Arim Choi
Ling 451
Prof. Bradlow
3/17/2006

Context Effects on the Vowel Spaces of Native Korean Speakers

1. Introduction

Previous work has shown that different speaking styles, such as “clear” versus conversational speech, show phonetic differences such as expansion of vowel space (Bradlow 2002). The question that will be addressed in this paper is whether different recording contexts elicit different speaking styles, particularly for non-natives, as their ability to shift speaking styles in their L2 language may be different from their L1 speech strategies. Using vowel space as a measure of change in speech, we will observe if native Korean speakers manifest a difference of speech between a neutral and discourse context. If there is a shift in vowel space between different speech contexts, the next step will be to identify the motivation for the change.

2. Methods

2.1 Subjects and Materials

The materials for this study were taken from the Wildcat corpus, a database of native and nonnative speech in different contexts designed to look at speech adaptation and other interactions between speakers of different backgrounds. All speakers were recorded separately in a single-speaker setting and instructed to read lists of stimuli designed to elicit a comprehensive collection of American English sounds. The stimuli were also designed to account for cross-context comparisons. The second part of the database consisted of recordings in which subjects performed a map task. One speaker directed the other speaker in to placing items on a map, or building a specific structure using the designated materials. Both speakers were given a chance to be the “giver” of the instructions on both tasks. The particular recordings chosen from the Wildcat Corpus were the Part A recordings by two native Korean speakers (K1 and K2), as well as their Part B recordings. Both of the Korean speakers had native Chinese speakers as their partners in part B (K1 with C1, K2 with C2).

2.2 Measurements

In order to make a comparison between the neutral recording context and the map task context, stimuli representing 3 cardinal vowels had to be chosen. Despite the design of the database to allow for cross-context comparisons, the inventory of comparable stimuli was rather sparse. The three vowels measured were /i/, /u/, and / /. The three tokens used for the vowel /i/ in both contexts were *heed*, *feed*, and *bean*. The three tokens for /u/ were *who'd/hude, juice*, and *tube*. Finding appropriate stimuli for / / was difficult as it was not the primary target vowel in the creation of the cross-context stimuli (/ / appeared to be the best-represented vowel for part B). The words chosen for part B, the discourse context, were *block*, *star*, and *top*. Their correlates in the Part A stimuli were *block*, *start*, *bot* for K1, and *dock* for K2. Unfortunately there was no *star* or *top* in the Part A stimuli, but *start* was similar, minus the final plosive and the duration of the vowel. The third stimulus for / / was different for K1 and K2, because *bot* was missing in the Part A stimulus for K2 (the word *boat* was repeated twice, instead). *Bot* and *dock* were chosen as the closest possible stimuli to *top* in Part A (the other words with / / either began or ended with a nasal, which would mean a different context for the vowel). Measurements were made using a Praat Script which extracted the F1, F2, and duration of the vowel. Segmenting *heed, keep, who'd/hude, juice, tube,* and *top* were straightforward, taking the vowel from a point where it began to regularize itself after the stopburst or release of the fricative. The vowels in b*ean*, *block* and *start* were more difficult to segment, due to the sonorant consonants either preceding or following the target vowel. The end of the /i/ in *bean* was determined by looking for the distinct shift in the waveform amplitude and shape which coincided with a corresponding shape in the spectrogram. A similar strategy was taken for measuring the / / in *block*. *Star* and *start* were by far the most difficult to segment due to the rhoticization of the / /. The vowel was only measured up to the point where the third formant had reached the nadir of its descent.

The vowel measurements for the neutral recording context were only made once, seeing as how the stimuli occur only once. For all if the discourse stimuli except *top,* measurements were taken for each occurrence during when the Korean speaker the "giver" for the map task. *Top* occurred consistently in the structure building task, so the measurements for each occurrence of *top* was made from this task.

3. Results

The complete neutral context vowel space for K1 and K2 are show in Figures 1 and 2, respectively. Given the current data, the vowel that appears to be the most inconsistent is the /u/. This variability is paralleled in the distribution of the /u/ in the discourse context for K1 and K2 (see Figures 3 and 4). Otherwise, there was not really a concrete pattern in the F1 and F2 of both the neutral context and discourse context vowels. An average formant values of the discourse tokens compared to the neutral context tokens is given in Figures 5 and 6 (for K1 and K2, respectively). There is less variability within the vowels for the discourse contexts in the data for both speakers, but this is probably a result from there being more samples of the tokens to average values from in that context. Conversely, there is variability of the /u/ and /   / vowels in the neutral context tokens, but this pattern does not follow for the /i/ vowels. They remain relatively closely clustered to each other and with the /i/ discourse tokens. One thing to note is that the variability of the /u/ vowel occurs mostly in the F2 dimension. Possibilities of why this occurred will be addressed in the discussion.

