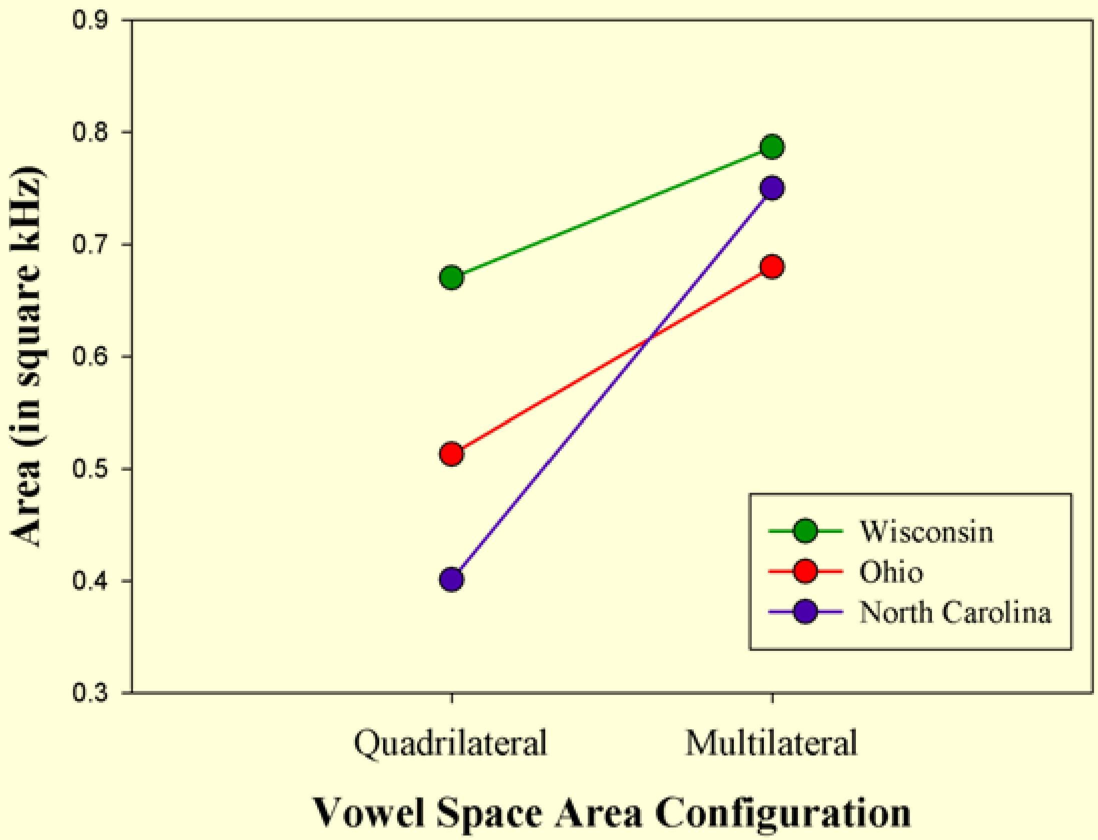


Quadrilateral vs. Multilateral Areas

- Finally, the quadrilateral and multilateral areas were compared using a mixed design ANOVA with the within-subject factor **temporal location** (20%, 35%, 50%, 65% and 80% points) and **vowel space configuration** (quadrilateral and multilateral) and the between-subject factors dialect (OH, WI and NC) and **age** (young adults and older adults).
- Not surprisingly, the overall multilateral areas (mean=0.719 kHz^2) were found to be significantly larger than the quadrilateral areas (mean=0.596 kHz^2) [$F(1,48)=1125$, $p<.001$, $\eta^2=0.959$].

Quadrilateral vs. Multilateral Areas: Dialect Differences

However, the most striking difference in areas of the quadrilateral and multilateral can be seen the in significant dialect by configuration interaction effect ($F(2,48)=81.4$, $p<.001$, $\eta^2=0.772$). The area change between the quadrilateral and the multilateral configurations is about the same for WI and OH, but dramatically larger for NC. **One cannot reliably predict the multilateral area from the quadrilateral**.

