Prosodic conditioning, vowel dynamics and sound change

Ewa Jacewicz, Joseph Salmons & Robert Allen Fox

1. Introduction

This paper reports on research aimed at bridging two important but disparate approaches to understanding variability and change in the realization of vowels: (1) phonetic work on how prosody shapes vowel characteristics and (2) diachronic and sociolinguistic work on ‘ chain shifting’ in vowel systems. We will examine two types of phonetic variation: changes in the dynamic formant patterns of vowels as a function of levels of prosodic emphasis and phonetic quality differences associated with regional dialects of American English. Understanding the nature and origins of such phonetic variation will, we believe, provide insight into the sources of sound change over time.

With production and perception data from two Midwestern dialects, we explore parallels between the effects of prosodic emphasis on vowel characteristics and the general directions of change over time and across languages observed for chain shifts. Those parallels shed light on a conundrum known as the ‘perseverance problem’ in chain shifts (Stockwell 1978), namely how generation after generation of speakers could continue to move vowels slowly but consistently in particular directions. The teleological notion that generations are all working toward the same goal in such changes must be rejected, and hypothesized social motivations, while promising, are incomplete. Furthermore, none of this helps to explain why similar shifts occur time and again across languages and periods. One possible structural explanation, for which a foundation is proposed in this paper, is that emphatic realizations cause some vowels to be perceived as more prominent (and thus more salient) in the continuously varying speech stream. These more emphatic productions may serve as the primary impetus for vowel shift over time in that more forceful realizations of vowels are transmitted across generations as less marked ones.

We see variation in vowel enhancement as interacting with vowel properties to bring about a shift as language is handed on from one generation to the next. As vowels are acquired by new generations, the older generation’s more emphatic renditions map to less prominent realizations in the speech of younger speakers, yielding gradual shift-like patterns of change. We illustrate this informally below, where the interlocking circles represent the varying realizations of the tense vowel /e/ under the relevant degrees of stress, with an older generation's prominent realizations mapping to a younger generation's prototypical form.

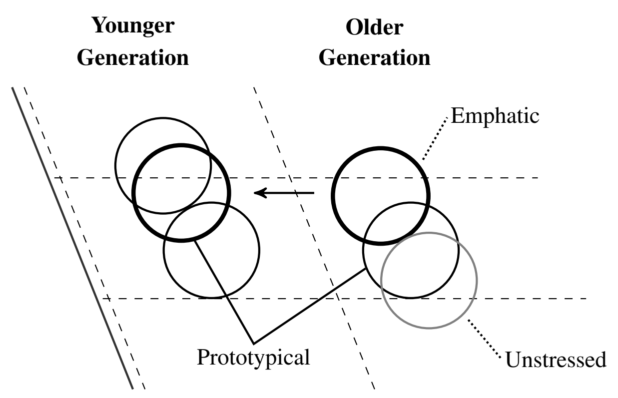

Figure 1: Schematic representation of chain shifts as a function of graded emphasis (or prominence) in vowel production involving two generations of speakers.

1.1. Regional differences in American English vowels and principles of vowel chain shifts

A body of work since Labov, Yeager and Steiner (1972) through Labov (1994, 2001) and Labov et al. (2006) argues that urban areas from Buffalo, New York to Madison, Wisconsin are undergoing a new chain shift (the Northern Cities Shift), including the raising of the vowel in words like *bad* and *bag*, which as a result can sound to speakers from other areas like *bed* and *beg* or even *bid* and *big*. In the Southeast, in contrast, lax front vowels, as in *bit*, are tensing, lengthening and even diphthongizing, so that *bit* can be pronounced like [biːjɪt], while tense front vowels, like in *beat*, tend to become laxer and more centralized. In the West, we find mergers notably of the vowels in *cot* and *caught*, so that *Don* and *Dawn* now sound identical for many speakers. Across much of the U.S. – outside of the Upper Midwest – the tense high back vowel /uː/ is fronting, and its mid counterpart /oː/ is fronting in the south and east. These and similar changes constitute a profound and rapid regional diversification of everyday American English speech.

These differences define the core of the emerging field of sociophonetics, and are central topics in sociolinguistics, dialectology and historical linguistics. Most theoretical discussions in the literature are organized around the notion of *chain shifts*, the tendency, in at least some languages, for individual vowels not to change in isolation from each other over time, but rather to move apparently linked together through the traditional vowel space. Such patterns have long been recognized, most famously in the so-called *Great Vowel Shift* of English, and the vast literature on chain shifts has been reviewed recently by Labov (1994, 2001), Gordon (2001, 2002) and others. A consistent pattern of chain shifting has 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"):

(1) Principle I: In chain shifts, long (tense) vowels rise.
Principle II: In chain shifts, short (lax) vowels fall.

These patterns are widespread and 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). When compared closely to the historical record of attested chain shifts, Principle I is broadly and securely attested – Labov (1994: 122, and elsewhere) alone gives 15 examples of long vowels rising in apparent chains. But Principle II is relatively ill-attested (Labov gives only two examples), and short vowels clearly show far more complex and variable patterns of change than this principle predicts. Take, for example, the already-mentioned Northern Cities Shift, currently in progress across urban areas from upstate New York westward to Wisconsin. In this shift /ɪ/ is generally lowering to /ɛ/, but /ɛ/ is generally thought to be backing to /ʌ/. At the same time, some studies like Gordon (2001) show widespread variation here. For example, /ɛ/ lowers or backs, while /ʌ/ backs, raises or lowers. We conclude that chain shifting of /ɛ/ typically involves lowering and/or backing.[1]

The present study investigates changes to the acoustic characteristics of vowels (including durational and spectral changes) spoken with graded degrees of emphasis in two closely related regional dialects of Midwestern American English, central Ohio and southern Wisconsin. Wisconsin English is assumed by sociolinguists to be in an early stage of the Northern Cities Shift whereas English spoken in central Ohio is not currently participating in any chain shift. It is therefore of interest whether changes to the acoustic structure of vowels caused by the variation in vowel enhancement are different in these two dialects. If diachronic principles are rooted in synchronic variation, we would expect long vowels to rise in the acoustic space in most prominent or emphatic productions. Although the extent of this raising in each dialect, as well as perceptual responses to such raising effects, are difficult to predict on the basis of broadly defined diachronic principles, we would generally expect more enhanced short vowels to lower and/or back in the acoustic space although changes in progress related to Northern Cities Shift in Wisconsin English may somewhat obscure these effects. The study considers acoustic and perception data in the search for synchronic evidence which could serve as a basis for explanation of diachronic principles of sound change.