Extrapolating a vowel space for each context using the average formant values for each vowel shows us that given the current data, we cannot concretely conclude that there is a significant effect of context on a nonnative vowel space (See Figures 7 and 8). There is no evidence of a distinct shift or an expansion of the vowel space, particularly for speaker K2. There is a large difference in the F2 of /u/ between the neutral and discourse contexts, but based on the limited amount of data, it is difficult to say that it is a significant difference in frequency.

The other dimension measurement taken from the vowels was their duration. Many of the target words in the discourse context were stressed, because they were key content words and highly salient for the speaker. The words in the discourse context were separated into stressed and unstressed, although in some cases when there was only one unstressed word of that token type, its result was generalized with the other tokens (See Figures 9 and 10). For the most part, unstressed vowels in the discourse context were shorter than the stressed ones as expected. There was not a clear pattern in the duration of vowels between the neutral and discourse contexts. Perhaps the only notable difference for both speakers K1 and K2 was the duration of the /u/ in a neutral context. This value is the only duration clearly greater than that of its respective discourse vowel.

4. Discussion

In general, it seems as though there is not sufficient enough data, particularly in the neutral context, to observe a concrete pattern in the data. One could make the argument that perhaps we see no real difference between the vowel spaces of the different contexts because speakers rely on the same style of speech when reciting a list of words in a single-speaker context and when using specific content words in a goal-oriented task where they must direct other people. One trend that was observed during the Part B recordings of both these Korean speakers is that they tended to pause and say the token word in an effort to emphasize it. Perhaps this strategy of making these words clear is similar to the way in which they read the words on the lists.

Another pattern in the speech of the Korean speakers during the map task was that they would say a word, and repeat it a few times afterward, still putting a stress on the word but trailing off, (e.g. "put the block at…Heed and Had street, Heed and Had street, Heed and Had street..okay?"). This was a pattern particularly characteristic of speaker K1. One thing to look at in the future will be whether this is a pattern particular to these two Korean speakers or if this repetition and stress strategy is something that both native and nonnative speakers of English employ during this map task. If this is just something exhibited by nonnative speakers, it could be a strategy to give themselves more chances to be understood if they are not as sure about their pronunciation being heard once. If both native and nonnative speakers, do it, then maybe it is because it is helpful for the listener to hear more than once, especially if the native speaker is confident in what they utter as being clearly understood the first time around.

I did not show in the figures the difference between stressed and nonstressed words in discourse context because looking at the data, it did not look like the stressed vowels were being particularly hyperarticulated. In some cases, it was the unstressed word that had higher formant values overall.

One thing this study clearly points out is the limited capabilities of the database. There were not enough cross-context words to provide solid evidence for what the shape of the vowel spaces should look like in each context. For example, there were a limited number of tokens with the /   / vowel. *Star* and *start* have different vowel durations and in general, both of the durations for these words were longer than for the other tokens in the /   / vowel category. Aside from segmenting difficulties, the nonuniformity of the stimuli may be a source of confounds. Another such case of questionable stimuli was with the /u/ vowel tokens. The variability in F2 results might have something to do with the fact that the words *tube* and *juice* are frequently palatalized. This was the particularly the case with the Korean speakers. Sometimes it almost sounded as though the vowel was being diphthongized. Thus the stimuli used in the database might have skewed the resulting measurements for /u/. If in fact the results for /u/ are legitimate, there are a number of factors that could be affecting this. The change could be a result from adaptation to their discourse partner. One way to verify this would be to construct a vowel space of their discourse partner and see if the change in the /u/ value happens to occur in the direction of the Chinese speaker’s /u/ (they are converging upon a mutual vowel). Another possibility is that the speaker is using a vowel more characteristic to their native language because their focus is more on completing the map task successfully and subconsciously forgoing pronunciation of the vowels in an American English manner. To verify if there is a resemblance to the L1 vowel space, additional data in the native language that is comparable to the English data must be obtained.

Failure to find contexts effects in this study does not rule out any future possibilities. As mentioned before, it might be that the map task context and the single-speaker context are not that different, which is why we don't observe any definite shift. It would be interesting to see if this lack of shift is found in both native and nonnative speakers. Also, the map task is only one context, and maybe more easygoing conversational condition might show different results. Another factor to consider is that both Korean speakers in this study were speaking to another non-native speaker. Maybe because they have more of a similar underlying sound system than English, there was no motivation to shift vowels around to accommodate the other speaker. Repeating this task with a native and nonnative may show very different results.

References

Bradlow, A. R. (2002). Confluent talker- and listener-related forces in clear speech production. In Gussenhoven, C. & Warner, N. (Eds.) Laboratory Phonology 7. Berlin & New York: Mouton de Gruyter. Pp. 241-273.

