Language Variation and Change, **18** (2006), 285–316. Printed in the U.S.A.
© 2006 Cambridge University Press 0954-3945/06 $9.50
DOI: 10.1017/S0954394506060145

Prosodic prominence effects on vowels in chain shifts

EWA JACEWICZ
ROBERT ALLEN FOX
The Ohio State University

JOSEPH SALMONS
University of Wisconsin–Madison

ABSTRACT

This study examines synchronic variation in vowels in an effort to advance our understanding of the "transmission problem" in language change, in particular, the cross-generational perseverance of vowel shifts. Seeking a connection to patterns and directions of shifts in vowel systems over time, we examine the role of a largely neglected parameter of structured heterogeneity: prosodic prominence. Experimental data from two Midwestern dialects of American English—central Ohio and south-central Wisconsin—show that, for the vowels studied here, the changes in vowel characteristics observed under higher degrees of prosodic prominence (or greater emphasis) correspond to the changes predicted by well-established principles of chain shifting. An acoustic study assesses variation in prosodic prominence by examining formant frequencies at multiple locations in the course of vowel duration, which provides information about vowel quality dynamics. A perceptual study determines listeners' sensitivity to the obtained acoustic variation, as manifested in specific patterns of vowel identification, confusions, and category goodness ratings. Finally, a prosodically based explanation of the transmission of sound change is described, which offers new connections between structural and social factors in sound change, notably the roles of "social affect" and speaker gender.

This article addresses a core puzzle about chain shifts, the type of sound change most studied in variation and change research: Why and how do similar chain shifts recur across languages and continue over generations, even centuries? The goal of the experiments presented here is to suggest that persistent structural factors, particularly prosodically conditioned variation in vowel characteristics, can contribute to "moving" a vowel in the chain in the direction of the dialect-specific sound change.

This article builds on earlier conference papers (Fox, Jacewicz, & Salmons, 2003; Jacewicz & Salmons, 2003; Jacewicz, Salmons, & Fox, 2004a, 2004b; Salmons & Jacewicz, 2004). Additional acoustic analyses and perceptual results from the data set presented here can be found in Jacewicz et al. (forthcoming) and Fox et al. (forthcoming). We thank three anonymous reviewers for helpful comments on a previous version of the manuscript. We also thank Kristin Hatcher, Jennifer Mercer, and Dilara Tepeli for help with collection of the perception data. Work supported by NIH NIDCD R01 DC006871-01 and NIH NIDCD R03 DC 005560 to The Ohio State University (Ewa Jacewicz, PI).

Chain shifting is a feature of the grammar of Germanic languages, a constant and dynamic fact at any given time in their history. The most famous example, the English Great Vowel Shift, was a massive sound change affecting the long vowels of English during the 15th to 18th centuries. Arguing against the traditional view that it was a single, unique event in the history of English, Stockwell (1978:337) wrote:

> The vowel shift occurred no more at the usually cited dates than at any other date in the documented history of English. That is, it *did* occur then, and also (equally, I believe) over the past 200 years, or over the 200 years between the birth of Alfred and the death of Aelfric, or any other period of that length. This kind of vowel shifting is a pervasive and persevering characteristic of vowel systems of a certain type.

If chain shifts persevere across generations, the question arises as to how is it possible that successive generations systematically over- or under-shoot phonetic targets in the same directions as their elders, steadily advancing the shift. In a search for an explanation, this article turns to structural factors. A closer examination of production and perception of selected vowels in Midwestern American English is undertaken to observe the size of the variation in vowel characteristics as a function of prosodic prominence of a vowel in an utterance. In particular, the study explores potential consequences of variation in prosodic prominence in transmission of vowels undergoing a set of changes known as the Northern Cities Shift (NCS) as opposed to vowels not involved in this shift. The experiments are designed to examine the role of more emphatic productions in vowel transmission and do not aim to explain how dialect-specific changes are initiated.

To that end, the article first introduces the notion of chain shifting and the phonetics of contemporary chain shifts in progress in American English. It then discusses the premise that prosodic prominence, a structural parameter of variation, plays an important role in transmitting such changes. The effects of variation in prosodic prominence on vowel characteristics are examined experimentally in a cross-dialect acoustic study, and listeners' responses to acoustic vowel changes are subsequently tested in a cross-dialect perception study. The results of these two studies are discussed in relation to broader issues in chain shifting.