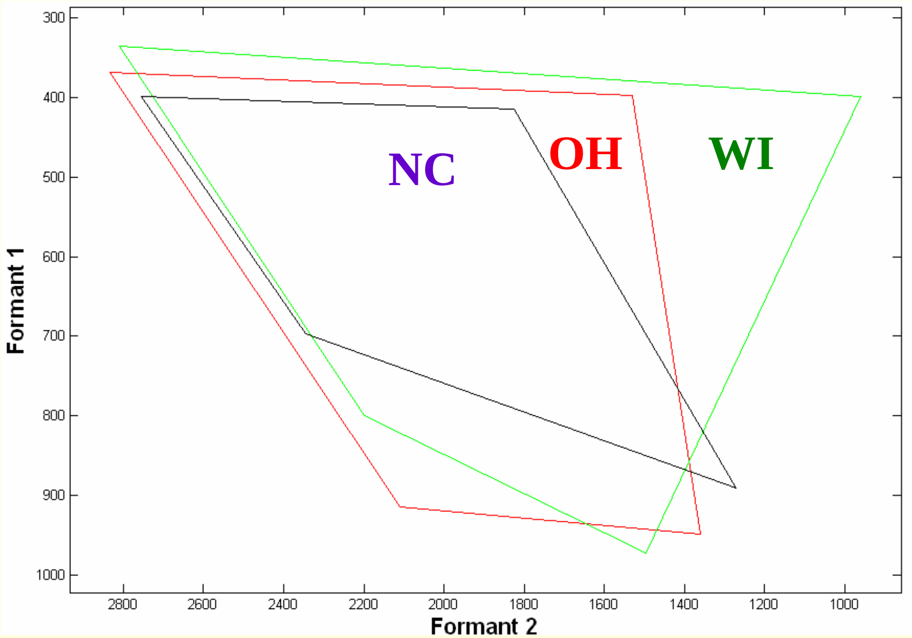


Areas of Normalized Vowel Spaces

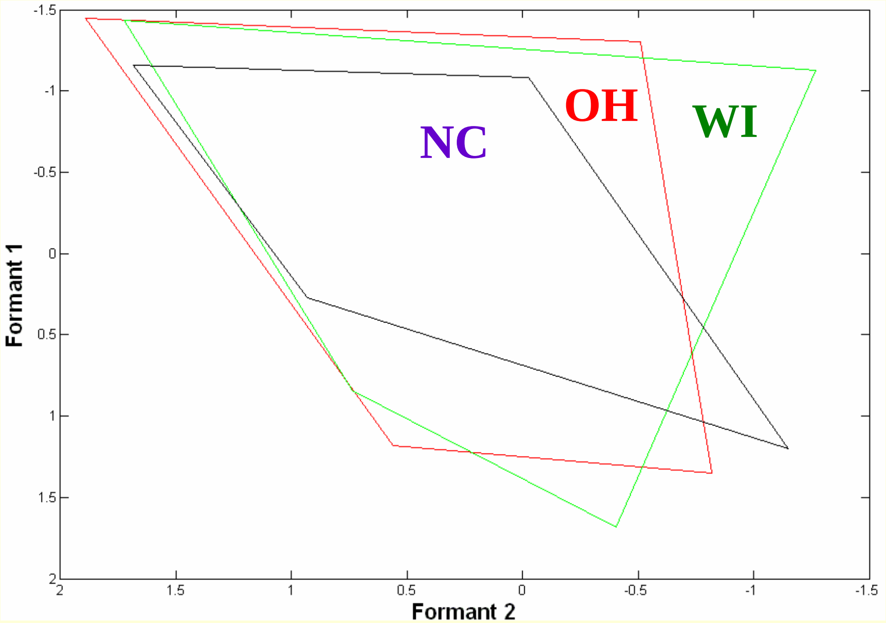
- When one examines acoustic vowels spaces, one is often forced to compare vowels produced by individuals with significantly different vocal tract lengths (e.g., comparing vowels produced by males and females or adults and children). To eliminate the individual differences produced by such variation, researchers often "normalize" the vowels using standard methods (e.g., as developed by Lobanov, Nearey and Watt & Fabricius). However, although these methods which change the relative positions of the vowels they may eliminate the dialect differences found here.
- Since the current study does not introduce significant vocal tract length variations (because all speakers are adult females) these data may allow a determination of whether such normalization scheme does indeed eliminate dialect differences in vowel space area. We used the Lobanov method to normalize each speaker's vowels and completed new ANOVAs.