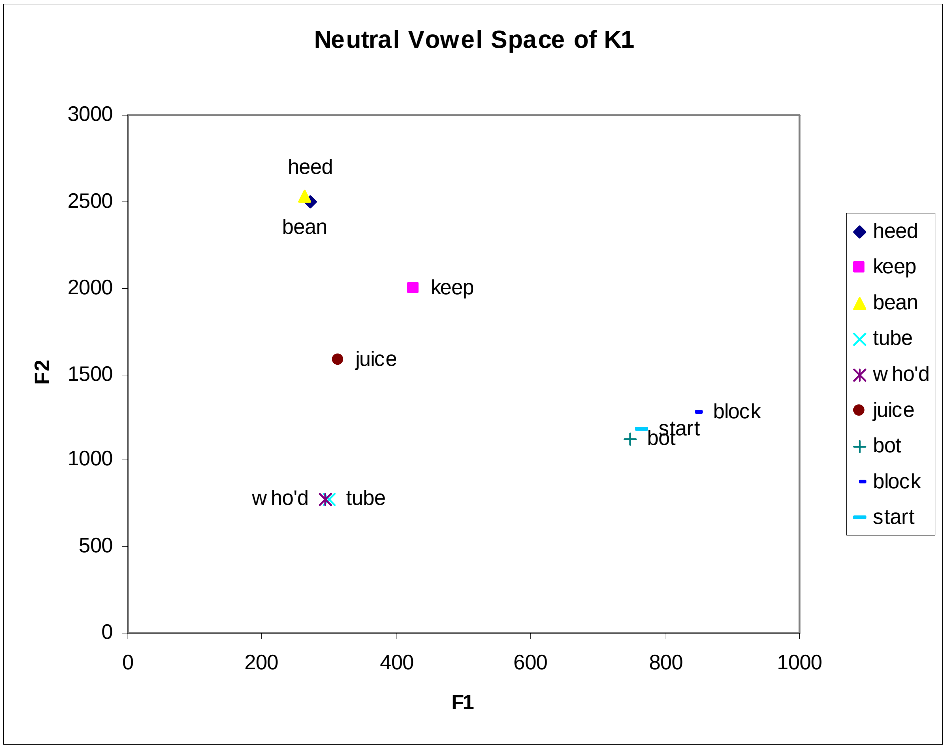

Figure 1. The neutral vowel space of K1. This figure shows the distribution of all 9 tokens taken from the single-speaker recording context.

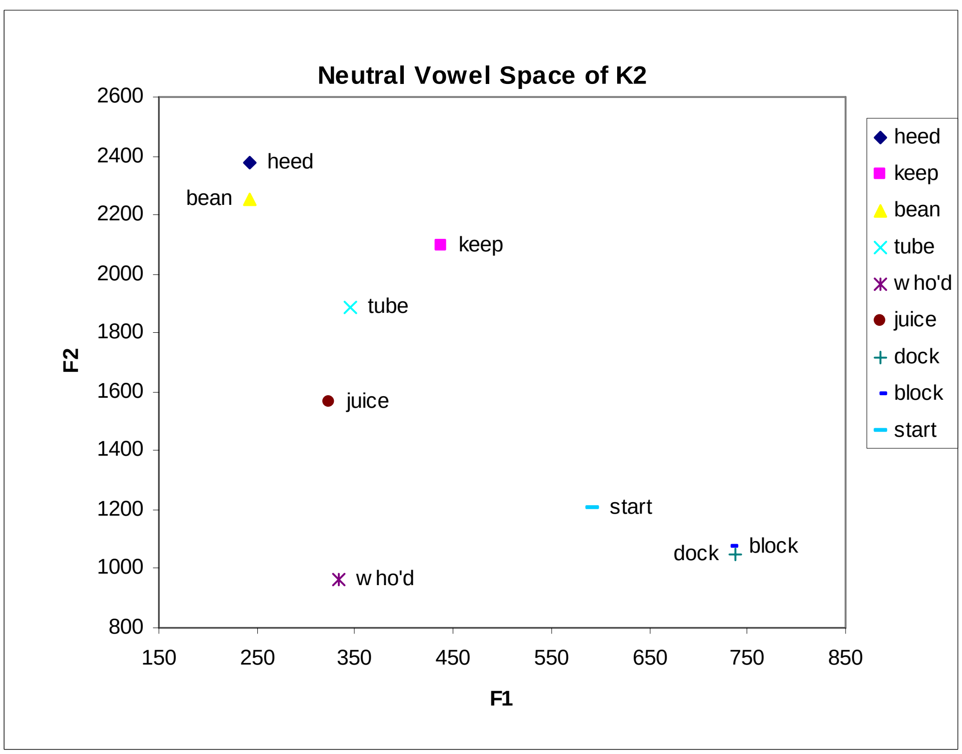

Figure 2. The neutral vowel space of K2. This figure shows the distribution of all 9 tokens taken from the single-speaker recording context.