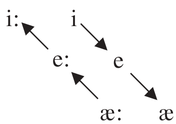
CHAIN SHIFTING AND THE DIVERSIFICATION OF AMERICAN ENGLISH

"Chain shifting" is the propensity of vowels to shift not in isolation from one another but in interlocking groups.[1] This has been recognized since Eduard Sievers (1881:202, original edition, 1876:130–131), who wrote about the direction of chain shifts,

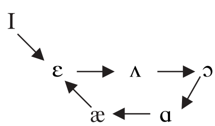
> that the articulations of a sound are more energetically and securely realized the more strongly one is conscious of it, i.e., the greater its intensity or its quantity. For a long vowel, this explains both the intensification of a specific articulation of the tongue... For a short vowel, by contrast, which requires only a momentary tongue

> movement, the actual amount of distance from the neutral position is easily not reached, i.e., a change of [short] vowels with stronger specific articulations into sounds with more neutral articulation is set in motion (which applies to tongue as well as to lip activity).[2]

This general trend is also captured in Labov's Principles I and II, where long/tense vowels raise and short/lax vowels lower, as illustrated below with three front long and short vowels (Sievers, 1876; Labov, 1994; and others):



Sievers pursued an essentially articulatory account of chain shifts, connected to duration, while researchers in the Labovian variationist tradition have focused on the role of the whole sound system or social motivations, among other factors. Placing these diachronic principles in a contemporary context, work on American English since Labov, Yaeger, and Steiner (1972), canonized in Labov (1994) and Labov et al. (2006), presented evidence that urban areas from Buffalo to Madison are undergoing a chain shift, including raising of the vowel in words like *bad* and *bag*, which can sound to speakers from other areas more like *bed* and *beg*. The changes postulated, the "Northern Cities Shift," (NCS) are illustrated below (after Labov, 1994):



bid > bed > bud > bawd > bod > bad > bed/bid

Labov (1994:122, and elsewhere) gave 15 examples of long vowels rising in chains (Principle I), but Principle II is, by comparison, poorly attested. He gave only two examples of short-vowel lowering, both of which also crucially involve backing, namely North Frisian /i/ to /a/ (not simply to /ɛ/ or to /æ/) and Vegliote /y/ to /ɔ/ (rather than to a front-rounded vowel) and /ɛ/ to /a/ (not to /æ/). Short vowels thus appear prone to backing, as well as lowering, as in fact also illustrated by backing of mid vowels in the NCS, a third example of Principle II. We then interpret Principle II as including lowering and/or backing.

THE PHONETICS OF CONTEMPORARY CHAIN SHIFTS IN PROGRESS

Let us turn now to how shifts in progress have been analyzed. Modern sociolinguistics is built around the notion of the orderly heterogeneity of language, as argued in the landmark work of Weinreich, Labov, and Herzog (1968:101): "The key to a rational conception of language change—indeed, of language itself—is the possibility of describing orderly differentiation in a language serving a community." Today, historical linguists agree with sociolinguists that language use in any community involves such "structured heterogeneity" and change draws heavily on that pool of synchronic variation. The sociolinguistic tradition, *by definition* focused on social factors of variation, has also obviously involved consideration of purely structural variables, including phonetic ones. The variation in prosodic prominence explored in this article is one such factor and a system-internal component of structured heterogeneity, which may provide a link to social and acquisitional factors in the propagation of shifts.

In studying vocalic chain shifts, the confounding effects of certain consonants (usually in the coda) on preceding vowels, have been noted and laid out most clearly by Labov, Yaeger, and Steiner (1972, I:52–91) for the conditioning of /æ/ raising.[3] Since that seminal work, contextual factors have been inconsistently taken into consideration or unreported in ways that would be assumed without discussion in contemporary phonetics or speech science research. For example, reported formant values in sociolinguistics are often given without information about adjacent consonants. Yet, immediate consonantal context may greatly contribute to the raising or fronting effects observed in the acoustic space. For example, Hillenbrand, Clark, and Nearey (2001) showed that /æ/ may be raised about 50 Hz in alveolar and velar consonant environments, which cause a downward shift in the first formant (F1) even when the vowel is measured at the "steady state." Even bigger changes to formant values at "steady state" have been reported for /u/, whose second formant (F2) may increase by 500–600 Hz when the vowel is embedded in alveolar consonants. Such increase in F2 values corresponds to considerable fronting of the vowel in the acoustic space, which disappears in the "neutral" /h_d/ environment. Sharpening our understanding of vocalic changes requires more systematic control of variation coming from consonant context, which may have a significant impact on the results.