SPA speech perception & acoustics labs

Mean Non-Normalized and Normalized Quadrilaterals at Vowel Midpoints for Dialects



Non-normalized

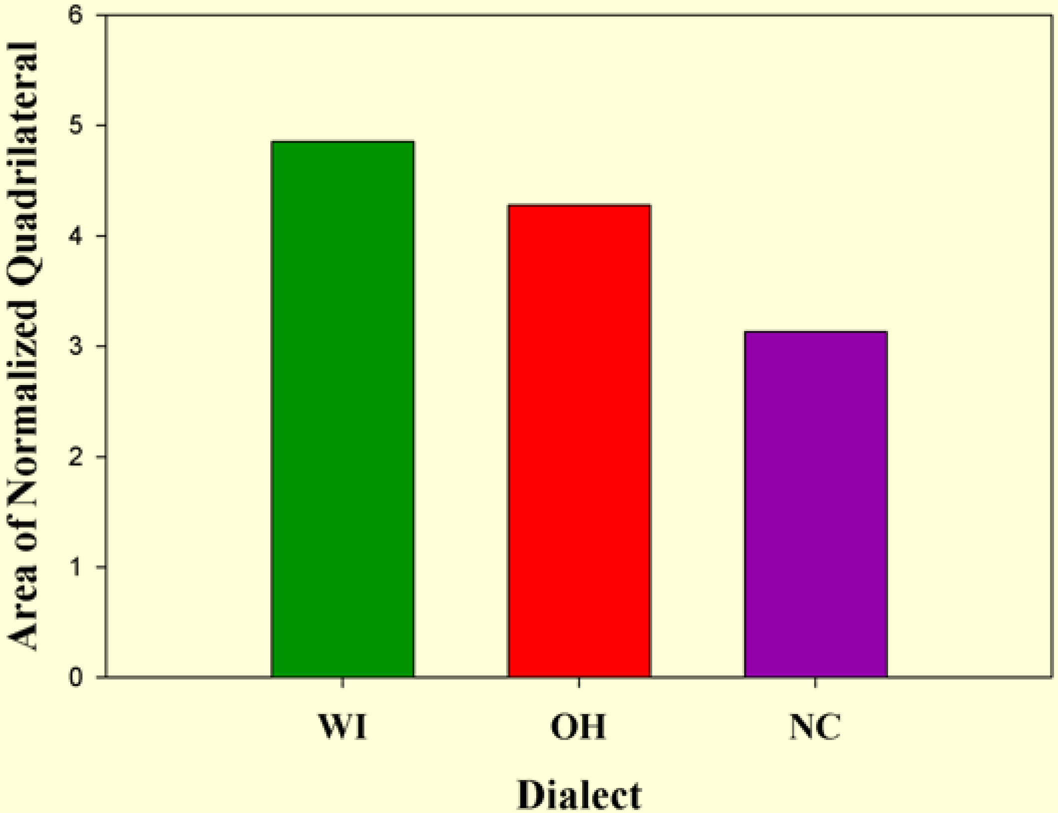


Normalized

speech perception & acoustics labs

Dialect and Normalized Quadrilateral Area

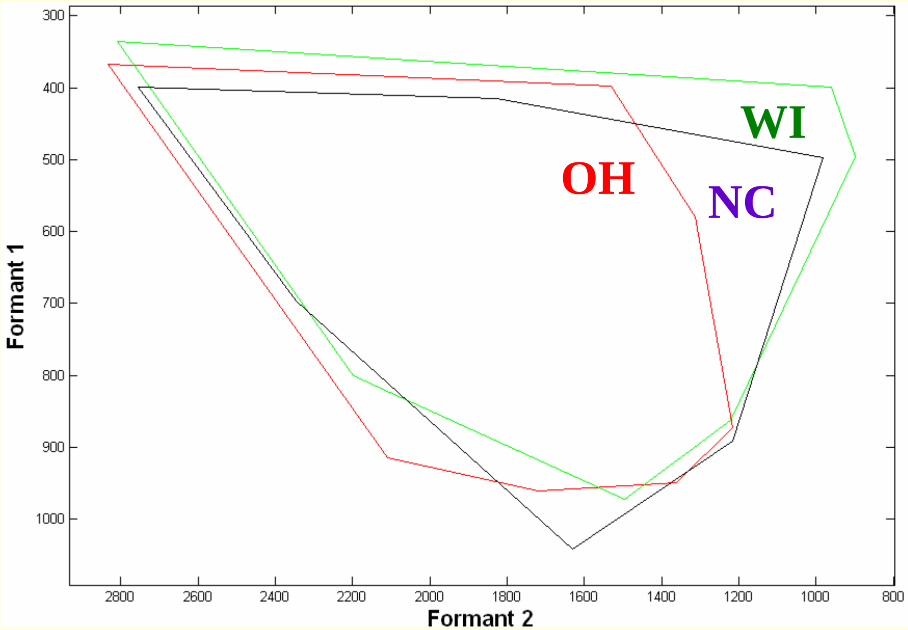
There was a significant difference in the size of the