2. The acoustic study

We examine the nature of acoustic variation that results from a specific prominence-defined position of a vowel in an utterance. We predict that vowels in the most prominent or emphatic positions in an utterance will be acoustically more enhanced than vowels in the positions of intermediate prominence, but the latter will still show more of such enhancement than vowels in the weakest position. In this paper, we refer to the position of highest prominence as the Utterance-initial position (U) followed by the Phrase position (P) and Syllable position (S), respectively.[2] The acoustic measures employed here include vowel duration and the frequencies of the first two formants, F1 and F2. Formant frequencies are measured at five different temporal points in the vowel to allow an approximation of formant movement in the course of vowel's duration. The prediction is that vowels in U-position are longer and their formant values are more extreme as compared to P- and S-positions, respectively.

2.1. Method

2.1.1. *Speakers*

Sixteen native speakers of two Midwestern varieties of American English participated in the experiment. Eight speakers (four males and four females) were born and raised in the Columbus, Ohio area, and eight speakers (four males and four females) were born and raised in the area of Madison, Wisconsin. Their ages ranged from 16 to 31 at the time of recording. All speakers were students, either high school students (Columbus) or undergraduate and graduate students of various majors enrolled at The Ohio State University or University of Wisconsin-Madison. The speakers were unaware of the purpose of the experiment.

2.1.2. *Stimulus materials*

The long vowel /e/ and the short vowel /ɛ/ were selected and placed in existing monosyllabic words in /b_t/ environments. Speakers produced the following sentences containing the words *bait/bet* along with distractor sentences in three distinct positions of prosodic prominence, here labeled U, P, S:

(2) *The strongest U-position (representing the stressed syllable in a noun phrase in utterance-initial position)*
‘Bait shop’ is what I said.
‘Bet some’ is what I said.

(3) *The intermediate P-position (representing the stressed syllable in a noun-phrase in non-initial position in an utterance)*
She said the bait shop was closed.
He said the bet slips were here.

(4) *The weakest S-position (representing the unstressed but unreduced syllable in a morphological compound)*
Shark bait with flavor seems hard to find.
Risky bets are nice but safe bets are better.

These three different sentence structures were used to elicit the productions of [beɪt] and [bɛt] with consistently different degrees of prosodic prominence, phonetically reflected in graded differences in emphasis. We make no theoretical claims regarding the phonological/prosodic structure of these phrases.[3]

2.1.3. *Procedure*

Each speaker read one sentence at a time in a random order for a total of three repetitions. Recording, procedure, and initial data processing were controlled by a program written in MATLAB. The sentences were recorded directly onto a hard drive at a 44.1-kHz sample rate while the speaker was seated in a sound-treated IAC booth. A head-mounted microphone (Shure SM10A) was used, placed at a 1-inch distance from the lips. Each speaker was instructed to read a sentence appearing on the monitor screen in a way typical of his/her conversational speaking style ("as you normally say it while talking to someone"). No other specific reading instructions were given to the speakers.

2.1.4. *Vowel measurements*

Vowel duration and formant movement across vowel duration served as primary measures of the effect of prosodic prominence on the acoustic changes in vowels. Vowel duration was measured from waveform with reference to a spectrogram, using TF32 speech analysis software (Milenkovic 2003). Vowel onsets and offsets included formant transitions. The initial measurement point was located at the first positive peak in the periodic waveform and the final measurement location was at the beginning of the stop closure. All segmentation decisions were later checked and corrected (and then re-checked) using a MATLAB program that displayed the segmentation marks superimposed over a display of the token's waveform. The frequencies of the first two formants of each vowel were then obtained at positions corresponding to 20%, 35%, 50%, 65% and 80% of duration of the vowel. These five formant measurements should exclude immediate consonant influences during CV or VC formant transitions. The stimulus tokens were downsampled to 11.025 kHz and premphasized (98%) prior to spectral analysis. Formant frequencies were estimated with a MATLAB program that utilized a 14-order LPC analysis with a 15 msec Hanning window centered over each measurement location.

2.2. Results

2.2.1. *Vowel duration*

Mean vowel durations are shown in Figure 2. As is evident, for both the long vowel /e/ and the short vowel /ɛ/ duration was sensitive to the degree of vowel enhancement. Higher prominence was reflected in longer durations, particularly in the strongest U-position. The differences between intermediate and weak positions were small but also consistent with the graded prosodic prominence except of /e/ in the S-position in Wisconsin productions. The dialectal differences were manifested in longer durations of vowels spoken by Ohio speakers.

To assess the significance of these results, a mixed design analysis of variance (ANOVA) was conducted (using the SPSS v. 13 software package) with the within-subject factor prosodic position (U, P, S) and the between-subject factors speaker gender (male, female) and speaker dialect (Ohio, Wisconsin). For all reported within-subjects significant main effects and interactions, the degrees of freedom were Greenhouse-Geisser adjusted when there were significant violations of sphericity. Post hoc analyses, when reported—both for vowel duration and formant analyses—were completed using additional ANOVAs on selected subsets of the data (with appropriate F-tests) and either the Tukey test (for between-subject factors) or GLM contrasts (for within-subject factors).

For the vowel /e/, there was a significant main effect of prosodic position ($F(1.7, 20.8) = 13.64$, $p < 0.001$, $\eta^2 = 0.532$) and post-hoc tests showed that, for each dialect, vowels in the U-position were significantly longer than vowels in either the P- and S-position, which did not differ significantly from each other. Thus, the observed graded differences between the P- and S-positions in the Ohio productions were not large enough to reach statistical significance. The effect of speaker dialect was not significant although the Ohio speakers produced longer vowels than the Wisconsin speakers due to the dialectal differences in the U- and P-positions. Similarly, although the effect of gender was not significant, the mean vowel durations were longer for the female speakers than for the male both in Ohio and Wisconsin.