Assessing the vocalic changes in sociolinguistics and dialectology involves plotting acoustic locations of vowels within the traditional F1/F2 grid.[4] That remains "the primary focus of instrumental vowel analysis" (Thomas, 2002:172), though some attention has been given in recent work to vowel duration. In this research, formants have typically been measured at one point for monophthongs and, more recently, at two points for diphthongs (e.g., Thomas, 2001:12). However, a more detailed analysis of acoustic vowel changes may be necessary to uncover information that may be missed without paying attention to their dynamic characteristics. Ash (2003:62–64) has recently called specifically for research into vowel dynamics in American dialects.[5]

While segmental conditioning has been well studied since the 19th century, attention in phonology, phonetics, and historical linguistics has turned in recent decades toward processes "above the segment," that is, at higher levels of prosodic structure. Yet none of these variables—the effects of consonantal context on the acoustic structure of vowels, the changes in formant values over the duration of a vowel, or the role of higher-level prosody in shaping the low-level realization of vowels—has been brought to bear on chain shifts, nor have they been generally exploited in dialectology and sociolinguistics. The most neglected of these is prosody, a topic to which we now turn.

A NEGLECTED STRUCTURAL PARAMETER OF VARIATION: PROSODY

Modern research on prosody from phonological, phonetic, and diachronic perspectives has created a foundation for introducing prosody as another dimension of structured heterogeneity. In this article, we argue that variability is prosodically structured, so that more and less emphatically pronounced vowels vary in predictable ways. That is, the articulatory and acoustic values of a given vowel are not static or absolute, but are rather situated in and profoundly shaped by prosodic structure. As a result, vowels in more prominent positions in an utterance not only reach their fullest articulation, but in a sense are hyperarticulated, while less prominent vowels are not merely reduced, but usually change their relative positions in the acoustic space. This variation is systematically present for all speakers and listeners at all times, and is a crucial factor for vocalic changes.

In sociolinguistics, the presence of prosodically conditioned variation has been noted by Labov (1994:195), who observed in passing that "the most highly stressed vowels tend to move farther in the direction of the change in progress" (see also Labov, 2001:12, as well as discussions of "style" and shift by Labov and others). In other passages, Labov's discussion is suggestive of additional connections to prosody. In his treatment of the speech of Carol Meyers (the pseudonym for a speaker recorded in a variety of settings by Labov's team), he paints a clear picture of "social affect" correlating with more advanced shift, in particular in the case of "complaints"—in contrast to her performance in "businesslike" situations (2001:439–442).

Labov's eloquent description of the social surroundings and speech styles can, we argue, be tied to a purely structural fact about the sounds of English:[6] The range of realizations of a particular vowel in a given speaker's dialect, from unshifted to advanced shift, is not fixed across different linguistic contexts. Rather, the degree of shift may be conditioned by prosody in that more emphatic tokens of a given vowel are more shifted, and less emphatic ones, less shifted. The correlation of "social affect" to vowel shifting in Carol Meyers' speech is secondary in our view. The more direct connection is between vowel change and prosodic prominence, which often goes hand in hand with more emotional speech. In short, we see emotional emphasis as one of the correlates of hyperarticulation (Lindblom, 1990), which causes changes to acoustic vowel characteristics similar to what we find in more prosodically prominent positions. As a result of these acoustic changes, vowels spoken more emphatically stand out from the continuously varying speech stream and provide listeners with more perceptual cues than vowels uttered less clearly. As such, vowels in prosodically prominent positions thus not only represent instances of more advanced shift, but should be more salient to listeners and learners. No study has tested this relationship systematically or rigorously, examining both the changes to acoustic vowel characteristics in speaker productions and listener response to such prosodically conditioned productions.

