Prosodically-Induced Phonetic Variation in Vowels: A Source of Language Change?

Robert Allen Fox
Ewa Jacewicz
Joseph Salmons

1. Introduction

Diachronic principles of vocalic chain shifting have been formulated on consistent patterns found in the transmission of sound across many generations of speakers and across many dialects and languages, both in the historical record and in sound changes in progress. But how and why do such changes span generations? One explanation, for which a foundation is established in this article, is that prosodically conditioned realizations cause some vowels to be perceived as more salient than others in the continuously varying speech stream. These more prominent productions may serve, we will argue, as a significant impetus for vowel shifts over time in that more forceful realizations of vowels are transmitted across generations, being reinterpreted by new generations as prosodically less marked ones.

Consistent patterns of chain shifting have been posited since at least the German phonetician Sievers (1876/1881) and the core observation has been often repeated down to the present, as by Labov (1994: 116, drawing on his "preliminary formulation"):

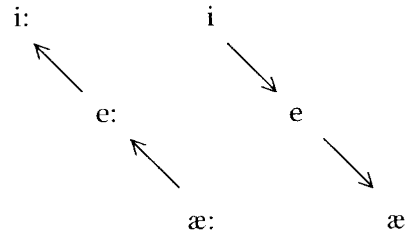
- Principle I: In chain shifts, long vowels rise.
- Principle II: In chain shifts, short vowels fall.

The characterizations *long* and *short* have regularly been exchanged for *tense* and *lax* or *peripheral* and *non-peripheral*, and these two princi-

Howard I. Aronson, Donald L. Dyer, Victor A. Friedman, Daniela S. Hristova and Jerrold M. Sadock, eds. *The Bill Question.* Bloomington, IN: Slavica, 2006, 87–110.

ples have often been represented graphically as in (1), below, using here three front vowels.

(1) Classic pattern of Chain Shifting (Sievers 1876/1881, Labov 1994, many others)



Such patterns provide a classic example of regular or Neogrammarian sound change (Labov 1994: 502–43), and they are widespread, even chronic, in the Germanic languages, along with some other families, to the point that it has been argued that instead of talking about particular chain shifts as happening at particular times, they should be regarded as an omnipresent, dynamic characteristic of vowels in a language like English (Stockwell 1978). We propose that these long-accepted principles result, at least in part, from variation in the prosodic context in which the vowels occur, and that perception plays a critical role in this process.

Specifically, we present acoustic data on the nature and size of variation in vowels resulting from differences in prosodic context and perceptual data which illustrate how these variations affect vowel quality. Our main focus is how listeners respond to the patterns of variation found in the production study. Although the transmission of sound over generations involves both speakers and listeners, perception of dialect-specific vowel characteristics and cross-dialectal differences in vowel features have been seldom studied by sociolinguists (*cf.* Thomas 2002). This study provides evidence that listeners use information in the acoustic vowel characteristics to identify vowels differently depending on their prosodic context. These cues are strong enough to affect vowel perception regardless of speaker dialect and speaker gender, which also make a significant contribution to how the vowels are perceptually classified.

2. Acoustic Characteristics of Vowels /e, ɛ/ Spoken in Central Ohio and Southern Wisconsin

In related work (Jacewicz *et al.*, forthcoming), we conducted an acoustic study to test the hypothesis that differences in prosodic prominence are one source of synchronic variation in vowels which may lead to vowel chain shifts over time. Variation is prosodically structured, so that more and less emphatically pronounced vowels vary in predictable ways as a function of hierarchical prosodic constituency. We present here only the selected findings of that study that bear directly on the focus of this article while a detailed description of these acoustic results can be found in Jacewicz *et al.* We examined the changes to the acoustic vowel characteristics that result when a word is produced in three different prosodically defined positions in an utterance. Eight young adults who were native speakers of American English spoken around Columbus, OH (four men and four women) and eight more speakers born and raised in the area of Madison, WI (four men and four women) served as speakers. We recorded each speaker producing sentences as in (2) containing the words *bait* or *bet* in three maximally distinct prosodic positions: the highest Utterance (U), intermediate Phonological Phrase (P), and the lowest Syllable (S):

(2) U-position: stressed and accented vowel

'*Bait shop*' is what I said.
'*Bet some*' is what I said.

P-position: stressed vowel

She said *the bait shop* was closed.
He said *the bet slips* were here.

S-position: full vowel

Shark bait with flavor seems hard to find.
Risky bets are nice but safe bets are better.

Acoustic measurements of vowels /e, ɛ/ contained in *bait* and *bet* included vowel duration and the frequency of the first two formants, F1 and F2, at several different points in its production (so that formant change over time—diphthongization—could be examined). We predicted that vowels in U-position would be longer and that their formant values would be more extreme as compared to P- and S-positions, respectively. This is because vowels in the strongest prosodic position are produced with greater emphasis (and, thus, with greater articulatory effort) than vowels in intermediate and weakest positions, respectively. We were interested in the extent to which these acoustic changes occurred independently of speaker dialect and gender. At the same time, we also anticipated dialect differences whose precise nature and size were the subject of our investigation.

Mean vowel duration data for /e, ɛ/ for both dialects are shown in Figure 1. Not surprisingly, the tense vowel /e/ is longer than the lax /ɛ/. Of particular interest is that duration of either vowel is sensitive to variation in prosodic prominence level. Generally, vowels in the U-position are longer than in either P- or S-positions, in accord with our predictions based on hierarchy of prosodic prominence. Unexpectedly, Ohio speakers were found to produce somewhat longer vowels than the Wisconsin speakers.

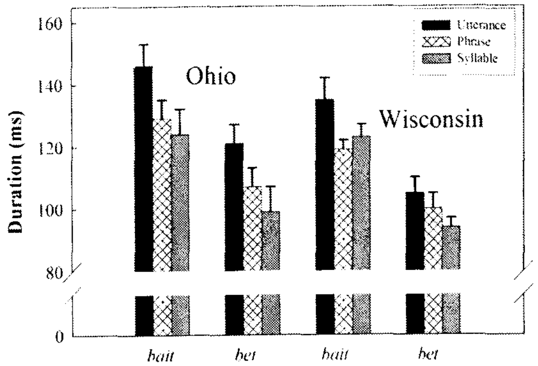


Figure 1. The effect of prosodic prominence on the duration of the vowels /e/ in *bait* and /ɛ/ in *bet*

For each group of speakers (Ohio and Wisconsin), the bar graphs represent mean values in three positions of prosodic prominence: Utterance, Phonological Phrase and Syllable. The values are collapsed across speaker gender. Error bars indicate one standard error.