For the vowel /ɛ/, the main effect of prosodic position was significant ($F(1.5, 18.2) = 9.4$, $p = 0.003$, $\eta^2 = 0.440$), with significantly longer vowels produced in the U- position than in either the P- or S-positions. This result is consistent with that for the vowel /e/, indicating that the graded differences

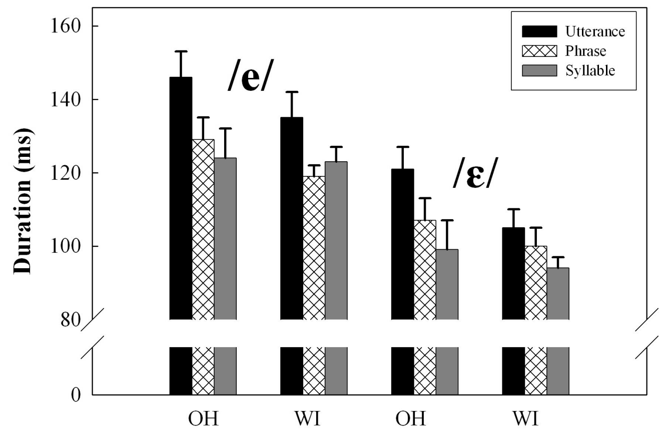

Figure 2: The effect of prosodic prominence on the duration of the vowels /e/ and /ɛ/ in the production of Ohio and Wisconsin speakers. The bar graphs represent mean values. Error bars indicate one standard error.

between the P- or S-positions for either dialect were not large enough to reach significance. Vowels spoken by Ohio speakers were again longer than vowels in Wisconsin productions although these differences were not significant. Speaker gender was significant for /ɛ/ ($F(1, 12) = 6.9$, $p = 0.022$, $\eta^2 = 0.364$), indicating that female vowels were longer across all prosodic positions than male vowels.

Overall, the duration results show that vowels in the highest prominence position were significantly longer than those produced in less prominent positions. At present, the remaining differences found in the means displayed in Figure 2, although suggestive with regard to graded effects of prosodic hierarchy, should be regarded as tendencies rather than representing a conclusive statement.

2.2.2. *Vowel formant pattern*

Figure 3 shows, for each prosodic prominence position, the average changes in F1 and F2 frequencies at five locations in the vowel. As can be seen, the specific formant frequency patterns for either vowel /e/ or /ɛ/ are different in the three prosodic positions. Comparing the locations of the vowels in the acoustic vowel space across the three prosodic positions, one can observe that the trend of greater *raising* and fronting of the vowel /e/ is associated with increased prosodic prominence. A similar effect is also evident for the vowel /ɛ/ although more extreme formant frequency values caused by more emphatic vowel productions corresponded to *lowering* and fronting of the vowel with increased prosodic prominence. As can be seen, frequency changes for F2 for /ɛ/ are small as compared to /e/ due to the non-diphthongal nature of this vowel.

As determined by a series of repeated-measures ANOVAs, enhancement due to the prosodic context had a significant effect on the frequencies of both F1 and F2 in the production of /e/ for both Ohio and Wisconsin speakers (see Table 1 in the Appendix for details concerning ANOVA results). Exploring the significant main effect of prosodic prominence on F1 changes, post-hoc analyses showed that the overall frequencies of F1 were significantly lower in the U-position than in the P-position, and the latter were still significantly lower than in the S-position. Consistently, the overall frequencies of F2 were significantly higher in the U-position than in the P-position, and higher in the P-position than in the S-position. These results are consistent with the graded positions of prominence and show that the degree in vowel enhancement influences changes in formant frequencies over the duration of the vowel in a predictable way.

In addition to the main effect of prosodic context, a significant interaction between prosodic position and vowel measurement location (except for F1 in the Wisconsin productions) showed that prosodic position systematically influenced the specific pattern of frequency change over the course of the vowel's duration. In the most prominent U-position, the vowels exhibited more diphthongal qualities in terms of the extent of frequency change (both F1 and F2) over their durations than vowels in P- and S-positions, respectively.

For the vowel /ɛ/, prosodic context had an effect on both formant frequency values and the nature of the frequency change in the course of the vowel's duration (see Table 2 in the Appendix for details concerning ANOVA results). The significant main effect of prosodic position for F1 for both Ohio and Wisconsin speakers revealed that the vowels in S-position had significantly lower overall mean frequencies, which corresponds to their raised positions in the acoustic space relative to the two more prominent prosodic positions. The effect of prosodic position was also significant for F2. The F2 values for /ɛ/ indicated that the vowel was gradually fronted in the course of its duration for the more prominent U- and P-positions and it was centralized

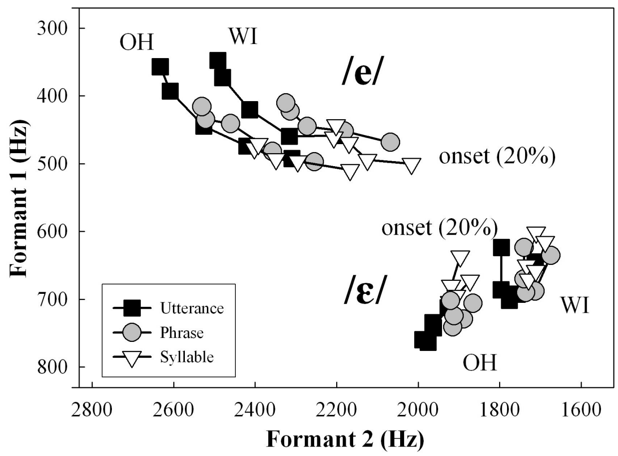

Figure 3: Mean F1 and F2 frequencies for /e/ and /ɛ/ at five temporal locations in the vowel, across prosodic positions. The 20% measurement location is marked as 'onset' and the lines connect each consecutive measurement for each prosodic level.

in the least prominent S-position. This directional change in F2 resulting from an interaction of prosodic context and a particular vowel measurement location could not be detected if F2 values were measured at one location only.

With regard to dialectal differences displayed in Figure 3, Wisconsin vowel productions appear to be more centralized in comparison with Ohio vowels. Because we do not have measures of entire vowel spaces for Ohio and Wisconsin at present, we do not know whether the dialectal differences include the shift of an entire vowel space toward the back for Wisconsin speakers. Based on the present data only, we see that both vowels /e/ and /ɛ/ are more fronted in Ohio relative to vowels spoken in Wisconsin.

2.2.3. *Vector length*

Analyses of F1 and F2 measures, separately, provide indirect insight into the total trajectory of vowel movement in the F1 by F2 plane. In order to examine the magnitude of this movement, we utilized the measure "vector length" or VL (Ferguson and Kewley-Port 2002, also see Hillenbrand et al. 1995) which is the unsigned linear distance (in Hz) between the start of the vowel (i.e., the location of the 20% point) and the end of the vowel (the 80% point) in the F1 by F2 plane. Diphthongal or diphthongized vowels will have longer vector lengths than will monophthongs, which represent their greater amount of frequency change. Consequently, we expect that vowels in more prominent prosodic positions will have longer vector lengths because of the greater F1 and F2 frequency change associated with increased prosodic prominence.