In our own work, to which we now turn, we define prominence in more relative terms. We do not take it to be exclusively a feature of the hierarchically ordered prosodic levels where each level consists of one or more constituents (e.g., Beckman, 1986; Beckman & Pierrehumbert, 1986). Although we agree that heads of such prosodic constituents are systematically hyperarticulated (de Jong, 1995, 2004), this is also consistent with insights from work on contrastive emphasis (Erickson, 1998, 2002; Erickson, Fujimura, & Pardo, 1998), clear speech (e.g., Ferguson & Kewley-Port, 2002; Krause & Braida, 2002, 2004; Picheny et al., 1986) and target undershoot (Lindblom, 1990; Moon & Lindblom, 1994). Common to all these lines of research is the observation that vowels uttered with greater enhancement differ systematically from vowels spoken more casually. In our view, the complexity of speech in everyday speaking situations almost certainly represents, in accord with the sociolinguistic position, a set of continua along several dimensions. A transition from one speaking style to another is dictated by a speaker's intention to emphasize or deemphasize specific content, which necessitates selection of most appropriate forms on the continuum from "very clear" to "very casual." Beyond individual differences between speakers, variation in speaking styles is subject to pragmatic intent and higher-level prosodic organization. The combination of characteristics of clear and casual speaking styles within an utterance is, by necessity, prosodically conditioned because any sequence of words serving human communication is shaped by prosody to convey a wealth of information including both lexical and nonlexical information, discourse functions such as focus and prominence, and/or emotions.

To arrive at experimentally testable degrees of prominence, we adopted in the present article a three-level prominence distinction, in the spirit of an early work by Trager and Smith (1951) who distinguished four different prominence categories representing relative strengths of vowels in utterances. In our work on vowels, we examine characteristics of three different levels of prominence: stressed and accented, stressed, and full vowels. It has been well established that stressed vowels have three main acoustic correlates that differentiate them from unstressed vowels: longer duration, greater intensity, and pitch (F0) movement (e.g., Fry, 1955; Harris, 1978; Rump & Collier, 1996; Sluijter & Van Heuven, 1996; Terken, 1991). Our present focus is on spectral changes, that is, formant frequency movement in the course of vowel duration. In this study, we examine changes to formant frequencies as a function of differences in prosodic prominence and hypothesize that such prosodically induced variation may contribute to dialect-specific vowel shifts.

CROSS-DIALECT PRODUCTION STUDY

We have collected data in two regions of the United States that appear to be undergoing distinct patterns of vowel shifting, central Ohio and south-central Wisconsin. Central Ohio is part of the Midlands dialect area (Labov et al., 2006: ch. 11). It is not thought to be participating in the classic chain shifts like NCS or the Southern Shift, but there are ongoing changes nonetheless. Based on our own interpretation of data presented in Thomas (2001:82–102) and other sources, it appears to us that /ɛ/ is moving downward in central Ohio, while /e/ appears to be approaching a more peripheral position in the vowel space.[7] South-central Wisconsin is widely assumed to be in the early stages of the Northern Cities Shift, meaning that /ɛ/ is moving back, centralizing. Little is known at the moment about the movement of /e/ in Wisconsin, although we are about to begin work on this, using real-time data collected over the last half century. For the moment, we expect prosodic effects to be moving it outward and upward, in accord with the core principle of chain shifting.

We focus here on the results of two studies that examined cross-dialect production and perception of prominence effects on two vowels, /e/ and /ɛ/, by speakers and listeners from Wisconsin and Ohio. Additional acoustic and perception results of this study can be found in Jacewicz et al. (forthcoming) and Fox et al. (forthcoming). The subset of the results presented here centers on aspects pertaining directly to the focus of the present article.

METHODS

Stimulus materials

The general approach for obtaining prosodically conditioned variable productions was similar to the one used in a number of articulatory studies initiated by Fougeron and Keating (1997). We introduced several modifications to the methods to elicit more natural pronunciation across three selected prosodic contexts from linguistically untrained speakers. The stimuli consisted of twelve sentences, six of which contained the words *bait* and *bet* in three positions of prosodic prominence: the highest being *Utterance* (U), the intermediate *Phonological Phrase* (P), and the lowest *Syllable* (S). In these prosodic contexts, the vowels in U-position are stressed and accented. In P-position, defined as the strong branch of a noun phrase, the vowels are normally stressed. In S-position, defined as the weak branch of a morphological compound, the vowels are unstressed but unreduced.

Utterance-initial position: stressed and accented vowel
1. '*Bait* shop' is what I said.
2. '*Bet* some' is what I said.

Phrase position: stressed vowel
3. She said [the [*bait* shop]] was closed.
4. He said [the [*bet* slips]] were here.