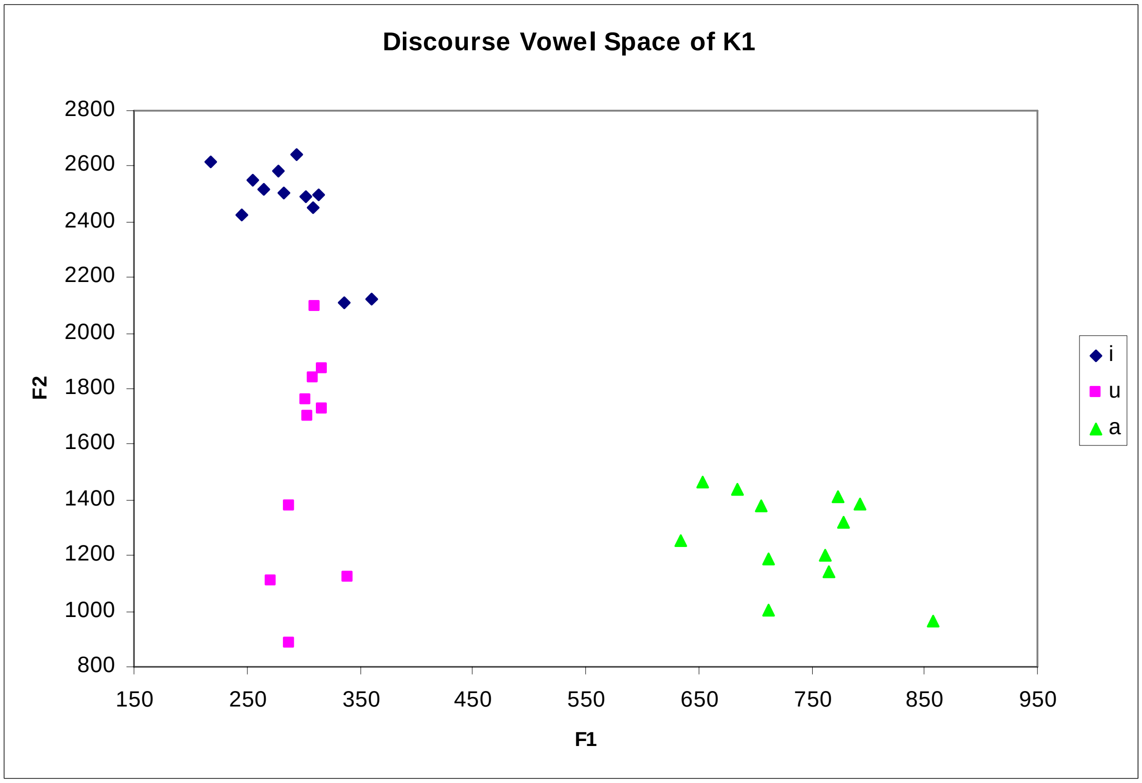

Figure 3. The discourse vowel space for speaker K1. This vowel space represents all the measurements of each of the tokens taken from the discourse context for K1.