The movements of the first two formants, F1 and F2, were examined by measuring their frequencies at five different locations corresponding to 20%, 35%, 50%, 65% and 80%-point in the course of each vowel's duration. Shown in Figure 2 are mean F1 and F2 frequencies of /e, ɛ/ plotted at each measurement location across the three prosodic positions. For both dialects, the nature of formant change over time was affected by prosodic position. For /e/, more prominent pronunciations related to higher prosodic prominence produced more extreme formant frequency values which corresponded to greater frequency change and greater fronting and raising of the vowel in the acoustic space. Although the variation for /ɛ/ was smaller, the effect of prosodic context was also manifested in that greater lowering and fronting of the vowel in the acoustic space was associated with increased prosodic prominence.

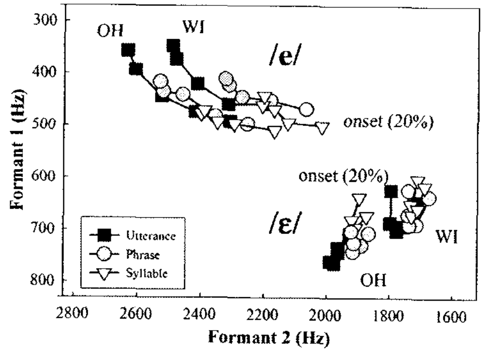


Figure 2: Mean F1 and F2 frequencies for /e/ and /ɛ/ at five locations in the vowel, corresponding to 20%, 35%, 50%, 65% and 80% of vowel duration

The frequency values are plotted for each prosodic position. The 20% measurement location is marked as "onset" and the lines connect each consecutive measurement locations for each prosodic context. For each speaker group, *i.e.*, Ohio (OH) and Wisconsin (WI), the mean frequency values are collapsed across speaker gender.

For the tense (or long) vowel /e/, our synchronic results provide consistent and close parallels to the behavior reported for long/tense vowels in diachronic chain shifts. First, the vowel rises under greater prosodic prominence, in step with the diachronic Principle I, which postulates that long vowels rise in historical sound change. Second, it also moves closer to the edge of the acoustic vowel space, or becomes more peripheral (Labov 1994), in a way reminiscent of the expansion of the vowel space found in work on clear speech and infant-directed speech. For lax (or short) /ɛ/, we have found consistent pattern with diachronic Principle II, that short vowels lower. This general pattern of prosodically induced variation was found in both Ohio and Wisconsin productions, despite dialect-related differences in their positions in the acoustic space. As can be observed in Figure 2, both vowels /e, ɛ/ spoken in Ohio are more fronted relative to Wisconsin vowels.

The pattern of variation across prosodic positions is also evident when the mean values are displayed separately for men and women. As shown in Figure 3 for /e/, formant values in women's productions are more extreme than in men's due to their shorter vocal tracts. Despite the difference, there is a consistent tendency of raising and

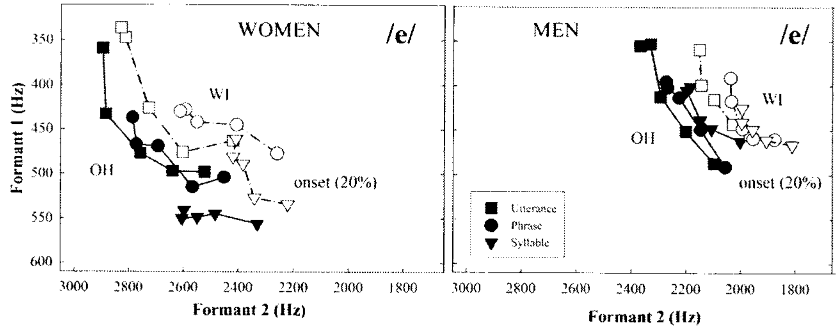


Figure 3: Mean F1 and F2 frequencies for /e/ in *bait* at five locations in the vowel, corresponding to 20%, 35%, 50%, 65% and 80% of vowel duration plotted separately for men and women

The frequency values across prosodic positions (Utterance, Phrase and Syllable) are represented by solid symbols for Ohio and open symbols for Wisconsin speakers.

fronting the vowel /e/ with each higher prosodic position for each gender group and for each dialect. Gender-related differences can be observed in the S-position relative to U- and P-contexts, however. Men's productions in the S-position show more overlap with the other prosodic contexts whereas women's vowels are clearly separated. This pattern cannot be explained on the basis of the differences in the vocal tract alone.

Figure 4 shows a similar display for the vowel /ɛ/. Although the internal pattern of formant movement is different for /ɛ/ than for /e/, one can observe an effect of prosodic context in the general trend of lowering and fronting the vowel with increased prosodic prominence regardless of speaker gender and dialect. As for the vowel /e/, gender-related differences can also be seen in the S-position, particularly in Ohio productions, where women's vowels in S-positions are separated from other prosodic contexts whereas men's vowels tend to overlap. As evident in Figures 3 and 4, the Wisconsin vowels are shifted toward the back relative to the Ohio vowels, occupying a more centralized position in the acoustic space.

In order to further examine the magnitude of formant movement, we utilized the measure "vector length" (Ferguson and Kewley-Port

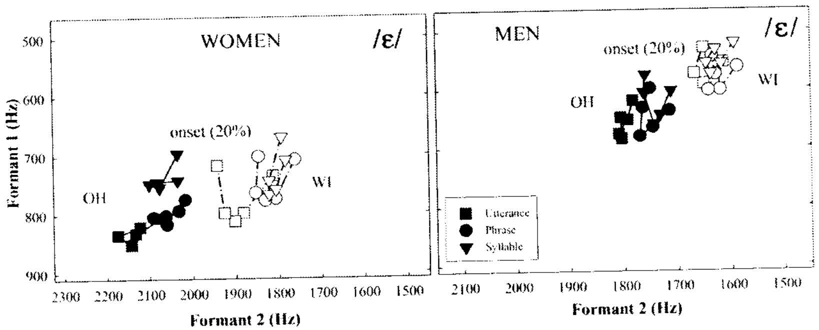


Figure 4: Mean F1 and F2 frequencies for /ɛ/ in *bet* at five locations in the vowel, corresponding to 20%, 35%, 50%, 65% and 80% of vowel duration plotted separately for men and women

The frequency values across prosodic positions (Utterance, Phrase, and Syllable) are represented by solid symbols for Ohio and open symbols for Wisconsin speakers.