quadrilateral areas as a function of dialect ($F(2,51)=51.5$, $p<.001$, $\eta^2=0.669$). Wisconsin had the largest area (4.86) followed by Ohio (4.28) and North Carolina (3.13). Again, Tukey tests indicated each pair was significantly different.



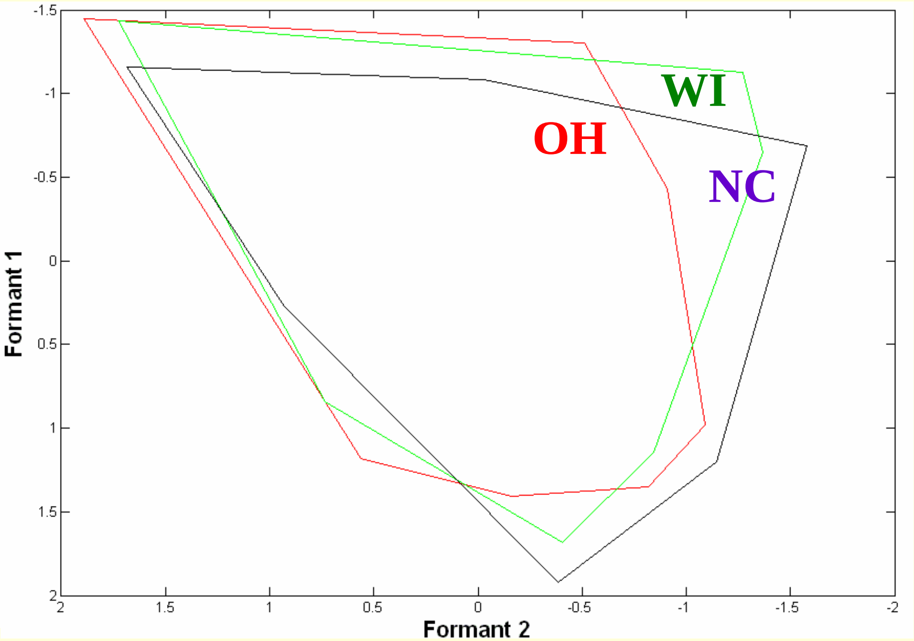
Normalization **did not eliminate** dialect differences in vowel quadrilateral area.