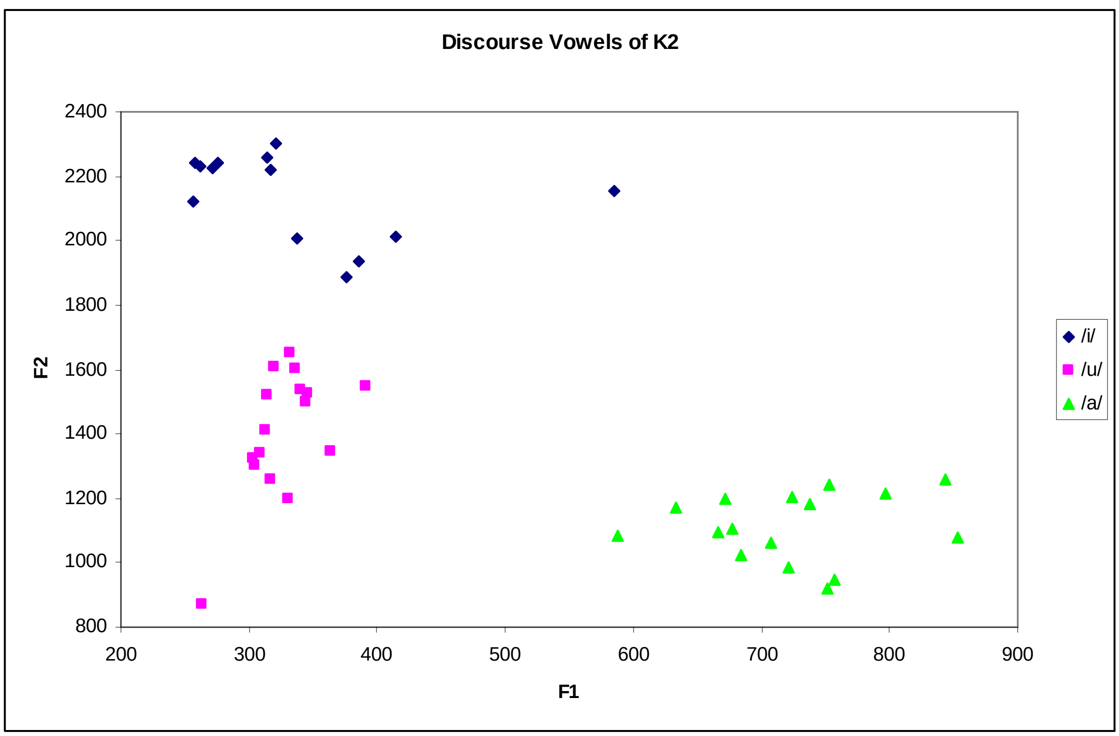

Figure 4. The discourse vowel space for speaker K2. This vowel space represents all the measurements of each of the tokens taken from the discourse context for K2.

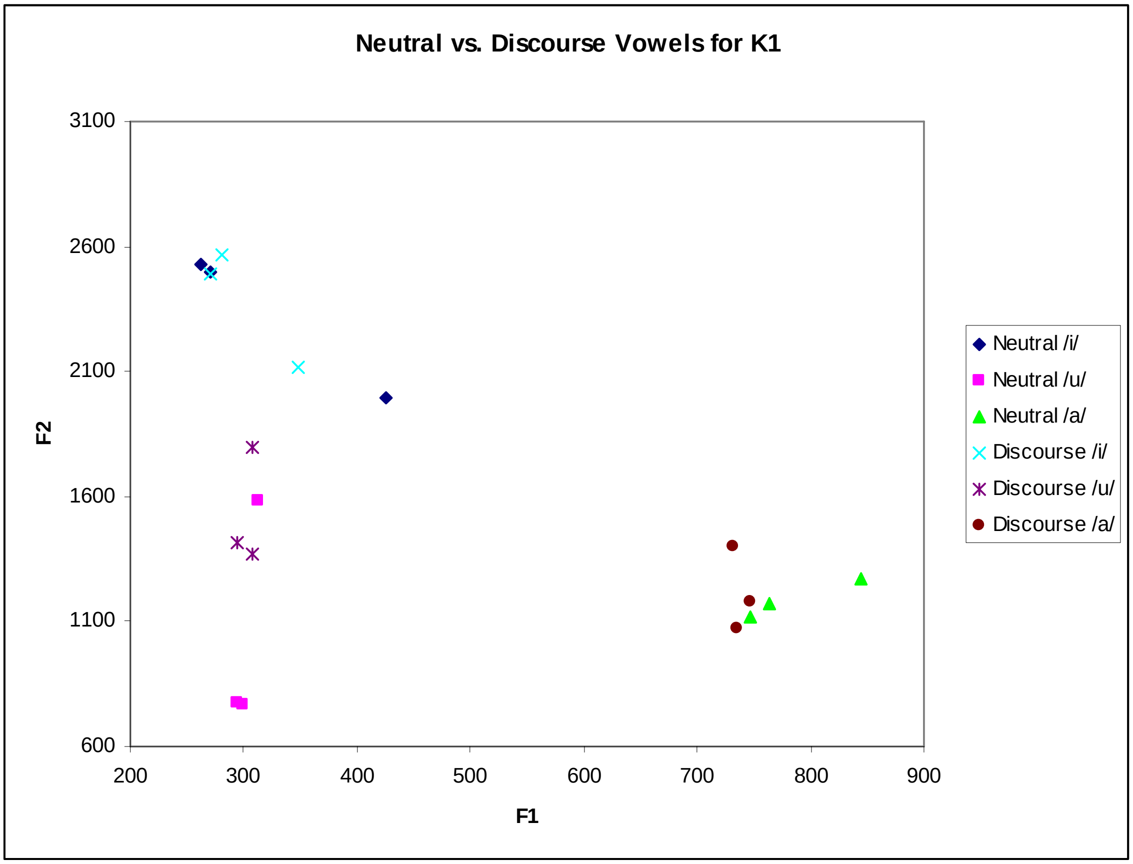

Figure 5. Neutral vs. Discourse vowel space for K1. This is a plot comparing the vowel spaces created by the neutral and discourse tokens.