2002, also see Hillenbrand *et al.* 1995) which is the unsigned linear distance (in Hz) between the start of the vowel and the end of the vowel in the F1 by F2 plane. Here we define vector length as the unsigned linear distance (in Hz) between the 20% and 80% measurements points. Diphthongal or diphthongized vowels will have longer vectors than will monophthongs. Therefore, we expect that vowels in more prominent prosodic positions will have longer vectors because of the greater F1 and F2 frequency change associated with increased prosodic hierarchy.

As shown in Figure 5, the vectors are longer for /e/, which is a diphhongized vowel, than for /ɛ/, which is monophthongal. Despite

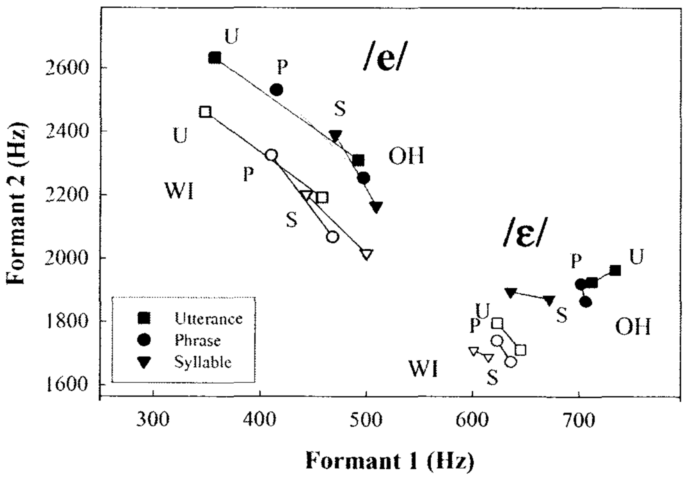


Figure 5: Change of F1 and F2 frequency between the 20%–80% measurement locations representing the "vector length" or the amount of spectral change

The frequency change is plotted for the vowels /e/ in *bait* /ɛ/ in *bet* in three prosodic positions: Utterance (U), Phrase (P) and Syllable (S). The symbols identifying the prosodic level are plotted at the 80% location and a line connects this point to the 20% measurement location. For each speaker group in Ohio (closed symbols) and Wisconsin (open symbols) the values are averaged across speaker gender.

the difference between the vowels, the vectors for both become progressively longer with increasing prosodic prominence. As expected, longest vectors in the highest U-position reflect more diphthongal vowel characteristics than the shortest vectors in the S-position. These relations are particularly evident for /e/. Although the vectors are drastically shorter for /ɛ/ reflecting the monophthongal characteristics of the vowel, one can nevertheless observe that their lengths increase slightly with the increase in the prosodic prominence in Wisconsin productions. Overall, the vectors across prosodic positions tend to be longer for Ohio vowels relative to the Wisconsin vowels, which indicates that Wisconsin vowels may be more monophthongal.

In sum, we found that acoustic characteristics of vowels differ in predictable ways as a function of variation in prosodic prominence. Vowels in higher prosodic positions are longer and have a greater amount of frequency change (*i.e.*, are more diphthongized and show greater formant movement) than vowels in lower prosodic positions. The prosodically induced variation is a source of a gradual shift of a vowel in the acoustic vowel space with increased prosodic prominence, which may be carried out through generations of speakers leading to a vowel shift over time. We also found some dialectal differences, namely that Wisconsin vowels were shorter and more centralized relative to the Ohio vowels. Finally, there seem to be gender-related differences in the productions of vowels across prosodic positions, particularly in the lowest prosodic context.

3. Perception Experiment

The goal of the perception experiment was to determine whether the observed acoustic variation in vowels resulting from prosodic context, speaker dialect and speaker gender has perceptual relevance. We asked three specific questions:

(i) Does prosodically induced acoustic change produce significant differences in perceived vowel quality?

(ii) Is there a perceived phonetic difference in the vowels produced by speakers from two related dialects, such as spoken in Ohio and Wisconsin?

(iii) Do acoustic differences between men's and women's productions in the two dialects produce concomitant differences in perceived vowel quality?

3.1. Listeners, Stimuli and Procedure

Twenty native speakers of American English born and raised in central Ohio (10 men and 10 women, aged 20–46 years) served as listeners. None of the listeners had a known history of speech and hearing disorder and none had participated as a speaker in the earlier study. The experimental stimuli consisted of all instances of the words *bait* and *bet* edited out of their prosodically structured sentences in the production study for a total of 244 (6 sentences x 16 speakers x 3 repetitions). The stimuli were randomized and presented to listeners seated in a sound-attenuating booth over high-quality headphones (Sennheiser HD600). Listeners were required to identify the vowel in each stimulus token using a forced-choice task, with the choices *beet*, *bit*, *bait*, *bet*, *bat* and *but*. These six key words appeared in separate response boxes on a computer screen (along with the appropriate phonetic symbols representing the vowel in each word) and listeners were instructed to click the mouse on the response box corresponding to the vowel heard. After identifying the vowel contained within a given stimulus token (by choosing one of six responses), the listener then rated the goodness of the token's vowel in terms of whether it represented either a good, fair or poor exemplar of the vowel category chosen. Again, listeners indicated their decision by clicking on the appropriate response box. In this experiment, the words *bait* and *bet* spoken by the Ohio and Wisconsin speakers were combined into one experimental block and were not presented separately for each dialect. This was done to prevent listeners from code-switching from one dialect to another and to force them to use the same perceptual criteria while responding to stimulus tokens from each dialect. If there are acoustic differences in vowel characteristics—as a function of prosodic context, dialect or speaker gender—which are perceptually salient, these differences should produce differences in perceived vowel quality reflected in either the choice of phonetic quality or in the goodness ratings.