Figure 4 shows the average VL differences for the vowel /e/ as a function of prosodic context. For the Ohio speakers, the VL of the vowels varied significantly by prosodic position ($F(1.9,\ 11.3) = 40.93$, $p < 0.001$, $\eta^2 = 0.872$), with the U-, P- and S-positions each differing significantly from the other. Longer VLs were found in the more prominent prosodic positions. The Wisconsin speakers exhibited similar patterns for VL. There was a significant prosodic context effect ($F(1.2, 7.3) = 12.76$, $p = 0.007$, $\eta^2 = 0.680$) with the U-, P- and S-positions differing significantly one from the other. As for the Ohio speakers, longer VLs were associated with greater prosodic prominence. The longer vectors reflected more diphthongal vowel characteristics, which significantly increased from the least prominent prosodic position to the most prominent, consistent with the prosodic hierarchy.

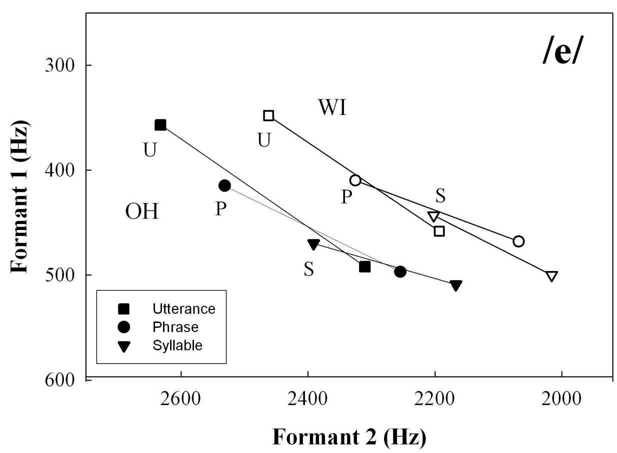

Figure 4: Change of F1 and F2 frequency between the 20% and 80% measurement points for the vowel /e/, representing the ‘vector length’ or the amount of frequency change across prosodic positions. The symbols identifying the prosodic level (U, P, S) are plotted at the 80% location and a line connects this point to the 20% location.

Based on a visual comparison of VLs for Ohio and Wisconsin vowels in Figure 4, one can observe that VLs in both U- and P-positions are longer for Ohio vowels. This indicates that the vowel /e/ spoken in Ohio is more diphthongal compared to that spoken in Wisconsin, at least in more prominent prosodic positions.

As might be expected, VLs for the monophthongal /ɛ/ were, in general, smaller than those for /e/ for both Ohio and Wisconsin speakers but there were some indications that prosodic context affected VLs for /ɛ/ as well (see Figure 5). However, these effects were not as strong and somewhat mixed in terms of directionality of the frequency change across prosodic positions. For both Ohio and Wisconsin speakers, ANOVA of the VL data yielded no significant effect as a function of prosodic context. As shown in Figure 5, there is much more variation in terms of VL differences between the two dialects, particularly for the P- and S-positions.

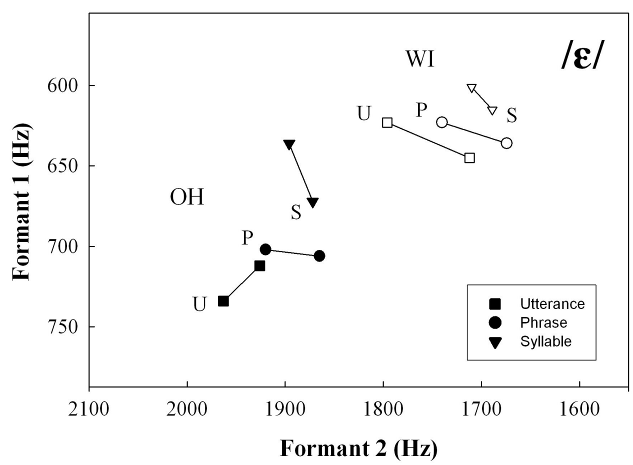

Figure 5: Change of F1 and F2 frequency between the 20% and 80% measurement points for the vowel /ɛ/, representing the 'vector length' or the amount of frequency change across prosodic positions. The symbols identifying the prosodic level (U, P, S) are plotted at the 80% location and a line connects this point to the 20% location.

3. Discussion of the acoustic study

Investigating differences in vowel enhancement as one source of systematic variation in vowels, this study shows that greater prosodic prominence can produce notable changes to both vowel duration and the frequencies of formants over the duration of a vowel. These changes result in progressive raising of /e/ and progressive lowering of /ɛ/ in the acoustic vowel space with each higher prominence position.

Considering duration, vowels in the prosodically strongest U-position had significantly longer durations than in either P- or S-positions. This was true both for the tense vowel /e/ and the lax vowel /ɛ/. The lack of significant differences between the two less prominent prosodic positions may indicate that duration alone is not a strong indicator of prosodic hierarchy despite seemingly straightforward predictions about temporal differences to be found across prosodic levels. It is worth noting that results of other studies are even less conclusive in detecting the relation between prosodic strengthening and vowel duration than the present data, including languages with considerably different accentual/prosodic systems. For example, Onaka (2003) reported no stable pattern of vowel duration in postboundary position (V2) across four prosodic domains for the two Japanese subjects in the study. In another study, Cho and Keating (2001) reported mixed results in the duration of V2 across prosodic domains in Korean. The differences in vowel duration as a function of prosodic position were statistically not significant and the direction of variation differed depending on the domain-initial consonant.

Although not significant, the dialectal differences in vowel duration between the Ohio and Wisconsin variants showed a systematic pattern, which has also been found in another study conducted with different speakers from these two dialects (Jacewicz et al. 2007). Namely, under comparable experimental conditions, Ohio speakers produced longer vowels than Wisconsin speakers. These results were rather unexpected as such cross-dialect differences in vowel duration have not been previously reported in American dialects, although this topic warrants further investigation.

Examining the formant pattern, we measured the frequencies of F1 and F2 at five different locations over the duration of each vowel to capture the dynamic nature and extent of frequency change and its potential variations resulting from prosodic context. This technique proved successful in tracking variation in formant frequencies at different prosodic positions. Measured in this way, the frequency data revealed strong significant effects of prosodic position, measurement location and, most importantly, the interaction between these two which indicated a progressive change in location of the vowel in the acoustic space under greater prominence such as raising and fronting of /e/ and lowering and fronting of /ɛ/.

The vector length analysis further assessed the amount of frequency change over each vowel's duration, showing that vowels in higher prosodic prominence positions become more diphthongal than vowels in lower positions. These changes were also evident for the monophthongal /ɛ/, especially in the Wisconsin data, although generally the vectors were much shorter as compared to the diphthongal vowel /e/.