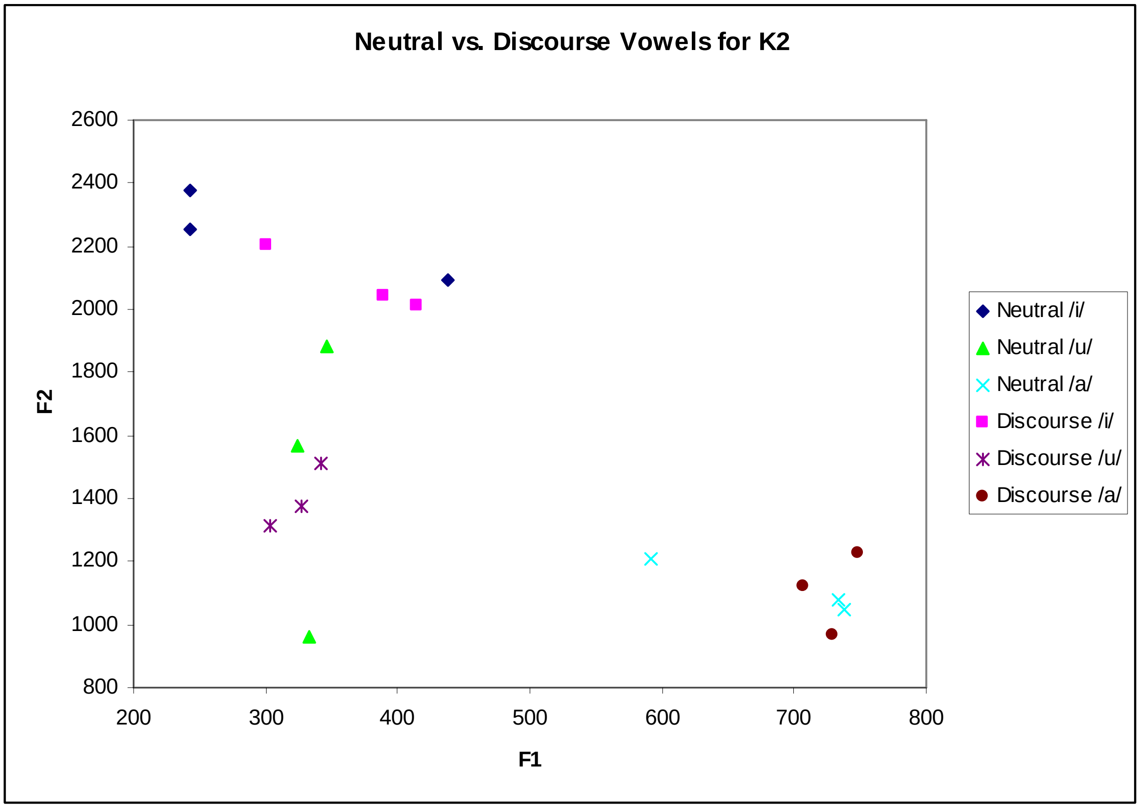

Figure 6. Neutral vs. Discourse vowel space for K2. This is a plot comparing the vowel spaces created by the neutral and discourse tokens.

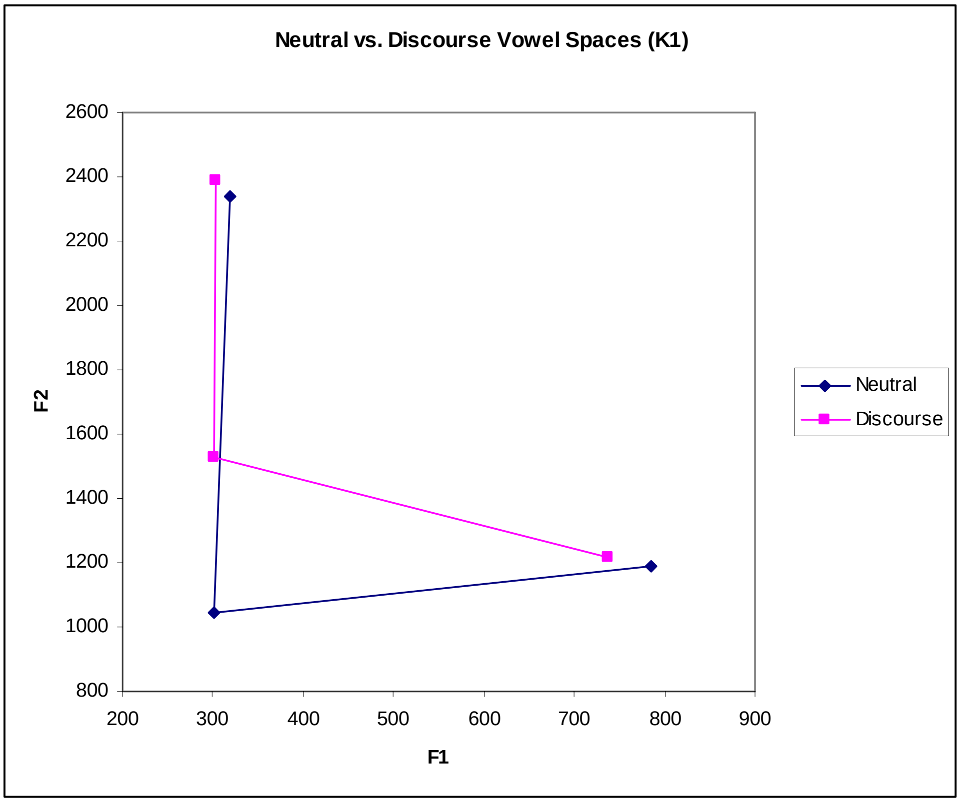

Figure 7. Neutral vs. Discourse vowel spaces for speaker K1. This figure is created by taking the average formant values for the vowels in each context.