Syllable position: full vowel
5. [Shark *bait*] with flavor seems hard to find.
6. [Risky *bets*] are nice but safe bets are better.

Although we have constructed the sentences to obtain graded differences in vowel prominence, one cannot assume that a given written text will yield a particular prosodic pattern (see on this point Shattuck-Hufnagel & Turk, 1996:228). To date, we have encountered only a few surprising occurrences: Some pauses within a sentence were unexpected, and phrases like *bait shop* were pronounced with separate stress on each word (*báit shóp*) by some speakers, rather than only on the first (*báit shop*). It has been suggested to us that the single quotes around phrases in sentences 1–2 may alter reading intonation, but we have not seen evidence of this. It is more likely that an effect could arise from different syntactic constructions, such as compound noun phrase versus verb phrase, as in *bait shop* versus *bet some*. In addition to the test sentences, eight sentences of similar structure containing the words *date* and *debt* and eight additional distractor sentences were included in the set. Each sentence occurred three times and the presentation order was randomized for each speaker.

Speakers

Eight Ohio speakers (four women, four men) and eight Wisconsin speakers (four women, four men) participated in the study. The Ohio speakers were born and raised in the area of Columbus, and the Wisconsin speakers in the area of Madison. 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.

Procedure

The sentences appeared on a monitor screen, one at a time, and were read at a moderate speaking rate. The recording was under control of a computer program in Matlab. The sentences were directly digitized onto the hard disk at 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 distance of one inch from the lips. The 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. A total of 288 vowel tokens in /b_t/ context were obtained for subsequent analyses (6 sentence types $\times$ 16 speakers $\times$ 3 repetitions).

Vowel measurements

Vowel duration and formant movement in the course of 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 hand corrected (and then rechecked) using another 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 consonant–vowel (CV) or vowel–consonant (VC) formant transitions and estimate vowel inherent spectral change in the form of formant movement. The stimulus tokens were down-sampled to 11.025 kHz and premphasized (98%) prior to spectral analysis. Formant frequencies were estimated with a Matlab program that utilized a 14-order linear predictive coding (LPC) analysis with a 15 ms Hanning window centered over each measurement location.

RESULTS

The data were examined for statistical significance using analysis of variance (ANOVA). In all ANOVA results in this article, for all reported within-subjects significant main effects and interactions, the degrees of freedom for the F-tests were Greenhouse-Geisser adjusted when there were significant violations of sphericity. In addition to the significance values, a measure of the effect size, partial eta squared (η^2), is also reported (η^2, whose value can range from 0.0 to 1.0, should be considered a measure of the proportion of variance explained by a dependent variable when controlling for other factors). *Post hoc* analyses, when reported—both for vowel duration and formants 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 General Linear Methods (GLM) contrasts (for within-subject factors). All statistical analyses were performed in the SPSS v.13 software package (SPSS, Inc., 1997).

Vowel duration

Shown in Figure 1 are mean vowel duration data for /e/ and /ɛ/ displayed by prosodic context, speaker gender, and dialect. These durations were analyzed using ANOVA with the within-subject factor prosodic level (U, P, S) and the between-subject factors speaker gender (male, female) and speaker dialect (Ohio, Wisconsin). The main effect of prosodic level was significant for /e/, $F(1.7, 20.8) = 13.64, p < .001, \eta^2 = .532$, and for /ɛ/, $F(1.5, 18.2) = 9.4, p = .003, \eta^2 = .440$. *Post hoc* contrasts showed that vowels in the most prominent Utterance position were significantly longer than vowels in the Phrase and Syllable positions. Female Wisconsin productions were the only exception to this pattern, showing no significant difference due to prominence position. This unusual pattern was reflected in the significant interaction between dialect, gender, and prosodic level for /e/, $F(1.7, 20.8) = 3.85, p < .043, \eta^2 = .243$.