Relating our acoustic results to diachronic vowel changes, we see strong parallels to the behavior reported for long/tense vowels in chain shifts. Most importantly, the vowel /e/ rises under greater prosodic prominence, in step with diachronic Principle I. For lax (or short) /ɛ/, the literature on chain shifts allows less firm predictions. Still, there are clear parallels, notably that we find lowering of /ɛ/ with greater prosodic prominence, in line with diachronic Principle II. The present results demonstrate ways in which both of these shifts might occur, suggesting that prosodically structured variation in vowels is one motivation for the direction of movement in chain shifts.

Although we expected similar acoustic effects as a function of vowel enhancement for both Ohio and Wisconsin productions, we also expected some differences related to the fact that Wisconsin vowels participate in the Northern Cities Shift whereas vowels spoken in central Ohio do not undergo any shift. We found Ohio vowels to be longer and more fronted relative to Wisconsin productions. In addition, the tense /e/ appears to be more diphthongal. These findings allowed us to formulate specific predictions about the perception of these dialectal differences in vowel characteristics.

For /e/, we can predict that the most emphatic productions in the positions of highest prominence will yield the highest identification of the vowel for Ohio productions but not for vowels spoken by Wisconsin speakers. This is because Ohio vowels represent clearer exemplars of the category relative to Wisconsin vowels, being longer, more fronted, and more diphthongal. Conversely, Wisconsin vowels may sound more monophthongal relative to Ohio vowels, which may result in confusions with neighboring monophthongs /i/, /ɪ/, or /ɛ/.

For the short /ɛ/, we predict more confusions with neighboring short vowels, particularly in the lowest prominence position. Since Ohio /ɛ/ is more fronted than Wisconsin /ɛ/, we would expect the vowel to sound more like an

/æ/ in Ohio productions but not as much when spoken by Wisconsin speakers. Because of its more centralized position in the acoustic space relative to Ohio vowels, Wisconsin /ɛ/ may be perceived more often as /ɪ/ or /ʌ/.

4. The perception study

A vowel identification experiment was conducted to verify our predictions about perceptual responses. Listeners from either Ohio or Wisconsin listened to the productions of both Ohio and Wisconsin speakers. For each dialect, listeners and speakers came from the same geographic areas.

4.1. Stimuli, listeners, and procedure

Stimuli consisted of all instances of the words “bait” and “bet” spoken by all 16 speakers in the production study (8 Ohio speakers and 8 Wisconsin speakers). The words were edited out of sentences (see 2.1.2.) and presented randomly to the listeners for identification. Twenty listeners who were born and raised in central Ohio (10 men and 10 women) and 9 listeners born and raised in southern Wisconsin (5 men and 4 women) responded to all 288 isolated words (2 vowels x 3 prosodic levels x 3 repetitions x 8 speakers x 2 dialects) in one block in one-alternative forced choice task with the choices “beet, bit, bait, bet, bat, but.” After identifying the vowel, the listener then rated its goodness in terms of whether it represented a good, fair, or poor exemplar of the vowel category chosen. Sound was delivered over Sennheiser HD600 headphones to a listener seated in a sound-attenuating booth. The experiment was under computer control, using a program written in MATLAB.

4.2. Results

Only the identification data are presented here as they are most relevant to the present focus. A more exhaustive account of the perception results including vowel goodness ratings, gender-related differences, and statistical treatment can be found in Fox et al. (2006).

Overall, listeners demonstrated sensitivity to the acoustic variation in vowels as a function of differences in prosodic prominence in identifying the vowels as speakers intended. Generally, vowels produced in lower prosodic positions were misidentified more often than vowels in more prominent positions.

4.2.1. *Identification of /e/*

Figure 6 shows the average identification of /e/ by Ohio listeners responding to Ohio vowels (left panel) and to Wisconsin vowels (right panel). Of interest are both identification rates and number of confusions with neighboring vowels. For Ohio productions, listeners responded to the graded differences in acoustic vowel characteristics across prosodic positions observed in the production data. The vowels were classified as /e/ most often when they occurred in the U-position and least often when they occurred in the S-position. For Wisconsin productions, identification rates were lower. Of particular interest is that vowels in the U-position were identified as intended by the speakers only 81% of the time as compared to 93% of the time for Ohio productions. This suggests that Ohio listeners perceived a dialectal difference for vowels produced in most prominent prosodic position, which represented the clearest exemplar of the category. We will return to this in general discussion.

Results from a repeated measures ANOVA showed that these differences were significant. In particular, the significant interaction between prosodic position and speaker dialect ($F(2, 31) = 4.72$, $p = 0.020$, $\eta^2 = 0.208$) explored in subsequent post-hoc analyses demonstrated that listeners' responses to all three prosodic positions were significantly different from one another for the Ohio vowels but not for Wisconsin. For the latter, although the U- and P-positions produced significantly more /e/ responses than the S-position, they were not significantly different from one another.

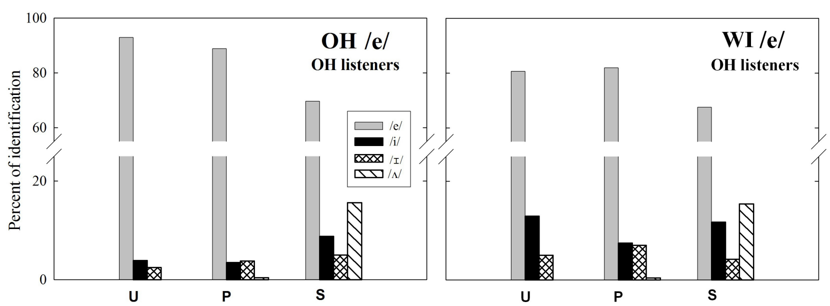

Figure 6: Identification of /e/ across three prosodic levels (U, P, S) for Ohio listeners responding to Ohio productions (left) and Wisconsin productions (right).

The results for Wisconsin listeners are shown in Figure 7. It is striking that vowels in Ohio productions were identified as intended by the speakers more often than vowels in Wisconsin productions (see the parallel findings of Labov and Ash 1997, also discussed in Labov 2001: 490). In particular, identification rate of Ohio /e/ in the U-position was much higher (97%) than that of Wisconsin /e/ (83%). Also, for Wisconsin productions, acoustic differences as a function of prosodic context did not yield differences in listeners' responses, and vowels in all three prosodic contexts were classified as intended by the speakers about 83% of the time. However, Wisconsin listeners perceived differences due to prosodic context in Ohio productions, and vowels in the lowest S-position were still perceived more often as intended by the speakers (87%) than vowels in Wisconsin productions.