3.2. Results for the Vowel /e/ in bait

As shown in Figure 6, the percentage of times that the vowel in the stimulus token *bait* was identified as the vowel /e/ was affected by the prosodic context in which the stimulus token had been produced. It was classified as /e/ most often when it occurred in the U-position and least often when it occurred in the S-position. This indicates that listeners responded to the graded differences in acoustic vowel characteristics across prosodic positions observed in the production data. Differences due to speaker dialect were also perceived, as evident from higher identification rates for vowels spoken by Ohio speakers, particularly in the U- and P-positions. The attunement to the Ohio dialect by Ohio listeners was also reflected in the substitution pattern for /e/ whose number of misidentifications as /i/ increased for vowels spoken by Wisconsin speakers.

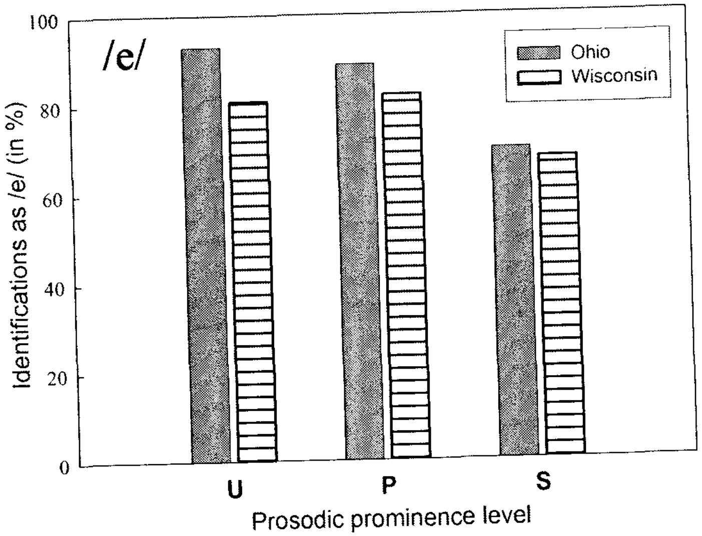


Figure 6: Identification of /e/

The figure shows responses of Ohio listeners to Ohio and Wisconsin productions of *bait* across three positions of prosodic prominence: Utterance (U), Phrase (P) and Syllable (S).

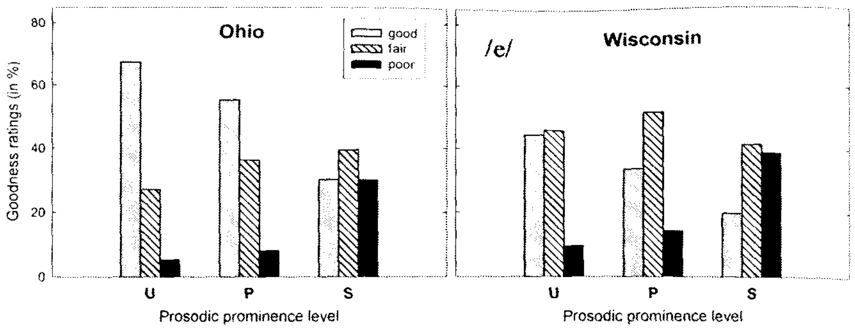


Figure 7: Goodness ratings for /e/ by Ohio listeners responding to Ohio (left) and Wisconsin (right) productions of *bait*

For each identification level across prosodic positions shown in Figure 6, the vowels are rated as good, fair or poor instances of /e/.

Both the prosodically induced variation and dialectal differences produced notable differences in the distribution of vowel goodness ratings. Figure 7 shows the goodness ratings for the response /e/ (*i.e.*, the percentage of times that listeners used goodness ratings of good, fair or poor, when they had identified the stimulus token *bait* as containing the vowel /e/). For vowels spoken by Ohio speakers, the U-position yielded the highest goodness ratings (in terms of the percentage of time the vowel was identified as representing a "good" exemplar of the vowel quality /e/), which decreased in P- and S-positions, respectively. The goodness ratings for Wisconsin vowels were much lower than for Ohio vowels in the U-position and a similar difference is evident in each of the lower prosodic contexts.

Acoustic differences as a function of speaker gender also produced perceived differences in vowel quality across prosodic contexts, as shown in Figure 8. For both Ohio and Wisconsin productions, the percentage of times that the vowel was identified as intended by the speaker were higher for vowels spoken by women than by men, particularly in U- and P-positions.

Goodness ratings provided further insights into the differences related to speaker gender. As displayed in Figure 9, vowels spoken by Ohio women received considerably higher ratings in the U- and P-

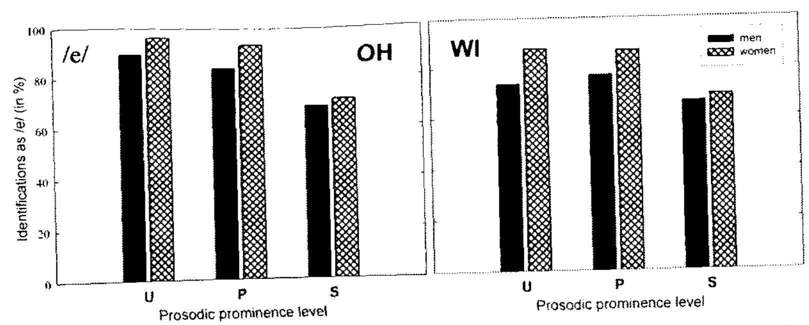


Figure 8: The effect of speaker gender on the perceptual identification of /e/ in *bait* across three positions of prosodic prominence: utterance (U), phrase (P) and syllable (S)

Ohio listeners responded to the vowels spoken by men and women in Ohio (left) and Wisconsin (right).