Also of interest is the general trend for female vowels to be longer than male vowels across the prosodic positions. These differences were not significant for /e/ although the overall mean duration values for female vowels were higher

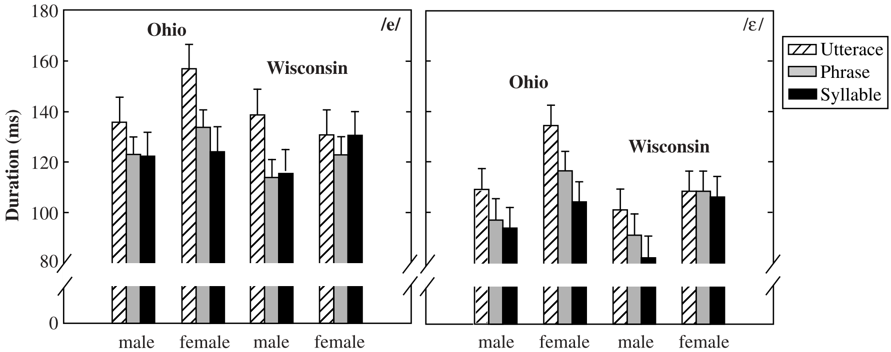


FIGURE 1. Mean vowel durations across three positions of prosodic prominence (Utterance, Phonological Phrase, and Syllable) in the vowels /e/ in *bait* (left) and /E/ in *bet* (right) for Ohio and Wisconsin male and female speakers. Error bars indicate one standard error.

than for male vowels in both Ohio (139 vs. 127 ms) and Wisconsin (129 vs. 123 ms). However, the effect of gender was significant for /ɛ/, $F(1, 12) = 6.9$, $p = .022$, $\eta^2 = .364$, showing that women's vowels were longer in all prosodic positions for both Ohio and Wisconsin. For this vowel, not only were the overall means for female vowels higher than for male vowels in Ohio (118 vs. 100 ms) and Wisconsin (108 vs. 91 ms) but the differences were consistent across all prosodic positions, as illustrated in Figure 1.

Finally, there were noteworthy dialectal differences in vowel durations. The effect of speaker dialect did not reach significance for either vowel, although the Ohio speakers produced, on average, longer vowels than did the Wisconsin speakers (133 vs. 126 ms for /e/ and 109 vs. 99 ms for /ɛ/).

Formant movement

In very general terms, spectral changes in the course of duration reflect the dynamic character of a vowel. The amount of formant frequency change varies with how diphthongal the vowel is, and true diphthongs such as /aɪ/ will show much greater formant movement than relative monophthongs such as /æ/. In the case of /e/, we expect greater formant movement than for /ɛ/ due to its clear diphthongized character in both Ohio and Wisconsin. We predict that the amount of formant frequency change will further vary with the degree of prosodic prominence of the vowel, the greatest being in the highest Utterance position, and will diminish with the decreasing prosodic prominence in Phrase and Syllable positions, respectively. For the vowel /ɛ/ we still predict similar graded differences, although the amount of frequency change is expected to be much less than for the vowel /e/.

In this article, we report the formant pattern for female productions only. As shown in Jacewicz et al. (forthcoming), women produced more exaggerated forms

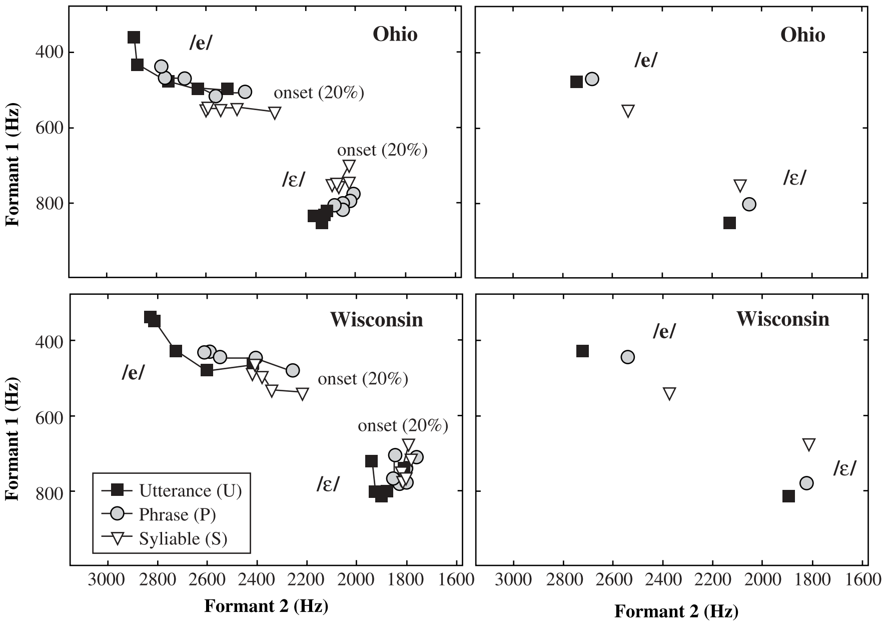


FIGURE 2. Overall mean F1 and F2 frequencies for /e/ in *bait* and /E/ in *bet* at five locations in the vowel corresponding to 20%, 35%, 50%, 65%, and 80% of vowel duration (left panels) and at a single measurement point at vowel midpoint (right panels) across three prosodic prominence positions. The 20% measurement location is marked as "onset" and the lines connect each consecutive measurement location for each prosodic context. The mean frequency values are plotted for four female speakers from Ohio and four female speakers from Wisconsin.