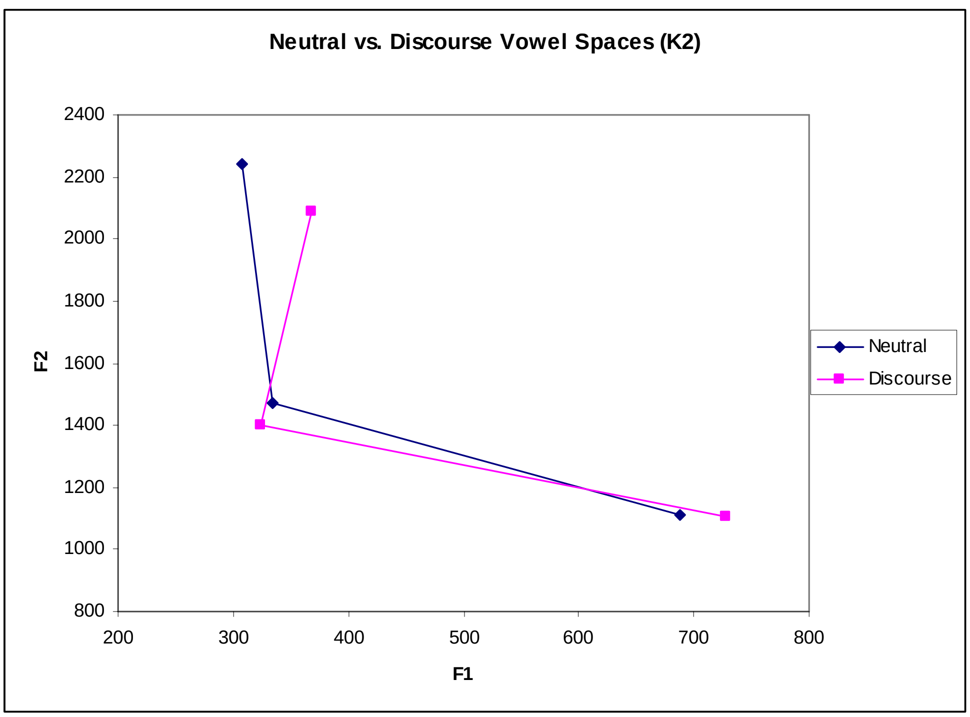

Figure 8. Neutral vs. Discourse vowel spaces for speaker K2. This figure is created by taking the average formant values for the vowels in each context.