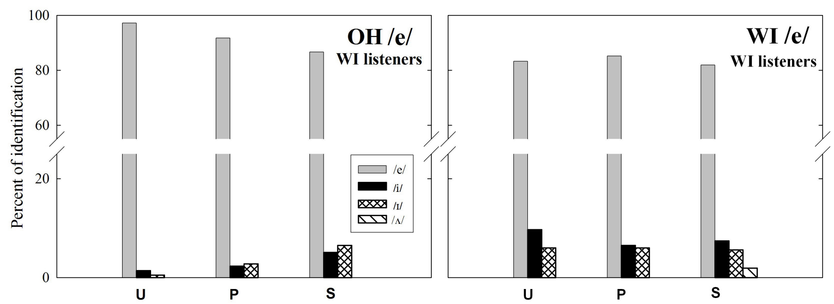

Figure 7: Identification of /e/ across three prosodic levels (U, P, S) for Wisconsin listeners responding to Ohio productions (left) and Wisconsin productions (right).

Comparing the substitution patterns shown in Figures 6 and 7, it is clear that vowels spoken by Wisconsin speakers were perceived more often as /i/ by both Ohio and Wisconsin listeners than vowels in Ohio productions. This was true particularly for vowels in the U-position. Substitutions by /ɪ/ were also prevalent, although less frequent than by /i/. A striking result was obtained for Ohio listeners' responses to the vowels in S-position. For both Ohio and Wisconsin speakers, the misidentified vowels were perceived most often as /ʌ/ and not as either /i/ or /ɪ/ as in the responses of the Wisconsin listeners. This outcome may be related to Ohio listeners' response to lesser formant movement for /e/ in the S-position as compared to the two more prominent positions. Such response most likely reflects dialectal differences in perceptual attunement to vocalic distinctions and needs to be addressed in future experiments.

4.2.2. *Identification of /ɛ/*

Displayed in Figure 8 is the average identification of the lax vowel /ɛ/ by Ohio listeners responding to Ohio and Wisconsin productions. The identification rates were very similar for both dialects. Consistently with prosodic hierarchy, vowels in the S-position were identified less often as intended by the speakers than vowels in higher prosodic positions. However, both U- and P-positions had comparable identification rates and there were no clear differences between them. The results of an ANOVA performed on the identification data showed that the effect of prosodic position was significant ($F(2, 31) = 12.15$, $p < 0.001$, $\eta^2 = 0.403$). Subsequent post hoc analyses indicated 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.

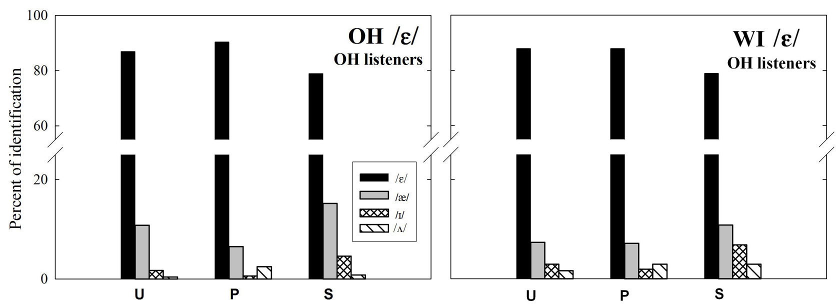

Figure 8: Identification of /ɛ/ across three prosodic levels (U, P, S) for Ohio listeners responding to Ohio productions (left) and Wisconsin productions (right).

As further shown in Figure 8, some dialectal differences were perceived. Notably, the number of substitutions by /æ/ for Ohio productions was higher than for Wisconsin productions. The substitutions for Wisconsin /e/ were more distributed across the neighboring vowels, including /æ/, /ɪ/, and /ʌ/.

Identification rates by Wisconsin listeners were slightly higher across prosodic positions and, unlike in the case of /e/, Wisconsin productions yielded a higher number of responses as intended by the speakers, particularly in the U-position. As evident in Figure 9, responses of Wisconsin listeners reflected the graded prosodic differences for Wisconsin productions but the identification pattern for Ohio vowels was similar to that for Ohio listeners. The results of an ANOVA showed that the main effect of prosodic position was significant ($F(2, 28) = 7.65$, $p = 0.002$, $\eta^2 = 0.353$), although significant differences were obtained only between the S-position and either U- or P-position, which did not differ from each other.

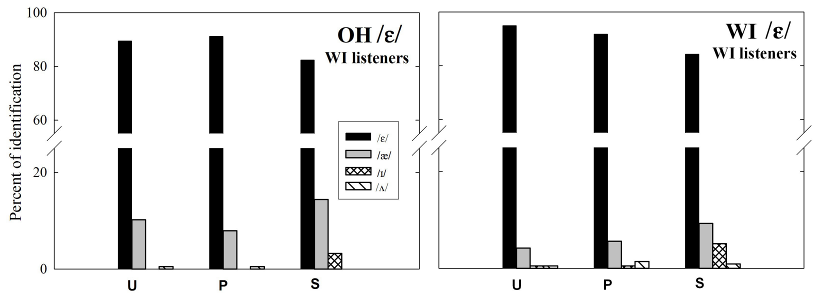

Figure 9: Identification of /ɛ/ across three prosodic levels (U, P, S) for Wisconsin listeners responding to Ohio productions (left) and Wisconsin productions (right).

As has been previously observed in the substitution pattern for Ohio listeners, the vowels spoken by Ohio speakers were confused with /æ/ more often than vowels spoken by Wisconsin speakers. For the small overall number of substitutions in classification of Wisconsin vowels, the responses were more distributed over the neighboring vowels /æ/, /ɪ/, and /ʌ/.

5. General discussion

In our exploration of possible synchronic sources of diachronic vowel changes, we examined differences in vowel characteristics as a function of variation in prosodic prominence.[4] We expected that more emphatic vowel productions would be reflected in greater vowel enhancement, which, over time, will serve as an impetus to sound change across generations of speakers from the same geographic area.

As a specific example, we selected two diachronic principles of sound change, the raising of tense vowels and possible lowering of lax vowels in chain shifts. Investigating changes to vowel characteristics as a function of differences in prosodic prominence, we found noteworthy synchronic parallels to the diachronic changes. Comparing the results from two closely related dialects of Midwestern American English, we aimed to find whether and to what extent the general trends are influenced by dialect-specific differences.