of vowels in terms of formant frequency values than men, and we selected female data for the present article for a better display of the variable effect of prosodic context. Figure 2 (left panels) shows the frequency values for the first two formants, F1 and F2, sampled at five points in time in a vowel. The data for /e/ and /ɛ/ are plotted separately for Ohio and Wisconsin productions. Repeated-measures ANOVAs with the within-subject factors prosodic context (U, P, S) and measurement point in the course of vowel duration (corresponding to 20%, 35%, 50%, 65%, and 80% point) and the between-subject factor speaker dialect were conducted for each formant (F1 and F2) of each vowel /e/ and /ɛ/.

F1 movement for /e/

For F1 in /e/, the main effect of prosodic context was highly significant, $F(1.5, 9.2) = 29.7, p < .001, \eta^2 = .832$. The effect of measurement point was also significant. $F(1.7, 11.1) = 11.5$, $p = .002$, $\eta^2 = .658$, as was the interaction between prosodic context and measurement point, $F(2.8, 16.9) = 4.5$, $p = .019$,

$\eta^2 = .427$. These results indicate that prosodic context affected both the nature and the amount of formant frequency change in the course of vowel duration. As can be seen in Figure 2, the vowel becomes progressively raised (as measured by its relative position in the acoustic space) with each higher level of prosodic prominence. This raising tendency was also reflected in the overall frequency means for F1 (averaged across all five measurement points), which were progressively smaller with increasing prosodic prominence (S = 523 Hz, P = 461 Hz, and U = 431 Hz). The effect of dialect was not significant for F1, indicating that the dialects did not differ along the vowel height dimension.

F2 movement for /e/

For F2 in /e/, the effect of prosodic context was highly significant, $F(1.5, 9.2) = 97.8$, $p < .001$, $\eta^2 = .942$, as was the effect of measurement point, $F(1.8, 10.9) = 73.7$, $p < .001$, $\eta^2 = .925$. The interaction between prosodic context and measurement point was also significant, $F(3.4, 20.6) = 5.3$, $p = .006$, $\eta^2 = .468$. These results indicate that prosodic context affected F2 movement and the amount of frequency change so that the vowel was progressively more fronted with each higher prosodic prominence. Although the main effect of dialect was not significant, the interaction between prosodic context and speaker dialect was significant, $F(1.5, 9.2) = 4.6$, $p = .048$, $\eta^2 = .433$, showing that the Ohio vowels were more fronted than the vowels in Wisconsin productions, as also reflected in the overall F2 means across all five measurement points (Wisconsin means: S = 2355 Hz, P = 2485 Hz, U = 2679 Hz vs. Ohio means: S = 2514 Hz, P = 2655 Hz, U = 2740 Hz).

F1 and F2 for /e/ measured at vowel midpoint

Measuring formant frequency at multiple locations over the course of the vowel provides an estimate of formant movement, the dynamic vowel property that characterizes the degree of diphthongization. This measurement technique is essential in assessing formant frequency change of a diphthongized vowel as a function of prosodic prominence. We can contrast these results to those obtained when analyzing, instead, the frequencies of the formants measured at vowel midpoint (the 50% measurement point)—a position in the vowel that roughly corresponds to the vowel "steady state" point (also referred to as 'vowel center', 'vowel target', or 'vowel nucleus') which many acoustic studies use as a single measure of vowel quality.