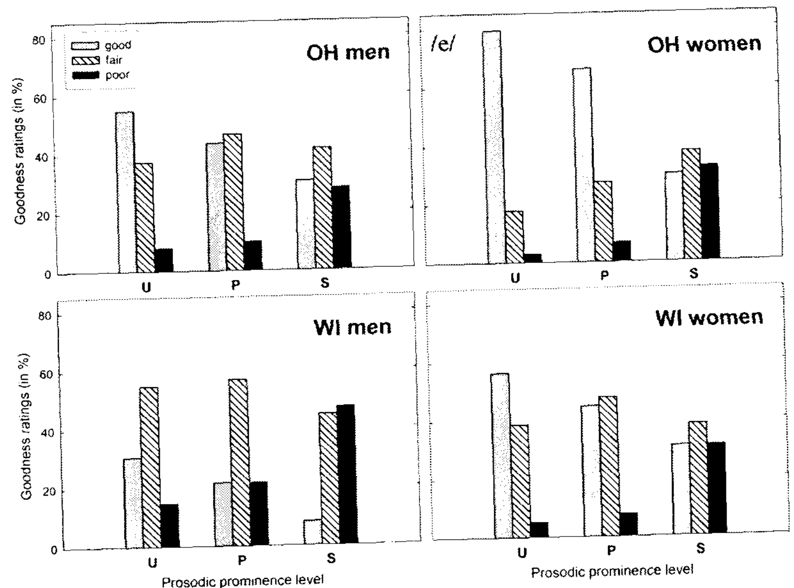


Figure 9: The effect of speaker gender on goodness ratings for /e/ in *bait* across three positions of prosodic prominence: utterance (U), phrase (P) and syllable (S)

Ohio listeners rated the vowels identified separately for Ohio men and women and Wisconsin men and women (see Figure 8) as good, fair, or poor instances of /e/.

positions, respectively, than vowels spoken by Ohio men. Although the goodness ratings were generally lower for Wisconsin vowels, the number of "good" responses decreased drastically for men. However, the ratings as "good" for vowels spoken by women were again higher, reaching a rating level comparable with vowels spoken by Ohio men.

As a whole, the results indicate that variation in prosodic context produced differences in perceived vowel quality. In addition, variations in speaker dialect and speaker gender also produced vowel quality differences. To determine the significance of this variation on vowel identification, the identification data were subjected to the appropriate statistical analysis. In particular, the number of /e/ responses to the *bait* stimulus tokens were examined using a repeated measures analysis of variance (ANOVA) with the within-subject factors prosodic context, speaker dialect and speaker gender. All three main factors were significant. There was a significant prosodic context effect ($F(1, 24) = 12.53$, $p = 0.001$) resulting from the fact that significantly fewer vowels produced in the S-position were identified as /e/ than in either U- and P-positions. However, the significant interaction between prosodic context and speaker dialect ($F(2, 31) = 4.72$, $p = 0.020$) demonstrated that the effect of prosodic context was somewhat different for Ohio as opposed to Wisconsin vowels. In particular, post hoc analysis showed that all three prosodic contexts were significantly different from one another (at the .05 level) in the Ohio vowels. The U-position produced significantly more /e/ responses than either the P- or S-position and the P-position produced more /e/ responses than the S-position. However, for the Wisconsin vowels, although the U- and P-positions produced significantly more /e/ responses than the S-position, they were not significantly different from one another.

There was a significant effect of speaker dialect ($F(1, 18) = 13.79$, $p = 0.002$) resulting from the fact that the Ohio vowels were more often identified as /e/ than the Wisconsin vowels, probably a result of a "mismatch" in dialectal vowel space between the Ohio listeners and the Wisconsin speakers. Speaker gender was also a significant main effect ($F(1, 18) = 22.94$, $p < 0.001$). In particular, there were more /e/ responses to the *bait* tokens when they were produced by a woman than by a man. The significant interaction between speaker gender and prosodic context ($F(2, 32) = 6.64$, $p = 0.005$) revealed that the difference in

the number of /e/ responses between the S-position and either the U- or P-position was greater for women than for men.

A separate analysis was done on the goodness ratings as these data may provide insight into more subtle perceptual differences. The goodness ratings were given the values of 3 for "good," 2 for "fair" and 1 for "poor" and then the mean value of these ratings was determined for all /e/ responses to the stimulus token *bait*. These mean ratings were then analyzed using ANOVA[1] with the within-subject factors prosodic context, speaker gender and speaker dialect. Again, all three main effects were significant. There was a significant main effect of prosodic context ($F(1, 24) = 43.67$, $p < 0.001$) as the mean goodness ratings decreased progressively as the prominence of the prosodic context diminished. Post hoc analysis showed that the mean goodness ratings were significantly higher (at the .05 level) for vowels produced in the U-position than in either the P- or S-positions, and that goodness ratings were significantly higher (at the .05 level) for vowels produced in the P-position than in the S-position. This pattern was found for both the Ohio and Wisconsin vowels. Listeners had significantly higher goodness ratings for Ohio vowels than for Wisconsin vowels ($F(1, 18) = 41.93$, $p < 0.001$) and significantly higher ratings for women's vowels than for men's vowels ($F(1, 18) = 27.21$, $p < 0.001$). However, there was a significant interaction between prosodic context and speaker gender ($F(2, 30) = 6.5$, $p = 0.007$); the difference in mean goodness ratings between the S-position and either the U- or P-positions was greater for women than for men. Although the mean goodness ratings were higher for vowels produced by women than for men in both dialects, the difference was significantly greater for the Wisconsin vowels than for the Ohio vowels producing a significant dialect by speaker gender interaction ($F(1,180) = 5.42$, $p = 0.032$).

[1] We have used ANOVA to analyze the goodness ratings despite the fact that one might argue, as did a reviewer, that these ratings represent ordinal rather than interval data and, thus, might be better analyzed using nonparametric statistics. However, it is common to use repeated-measures ANOVA to analyze such perceptual data, including Equal Appearing Interval Scales and Likert Scales (see, for example, Grabe *et al.*, 1997). Given that fact that ANOVA is such a robust analysis technique, we are confident in the validity of the significant results obtained.

3.3. Results for the Vowel /ɛ/ in bet

As shown in Figure 10, variations in the prosodic context for the stimulus *bet* did not produce the same level of variation in the identification of the vowel as was seen for the token *bait*. However, in keeping with the prosodic hierarchy, the vowel in the S-position was identified as /ɛ/ less often than in the U- and P-positions (which have comparable identification rates). Across all three prosodic contexts, there were no clear differences in identification as a function of dialectal differences, nor were there any clear effects of speaker gender.