speech perception & acoustics labs

Mean Non-Normalized and Normalized Multilaterals at Vowel Midpoints for Dialects



Non-normalized

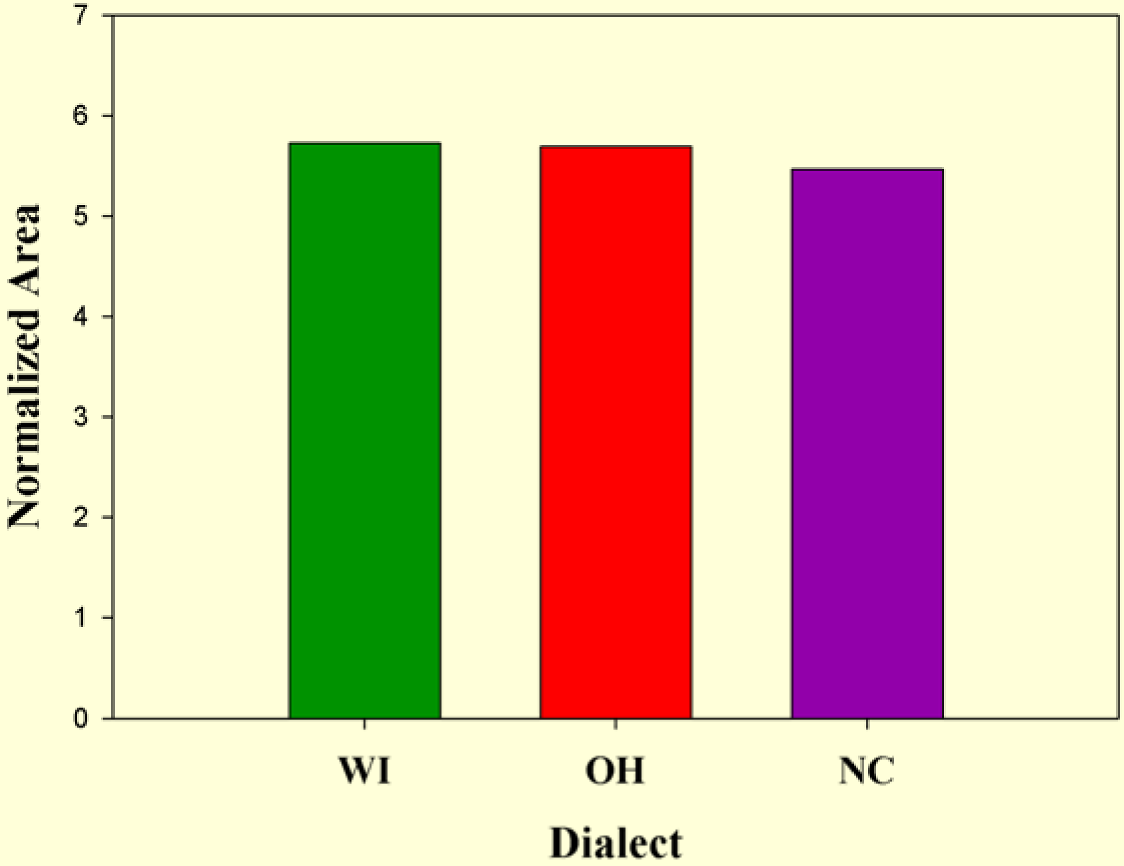


Normalized

SPA speech perception & acoustics labs

Dialect and Normalized Multilateral Area

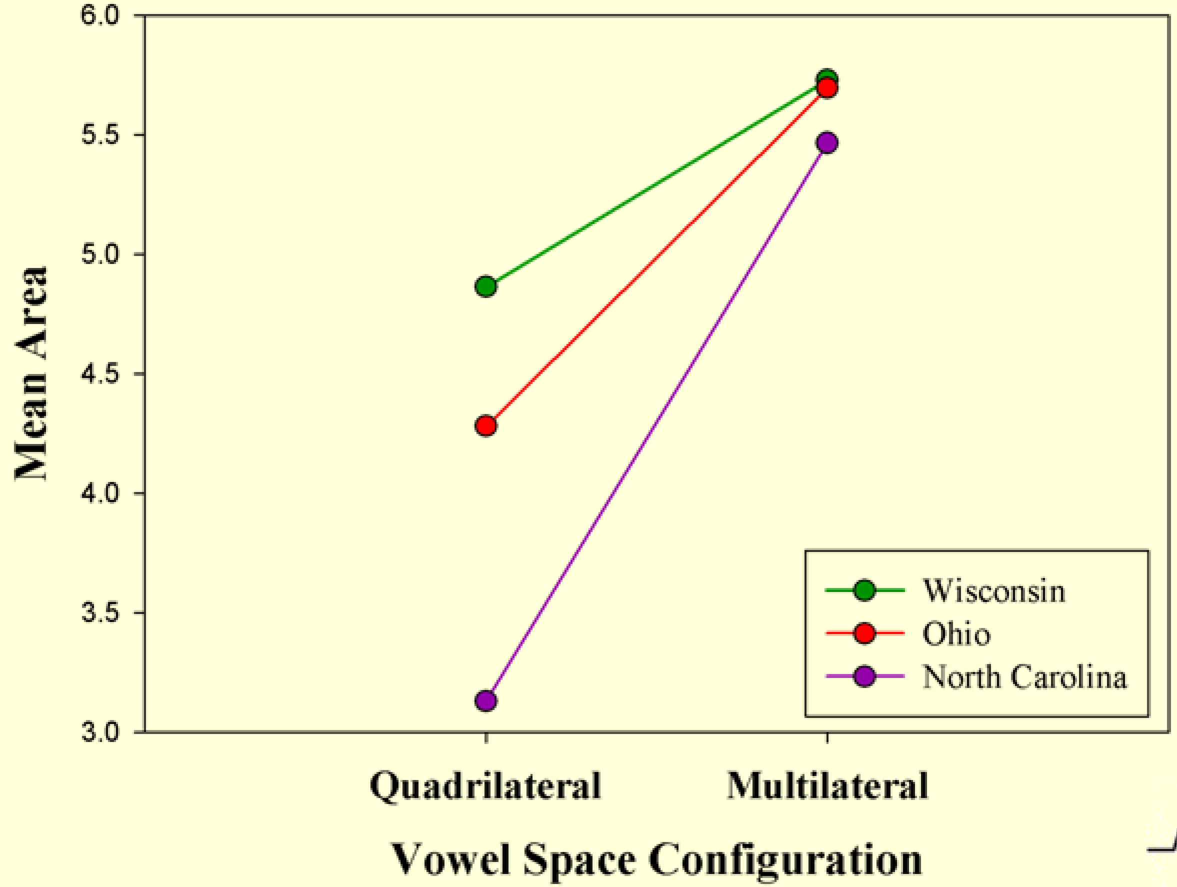
However, there was no significant difference in the size of the multilateral areas as a function of dialect ($F(2,51)=51.5$, $p<.001$, $\eta^2=0.669$). This result differs from that found with the non-normalized vowel space.



Normalization thus **did eliminate** dialect differences in the vowel multilateral area.

Normalized Quadrilateral vs. Multilateral Areas: Dialect Differences

This figure demonstrates the elimination of the vowel space area differences in the multilateral as a function of normalization.



SPA
speech perception & acoustics labs

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 and other temporal location differences were evident. This pattern of results suggests that the size of the vowel space corresponds to the dynamic changes in formant trajectories.

SPA speech perception & acoustics labs

Discussion (cont.)

- The results for the multilateral vowel space analyses indicate that some differences in the vowel space areas observed for the quadrilateral vowel space persist even when all perimeter vowels are considered (and it is not just the backed /u/ in WI that produces the differences in the size of the working vowel space area).
- The main effect of age for the quadrilateral was not significant, indicating that there was no difference in vowel space as a function of aging. However, the significant dialect by age interaction suggests dialect-specific diachronic changes in the vowel systems.

Discussion (cont.)

- Normalization did not eliminate the significant dialect differences in vowel quadrilateral areas. However, it did eliminate dialect differences in the vowel multilateral areas. This inconsistency suggests that formant normalization (at least the type of normalization introduced by Lobanov) should not be used or should be used with caution when comparing the size of the different speakers' working vowel spaces. When we do not expect significant differences among speakers in the lengths of their vocal tracts we should compare their working vowel spaces without normalization.

Discussion (cont.)

- Corner vowels (and the quadrilateral vowel space) may provide a good estimate of diachronic vowel change if such changes involve the fronting of /u/ or raising or lowering of /ae/ in American English. However, these processes may not affect the larger vowel space area involving the entire vowel system.