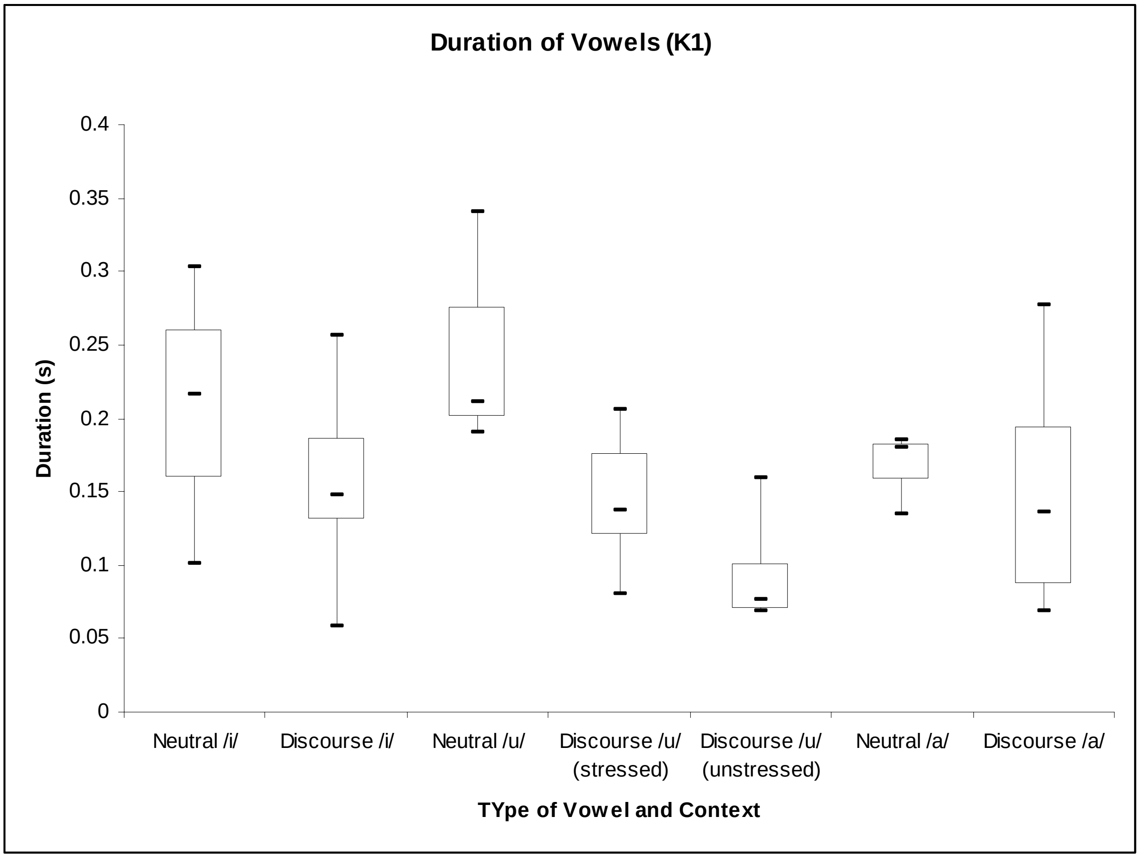

Figure 9. Duration of vowels across context for K1. For the discourse tokens that are not separated into stressed and unstressed categories, there was only one unstressed word. The end bars for each category are the maximum and minimum durations, the block represents everything between the 1st and 3rd quartile, and the bar within the block represents the median of the data.

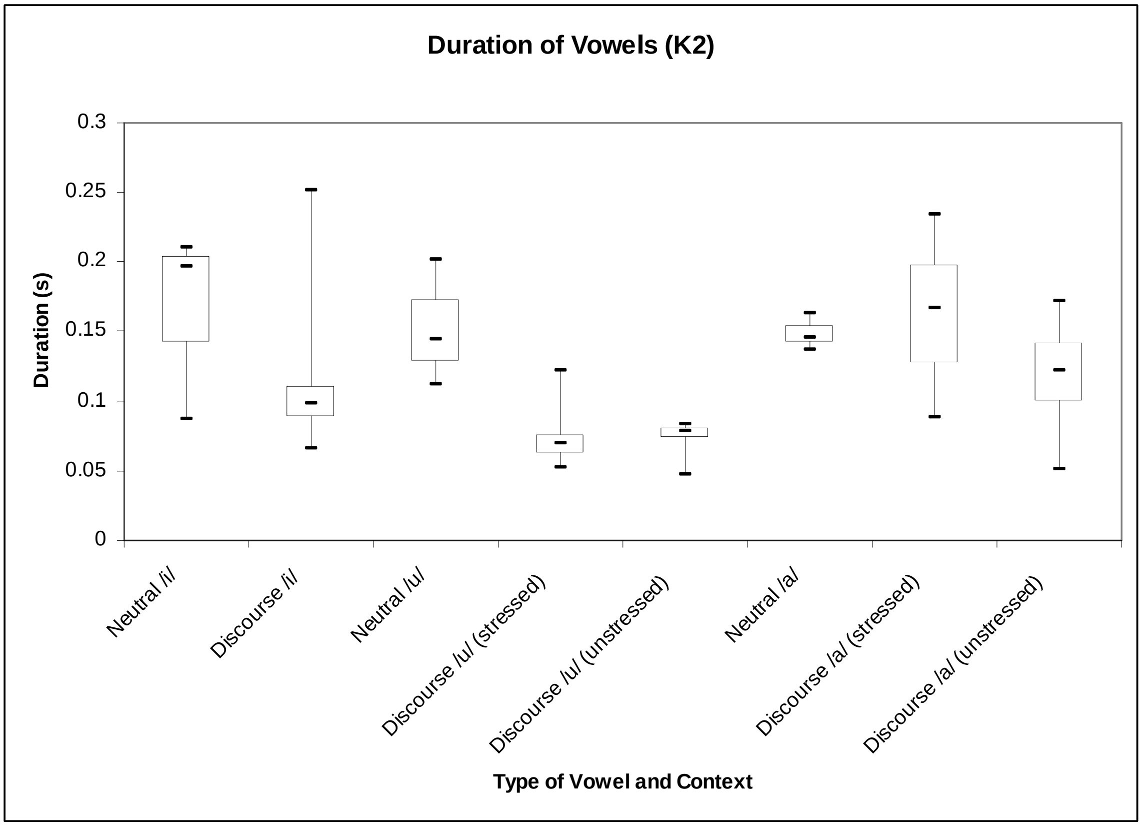

Figure 10. Duration of vowels across context for K2. For the discourse tokens that are not separated into stressed and unstressed categories, there was only one unstressed word. The end bars for each category are the maximum and minimum durations, the block represents everything between the 1st and 3rd quartile, and the bar within the block represents the median of the data.