Measured acoustically, we found Ohio vowels /e/ and /ɛ/ to be longer and more fronted relative to Wisconsin vowels. We predicted that both duration differences and differences in the acoustic location may affect how the vowels are perceived by listeners born and raised in the same geographic areas as the speakers and may also affect cross-dialect vowel perception.

Measuring formant frequency changes at multiple locations equidistant in the course of the vowel's duration, we were able to detect gradual fronting and raising of the tense vowel /e/ with each position of higher prosodic prominence. This is because acoustic correlates of vowel enhancement are manifested in more extreme frequency values and, in the case of a diphthongal vowel, a greater degree of diphthongization. As could be expected, longer durations of Ohio vowels caused a greater amount of frequency change as compared to Wisconsin vowels.

An apparent contradiction with respect to raising and fronting of the vowel /e/ emerged from the data. Namely, since sound change involves both speaker and listener, the more enhanced instances of /e/, although raised and fronted in the acoustic space, should actually result in a better identification of the vowel as /e/ with each higher prosodic position. This would suggest a change in the acoustic location and no perceived vowel category change under greater amount of enhancement. Would greater prosodic prominence create a clearer exemplar of a vowel or would it cause a vowel category shift?

Perception results showed that either outcome is possible. For Ohio vowels, greater enhancement resulted in higher identification rates and lesser confusions with neighboring vowels for both Ohio and Wisconsin listeners. Moreover, longer, more fronted and more diphthongal instances of Ohio vowels were favored by Wisconsin listeners over the vowels of their own dialect, yielding higher identification rates. For Wisconsin vowels, however, more enhanced vowel variants did not provide clearer exemplars for the listeners. To the contrary, both Wisconsin and Ohio listeners perceived them more often as /i/ or /ɪ/. This suggests that Wisconsin /e/ is more likely to undergo a perceptual category shift under greater enhancement, sounding to speakers of other, non-shifting dialects more raised and fronted. This is in line with diachronic Principle I and suggests that the principles of vowel changes formulated in the past have a synchronic basis.

The question arises as to why the tense vowel /e/ tends to rise in Wisconsin English and not in Ohio. Although we can only speculate at present, the raising of /e/ may be related to the other changes undergoing in Wisconsin vowel system. While it is known that Northern Cities Shift affects short vowels only, the acoustic characteristics of long vowels have not been explored sufficiently and perception of long vowels has not been tested to rule out possible changes to long vowels in the dialect of south-central Wisconsin. Consequently, since English spoken in central Ohio is unaffected by any chain shift, we would not necessarily expect the raising of /e/ in Ohio.

Turning our attention to the second vowel investigated in this study, the lax /ɛ/, we found smaller formant frequency changes as a function of prosodic context for both Ohio and Wisconsin productions. In both dialects, the vowel is a true monophthong and we thus do not expect it to become diphthongized in more emphatic productions. Variation in prosodic prominence affected primarily vowel duration although small changes to formant frequency in more prominent positions were also observed. As a consequence of these frequency changes, the vowel was slightly lower and more fronted in more emphatic productions of both Ohio and Wisconsin speakers.

Before we discuss our perception results for /ɛ/, we need to address the issue of intelligibility scores for this vowel in other reported studies. In a study conducted in Michigan (Hillenbrand et al. 2001), listeners identified eight vowels, including /ɛ/, from a word list read by the speakers of the same Northern Cities dialect. The vowels were embedded in various consonant contexts. The intelligibility was lowest for the vowels /æ/ and /ɛ/ and the average identification rate for these vowels was about 90%. The most common misidentifications of tokens that were intended as /ɛ/ were /ɪ/ or /æ/. Comparing our results with Hillenbrand et al. (2001) data, the identification rates for /ɛ/ in the lowest prominence position were much lower, reaching 79% for Ohio listeners and 83% for Wisconsin listeners. This underscores that variation as a function of prosodic context may have great effects on vowel intelligibility and clearly supersedes variation typically found across consonant contexts.

It is important to note that the instances of vowel /ɛ/ spoken by Ohio speakers were misidentified most often as /æ/ by both Ohio and Wisconsin listeners. However, Wisconsin misidentifications of /ɛ/ were more distributed among neighboring vowel categories /æ/, /ɪ/, and /ʌ/ in the responses of both Ohio and Wisconsin listeners. This indicates that listeners were able to detect acoustic differences between Ohio and Wisconsin instances of /ɛ/. More fronted Ohio vowels yielded more confusion with /æ/ and more centralized Wisconsin vowels were more often confused with other neighboring vowels.

Diachronic Principle II, that short vowels fall in chain shifts, finds less clear synchronic support in light of the present results, although our data support the observed pattern of /ɛ/ backing currently ongoing in a number of American dialects, and therewith our revision of Principle II. Although the more emphatic productions of /ɛ/ did cause lowering of the vowel in the acoustic space, the perceptual results of this lowering in terms of a larger number of substitutions by /æ/ were somewhat greater for Ohio vowels which are assumed to be unaffected by any chain shift. Both the centralized acoustic location of Wisconsin /ɛ/ and substitutions by neighboring vowels in the perception of /ɛ/, particularly in the least prominent S-position, would suggest a gradual backing of this vowel in Wisconsin, which is consistent with changes in vowel characteristics in the Northern Cities dialects. More experimental work is needed to examine more thoroughly diachronic Principle II which, as already mentioned at the outset, seems to be ill-attested, though our results are more consistent with the backing tendency of /ɛ/ found in the Northern Cities Shift.

6. Conclusions

This study explored acoustic changes to vowels as a function of variation in prosodic prominence and examined listeners' responses to these changes with the aim of understanding how sources of systematic variation in speech can lead to dialectal vowel change over time. Generally, the results shed light on the question of how vowels change in particular directions over time. More emphatically produced vowels correlate in our data with greater vowel enhancement, which serves as a primary impetus to dialect-specific perceptual judgments about vowel quality. In the case of tense diphthongal /e/, dynamic rather than static vowel characteristics determine the nature and direction of the change. They also contribute to a lack of such change in a dialect which is not undergoing chain shifting. A more complex relationship between acoustic variation in vowel enhancement, its perception, and dialectal differences emerged for the lax monophthongal vowel /ɛ/, although the data are rather suggestive as to the direction of the change in the dialect undergoing Northern Cities Shift.

Overall, then, our results suggest a likely connection between vowel enhancement and vowel changes of the sort posited for chain shifts. We are now beginning work to pursue such connections at the seam between generations, namely in the speech of grandparents, parents, and children.