However, the vowel goodness ratings may be able to demonstrate more subtle phonetic differences than vowel identifications; in particular, the goodness ratings will allow us to examine within- category

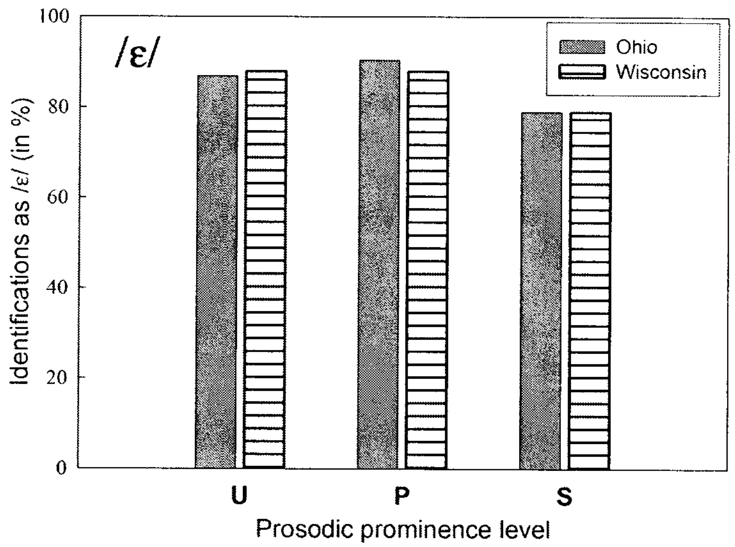


Figure 10: Perceptual identification of /ɛ/

The figure shows responses of Ohio listeners to Ohio and Wisconsin productions of *bet* across three positions of prosodic prominence: utterance (U), phrase (P) and syllable (S).

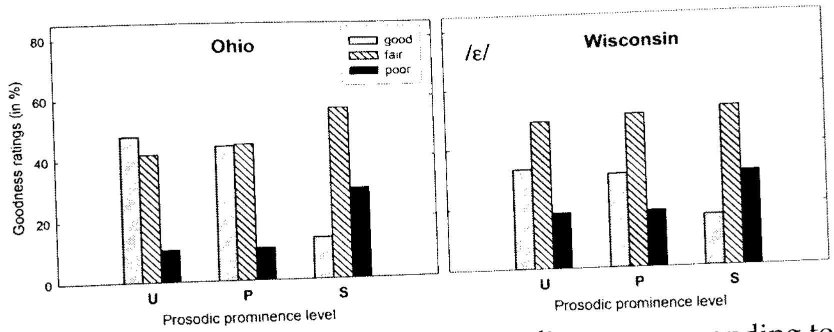


Figure 11: Goodness ratings for /ɛ/ by Ohio listeners responding to Ohio (left) and Wisconsin (right) productions of *bet*

For each identification level across prosodic positions shown in Figure 10, the vowels are rated as good, fair or poor instances of /ɛ/.

variations. Figure 11 shows that for both dialect groups, fewer /ɛ/ responses were rated as "good" as the prosodic context lowered, although the difference between the U- and P-positions is not as large as for the token *bait*. This result matches the identification results found for *bait* and indicates that the prosodically induced acoustic changes produced perceptible within-category vowel quality differences in the stimulus token *bet*. The goodness ratings also provide evidence of a dialect difference as well. In particular, Ohio vowels received more "good" ratings for /ɛ/ responses in both the U- and P-positions than did Wisconsin vowels.

Acoustic differences due to speaker gender did not produce any notable effects on the identification results, as evident in Figure 12. Only relatively minor differences were found in P- and S-positions for Ohio vowels, where men's productions were identified slightly more often as /ɛ/ than were the women's vowels. The reverse situation is found for the Wisconsin vowels in that tokens in the U-position were more often identified as /ɛ/ when spoken by women as compared to men.

However, the pattern of goodness ratings in Figure 13 shows that perceptual differences due to speaker gender and speaker dialect were obtained. Women's vowels produced a larger number of "good" rat-

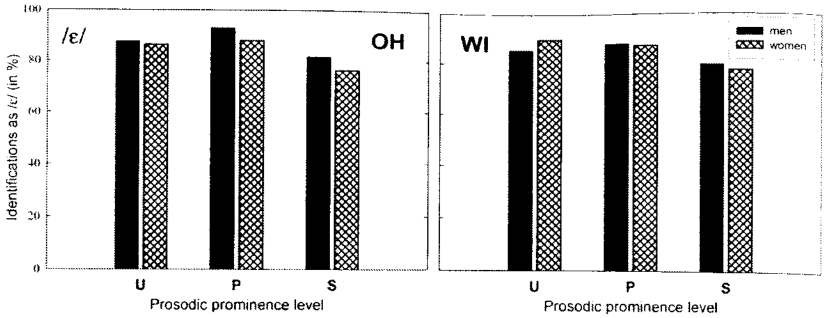


Figure 12: The effect of speaker gender on the perceptual identification of /ɛ/ in bet across three levels of prosodic prominence: utterance (U), phrase (P) and syllable (S)

Ohio listeners responded to the vowels spoken by men and women in Ohio (left) and Wisconsin (right).