Acknowledgements

We thank Kristin Hatcher, Jennifer Delahanty and Dilara Tepeli for help with collection of the perception data. Special thanks go to Jeff Murray for editorial help. Work supported by NIH NIDCD R01 DC006871-01.

Notes

1. An anonymous reviewer reminds us that other changes do not behave in accordance with Principle II, such as a lax vowel raising pattern in New Zealand (Watson et al. 1998, Watson et al. 2000).
2. There are of course many and varied phonological approaches to prosodic prominence within sentences, the most important reviewed in Selkirk (1995). Our approach here, using three levels from utterance-initial to unstressed forms (but with non-reduced vowels) and an intermediate one to produce graded degrees of emphasis, is intended to avoid those theoretical controversies for the moment to the extent possible.
3. In fact, these stimuli are not entirely structurally parallel. In particular, as one reviewer also notes, *shark bait* forms a phonological word while *risky bets* does not. We have since revised the stimuli extensively and begun testing them. We hope to report the results soon.
4. There is a long tradition in phonetics and historical linguistics of connecting sound change to patterns of hypo- and hyper-articulation, and similar terms. In the introduction, above, and in related work (Jacewicz et al. 2006), we have suggested a reason for the apparent correlation between prosodically-driven 'exaggeration' and vocalic chain shifting, rooted in language learning.

Appendix

Table 1: ANOVA results for the tense vowel /e/ showing significant main effects and interactions for F1 and F2 measurements.

Effect	dF	dFe	F	p	η^2
OHIO					
Prosodic level_F1	1.3	8.0	6.80	0.025	0.531
Measurement location_F1	1.6	9.7	21.11	<0.001	0.779
Prosodic level_F1 x measurement location_F1	2.7	16.3	5.86	0.008	0.494
Prosodic level_F2	1.4	8.5	28.40	<0.001	0.826
Measurement location_F2	1.3	7.5	43.60	<0.001	0.879
Prosodic level_F2 x measurement location_F2	2.1	12.9	6.460	0.010	0.519
WISCONSIN					
Prosodic level_F1	1.6	9.4	13.86	0.002	0.698
Measurement location_F1	1.7	10.1	18.05	0.001	0.751
Prosodic level_F1 x measurement location_F1	ns				
Prosodic level_F2	1.7	10.3	55.35	<0.001	0.902
Measurement location_F2	2.4	14.6	72.18	<0.001	0.923
Prosodic level_F2 x measurement location_F2	2.4	14.3	4.95	0.019	0.452

Table 2: ANOVA results for the lax vowel /ɛ/ showing significant main effects and interactions for F1 and F2 measurements.

Effect	dF	dFe	F	p	η^2
OHIO					
Prosodic level_F1	1.9	11.4	5.91	0.016	0.496
Measurement location_F1	2.2	13.0	8.64	0.004	0.590
Prosodic level_F1 x measurement location_F1	ns				
Prosodic level_F2	1.7	10.4	9.78	0.003	0.620
Measurement location_F2	ns				
Prosodic level_F2 x measurement location_F2	ns				
WISCONSIN					
Prosodic level_F1	1.6	9.4	5.04	0.038	0.457
Measurement location_F1	2.1	12.7	44.88	<0.001	0.882
Prosodic level_F1 x measurement location_F1	ns				
Prosodic level_F2	1.8	10.6	9.20	0.006	0.605
Measurement location_F2	ns				
Prosodic level_F2 x measurement location_F2	2.2	12.4	6.47	0.010	0.519

References

Cho, Taehong and Patricia A. Keating
2001 Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29: 155-190.

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: 259-271.

Fox, Robert Allen, Ewa Jacewicz, and Joseph Salmons
2006 Prosodically induced phonetic variations in vowels: A source of language change? In *The Bill Question*, Howard I. Aronson, Donald L. Dyer, Victor A. Friedman, Daniela S. Hristova, and Jerrold M. Sadock (eds.), 87-110. Bloomington, Ind.: Slavica.

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: Duke University Press.

Gordon, Matthew
2002 Investigating chain shifts and mergers. In *The Handbook of Language Variation and Change*, J.K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 244-266. Maldon, Mass. and Oxford: Blackwell.

Hillenbrand, James M., Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler
1995 Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97: 3099-3111.

Hillenbrand, James M., Michael J. Clark, and Terrance M. Neary
2001 Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America* 109: 748-763.

Jacewicz, Ewa, Joseph Salmons, and Robert Allen Fox
2006 Prosodic prominence effects on vowels in chain shifts. *Language Variation and Change* 18: 285-316.

Jacewicz, Ewa, Robert Allen Fox, and Joseph Salmons
2007 Vowel duration in three American English dialects. *American Speech* 82: 367-385.

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

Labov, William
2001 *Principles of Linguistic Change, 2: Social Factors.* Oxford: Blackwell.

Labov, William and Sherry Ash
1997 Understanding Birmingham. In *Language Variety in the South Revisited*, Cynthia Bernstein, Thomas Nunnally, and Robin Sabino (eds.), 508-573. Tuscaloosa: University of Alabama Press.

Labov, William, Sherry Ash, and Charles Boberg
2006 *The Atlas of North American English: Phonetics, Phonology and Sound Change.* Berlin and New York: Mouton de Gruyter.

Labov, William, Malcah Yeager, and Richard Steiner
1972 *A Quantitative Study of Sound Changes in Progress*. Philadelphia: U.S. Regional Survey.

Milenkovic, Paul
2003 TF32 software program. University of Wisconsin, Madison.

Onaka, Akiko
2003 Domain-initial strengthening in Japanese: and acoustic and articulatory study. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Maria-Josep Solé, Daniel Recasens, and Joaquín Romero (eds.), 2091-2094. Barcelona, Spain.

Selkirk, Elizabeth
1995 Sentence prosody: Intonation, stress, and phrasing. In *The Handbook of Phonological Theory*, John A. Goldsmith (ed.), 550-569. Oxford: Blackwell.

Sievers, Eduard
1876 *Grundzüge der Lautphysiologie: Zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. 2nd ed. *Grundzüge der Phonetik*. Leipzig: Breitkopf and Härtel. Second edition 1881.

Stockwell, Robert P.
1978 Perseverance in the English vowel shift. In *Recent Developments in Historical Phonology*, Jacek Fisiak (ed.), 337-348. The Hague: Mouton.

Watson, Catherine I., Jonathan Harrington, and Zoe Evans
1998 An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics* 18: 185-207.

Watson, Catherine I., Margaret Maclagan, and Jonathan Harrington
2000 Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change* 12: 51-68.