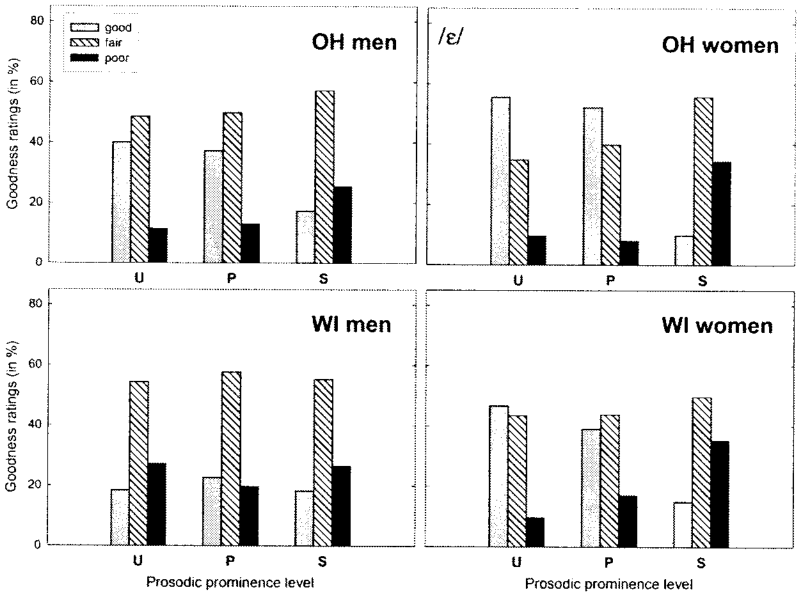


Figure 13: The effect of speaker gender on goodness ratings for /ɛ/ in bet across three positions of prosodic prominence: utterance (U), phrase (P) and syllable (S)

Ohio listeners rated the vowels identified separately for Ohio men and women and Wisconsin men and women (see Figure 12) as good, fair or poor instances of /ɛ/.

ings when the token was identified as /ɛ/ in U- and P-positions than did men's vowels for both dialects. This difference was somewhat greater for Ohio than for Wisconsin vowels. This pattern is consistent with the goodness ratings for the vowel /e/, where women's productions in both dialects and vowels spoken by Ohio speakers produced more "good" ratings.

In summary, variation in prosodic context, speaker dialect and speaker gender did not affect the identification of the *bet* tokens to the same extent they affected the *bait* tokens. However, perceived differences were manifested in goodness ratings, indicating that (1) listeners were able to detect differences between vowels in S-positions relative to both higher prosodic contexts, and that (2) Ohio vowels were perceived more often as "good" exemplars of this vowel category than were the Wisconsin vowels. In addition, for both dialects, vowels produced by women produced higher goodness ratings.

As was done for the *bait* token, both the identification and mean goodness ratings data were analyzed using a repeated-measures ANOVA with the within-subject factors prosodic context, dialect and speaker gender. The results of an ANOVA performed on the identification data showed that only the effect of prosodic context reached statistical significance ($F(2, 31) = 12.15$, $p < 0.001$), indicating that vowels in the S-positions were identified significantly less often as /ɛ/ than vowels in either U- or P-position, which did not differ from each other.

However, the results of another ANOVA performed on mean goodness ratings showed that all three main effects—prosodic context, speaker dialect and speaker gender—were significant, and produced measurable effects on listener judgments. There was significant effect of prosodic context ($F(1, 26) = 86$, $p < 0.001$) and post hoc analysis indicated that vowels in S-position received significantly lower goodness ratings (at the .05 level) than vowels in either the U- and P-positions (which did not differ from each other significantly). The significant effect of speaker dialect ($F(1, 18) = 9.64$, $p = 0.006$) showed that Ohio vowels received higher goodness ratings than did Wisconsin vowels. A significant interaction between prosodic context and dialect ($F(2, 30) = 9.03$, $p = 0.001$) revealed that differences in goodness ratings between S-position and either U- or P-position were greater for Ohio vowels than for Wisconsin vowels. The interaction effect can be interpreted as indicating that the differences produced by the prosodic hierarchy were reduced when there was a mismatch between the dialects of the speakers and listeners.

There was also a significant main effect of speaker gender ($F(1, 18) = 33.93$, $p < 0.001$) resulting from the fact that vowels spoken by women received higher goodness ratings than vowels spoken by men. However, the significant interaction between prosodic context and speaker gender ($F(2, 31) = 20.8$, $p < 0.001$) showed that the differences between the S-position and either the U- or P-position were greater for women than for men.

4. Summary and Discussion

We examined acoustic variation in vowels as a function of prosodic context together with the perceptual consequences of that variation. This furthers our understanding of the role of prosodic prominence in synchronic variation which may lead to vowel chain shifts over time by means of sound transmission across generations. The results of the production study show that systematic changes occur both in the spectral and durational characteristics of vowels as a function of prosodic structure. In particular, vowels produced in the U-position are more likely to exhibit formant values farther away from the center of the vowel space, to show greater spectral change over time (diphthongization), and to be longer. In turn, listeners demonstrated sensitivity to these acoustic variations in terms of both vowel identification and goodness ratings. Generally, vowels in lower prosodic positions were misidentified more often and represented poorer exemplars of vowel category than vowels produced in higher prosodic positions.

Durational differences were also closely associated with changes in the prosodic hierarchy. These differences likely contributed to the perceptual effects discussed, especially between vowels in U- or P-positions relative to S-position, which affect the amount of frequency change in the course of duration of a vowel. That is, as shown in the acoustic data, vowels in higher prosodic positions are longer and exhibit more frequency change, which corresponds to their greater diphthongization as compared to vowels in the lower prosodic positions. Apparently, this cue is important to the listeners in making decisions about vowel quality, which is also in line with the experimental results showing that listeners rely on formant frequency movements throughout the vowel rather than on static vowel properties (Hillenbrand and Nearey 1999).

We also examined the effects of speaker dialect and speaker gender on this type of variation, both in terms of vowel production and vowel perception. The Ohio vowels and the Wisconsin vowels differed in terms of position within the acoustic vowel space and in terms of extent of formant change over time. Results revealed that listeners raised in one dialect area misidentify vowels in another dialect more often than vowels from their own dialect, even when the two dialects are closely related. This was evident particularly for the vowel /e/. However, even if the vowels were identified correctly as intended by the speaker, listeners were still able to detect more subtle differences between the vowels in the two dialects, as shown by the goodness ratings for either /e/ or /ɛ/. Vowels spoken by Ohio speakers were rated by Ohio listeners as good exemplars of the vowel quality more often than vowels spoken by Wisconsin speakers. Although Ohio listeners responded well to prosodic variation in Wisconsin vowels, the vowels did not sound to them as good as the vowels from their own dialect spoken in Ohio. This indicates that attunement to the "native dialect" is strong and listeners are able to detect even subtle dialectal differences.

Speaker gender may play a role in the position of vowels within the vowel space, and women may demonstrate more "extreme" positions—especially in the U-position—than do men. Gender-related acoustic variation was shown to play a significant role not only in vowel goodness ratings across prosodic positions but also in the perception of vowel quality, as the results for the vowel /e/ show. Uniformly, vowels spoken by women were rated as "good" exemplars of vowel quality more often than vowels spoken by men and, for /e/, were misidentified less often than vowels spoken by men. This effect was evident for both Ohio and Wisconsin speakers, indicating that listeners make use of the acoustic differences between men's and women's productions in making perceptual decisions. The present perception results support the sociolinguistic work on sound change, including vocalic chain shifts, which overwhelmingly points to women as the "leaders" of change, that is, as producing more advanced forms of a given change (see Labov 2001 for a detailed review of literature on this issue). Also, Kuhl and colleagues (Liu *et al.* 2003) showed that women produce even greater acoustic changes in vowels when speaking to infants than to other adults. This underscores the role of women in sound transmission across generations as they provide more advanced forms of vowel change in their input to language acquired by children (see also Foulkes *et al.* 2005).

Addressing the issue of the role of prosodic variation in vowels as one motivation for sound change, the present results provide at least partial explanation for two basic, long accepted principles of vocalic chain shifting, namely that long vowels rise and short vowels fall. The synchronic variation associated with greater and lesser prosodic prominence, as observed in our acoustic and perception data, opens a way to explicate the gradualness of the phonetic vowel change across generations of speakers coming from the same geographic area.

For the tense (or long) vowel /e/, our synchronic results parallel the behavior reported for long/tense vowels in chain shifts. First, the vowel gradually rises under greater prosodic prominence, in step with the diachronic Principle I. Second, this vowel shows much greater format movement over the duration of the vowel at higher levels of prominence, which parallels a key aspect of chain shifts, diphthongization of long vowels in the course of the changes.

While it is clear that chain shifts can include the development of diphthongs from long vowels, there seems to be little agreement about precisely how or why this happens. To gain more insight into the variation in the diphthongal properties of vowels, we have employed in our acoustic study a measure of formant frequency change over the course of vowel production, called "vector length." This measure serves as an estimate of how "diphthongized" a long vowel might become in the most prominent prosodic positions. The perception data provide evidence that listeners are sensitive to this type of change in the degree of diphthongization and use this information in making decisions about vowel quality and goodness.

For lax (or short) /ɛ/, the literature on chain shifts allows less firm predictions. Most notably, Principle II, that short vowels lower, is relatively ill-attested (following Labov 1994) and the securely attested

patterns of change, either historical or underway at present, show great variability (as shown especially by Gordon 2001). Still, there are clear parallels, notably that we find lowering of /ɛ/ with greater prosodic prominence, in line with diachronic Principle II. The present perception results show that prosodic structure does influence listeners' responses, even if the vowel is relatively monophthongal and much shorter than the vowel /e/.

The systematic variation resulting from the prosodic structure affects vowel perception regardless of speaker dialect and speaker gender. We found that dialect-specific patterns of formant movement and gender-related acoustic differences did not obscure the prosodic effects. These effects will always be present in the speech of each generation of speakers. The cross-generational sound transmission may occur gradually, as more prosodically prominent forms in speech of older generation are perceived as a norm by each younger generation of speakers. While additional experimental evidence needs to be provided for this position, we conclude that prosodic prominence has a potential to provide a synchronic principle parallel to the diachronic principles of chain shifting.

References

Ferguson, Sarah Hargus, and Diane Kewley-Port. 2002. "Vowel Intelligibility in Clear and Conversational Speech for Normal-Hearing and Hearing-Imparted Listeners," *Journal of the Acoustical Society of America* 112: 259–71.

Foulkes, Paul, Gerard Docherty and Dominic Watt. 2005. "Phonological Variation in Child-Directed Speech," *Language* 81: 177–206.

Grabe, Esther, Carlos Gussenhoven, Judith Haan, Erik Marsi and Brechtje Post. 1997. "Preaccentual Pitch and Speaker Attitude in Dutch," *Language and Speech* 41: 63–85.

Gordon, Matthew. 2001. *Small-Town Values and Big-City Vowels: A Study of the Northern Cities Shift in Michigan* [= *Publications of the American Dialect Society* 84]. Durham, NC: Duke University Press.

Hillenbrand, James, Laura A. Getty, Michael J. Clark and Kimberlee Wheeler. 1995. "Acoustic Characteristics of American English Vowels," *Journal of the Acoustical Society of America* 97: 3099–11.

Hillenbrand, James M., and Terrance M. Nearey. 1999. "Identification of Resynthesized /hVd/ Syllables: Effects of Formant Contour," *Journal of the Acoustical Society of America* 105: 3509–23.

Jacewicz, Ewa, Joseph Salmons and Robert A. Fox. Forthcoming. "Prosodic Conditioning, Vowel Dynamics and Sound Change," *Variation in Phonetics and Phonology*. Lahiri, A. (ed.). Berlin: Mouton de Gruyter.

Labov , William. 1994. *Principles of Linguistic Change. 1: Internal Factors*. Oxford: Blackwell.

________. 2001. *Principles of Linguistic Change. 2: Social Factors*. Oxford: Blackwell.

Liu, Huei-Mei, Patricia Kuhl and Feng-Ming Tsao. 2003. "An Association between Mothers' Speech Clarity and Infants' Speech Discrimination Skills," *Developmental Science* 6(3): F1–F10.

Sievers, Eduard. 1876/1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf and Härtel.

Stockwell, Robert. 1978. "Perseverance in the English Vowel Shift," *Recent Developments in Historical Phonology*, Fisiak, J. (ed.). The Hague: Mouton, 337–48.

Thomas, Erik R. 2002. "Sociophonetic Applications of Speech Perception Experiments," *American Speech* 77: 115–47.

Robert Allen Fox and Ewa Jacewicz:
Dept. of Speech and Hearing Science
The Ohio State University
110 Pressey Hall, 1070 Carmack Road
Columbus, OH 43210-1002
fox.2@osu.edu
jacewicz.1@osu.edu

Joseph Salmons:
Center for the Study of Upper Midwestern Cultures
University of Wisconsin-Madison
901 University Bay Drive
Madison, Wisconsin 53705-2269
jsalmons@wisc.edu