Before we discuss our results at vowel midpoints, it needs to be pointed out that sampling formant frequency at the steady-state time has been a prevailing and traditional method stemming from speech science, whereby the obtained measurements have often served as values for synthetic versions of static vowels. Although it has been established in a great number of studies that information in the direction and slope of formant transitions and inherent spectral change contribute to vowel identity (e.g., Gay, 1978; Lindblom & Studdert-Kennedy, 1967; reviews in Hillenbrand & Nearey, 1999; Nearey, 1989; and Strange, 1989), the dynamic frequency changes at vowel transitions are typically excluded from the estimation of vowel "quality." In more recent studies (and depending on experimental focus), sampling vowels at two locations is not uncommon (e.g., in Hillenbrand et al., 1995, the measurements were taken at 20% and 80% of vowel duration). Measuring formant frequencies at two temporal points, midpoint and offset, has been evoked to capture essential dialectal information in diphthongal offsets (e.g., Thomas, 2001) and has also been applied to a larger set of vowels (Anderson, 2003). A three-point estimate at 20%, 50%, and 80% has been used in studies on clear speech (e.g., Bradlow, 2002). Sampling diphthongal trajectories at multiple locations was introduced by Fox (1983) who used four equidistant temporal points to estimate formant movement. Another multiple-location technique for measuring diphthongal trajectories by utilizing a frequency constant, ΔF, was developed by Jacewicz et al. (2003).

Our five-point measurement technique presented in this article captures the variation in spectral changes across vowels of variable durations. The use of relative measurement points in the course of vowel duration, rather than fixed values, of say, 25 ms from the onset/offset, has the advantage of estimating spectral changes resulting from sources other than the effects of immediate consonant context of the vowel. That is, variation as a function of prosodic context is well preserved for each consonant environment of the vowel, whose effects maintain across prosodic contexts and throughout the duration of the vowel.

Returning to the present results, Figure 2 (right panels) displays the frequency values for the first two formants, F1 and F2, measured at vowel midpoint. The data are plotted separately for Ohio and Wisconsin productions. Results of ANOVAs with the within-subject factors prosodic context (U, P, S) and the between-subject factor speaker dialect conducted on midpoints of F1 and F2 revealed a significant main effect of prosodic context both for F1, $F(1.9, 11.5) = 7.5, p = .009, \eta^2 = .554$, and for F2, $F(1.7, 10.4) = 40.6, p < .001, \eta^2 = .871$. This indicates that the effects of prosodic prominence were also detectable at vowel midpoints and were thus consistent with the general direction of vowel raising and fronting. However, this analysis did not show any dialect differences for F2 that were significant only when formant frequencies were measured at multiple locations in a vowel. Clearly, the single measurement at vowel midpoint not only fails to provide information about the diphthongal character of the vowel, but in the present data, does not demonstrate the fact that Ohio vowels were more fronted relative to the Wisconsin vowels. For example, the F2 means for U-levels at vowel midpoint were 2728 Hz for the Wisconsin /e/ and 2758 Hz for the Ohio /e/, as compared to 2679 Hz for Wisconsin and 2740 Hz for Ohio when F2 means were averaged across all five measurement points in the course of the duration.

F1 frequency change for /ɛ/

For the vowel /ɛ/, when frequency changes to F1 were measured at five locations, the effects of prosodic context and measurement point were significant, $F(1.9, 11.7) = 6.97, p = .010, \eta^2 = .538$ and $F(1.9, 11.3) = 15.4, p = .001, \eta^2 =$

.720, respectively. Although the nature of spectral change is different for /ɛ/ than for /e/, and the amount of formant frequency change is much smaller (compare Figure 2, left panels), these results indicate that prosodic context affects formant frequency of a relatively monophthongal vowel as well. The frequency of F1 increased progressively with each higher prominence position, which corresponded to a progressive lowering of /ɛ/ in the acoustic space. A significant interaction between measurement point and dialect, $F(1.9, 11.3) = 4.1$, $p = .048$, $\eta^2 = .406$, suggests that there might be some minor differences in the nature of formant movement between Ohio and Wisconsin female productions, although the effect does not seem to be large. The main effect of dialect was significant, $F(1, 6) = 19.7$, $p = .004$, $\eta^2 = .766$, indicating that the Ohio /ɛ/ was lower in the acoustic space than the Wisconsin /ɛ/ (the overall F1 means for Ohio were 788 Hz and for Wisconsin 749 Hz).

F2 frequency change for /ɛ/

For F2 in /ɛ/, the main effects of prosodic context and measurement point were significant, $F(1.7, 9.9) = 26.4$, $p < .001$, $\eta^2 = .815$ and $F(2, 11.8) = 6.6$, $p = .012$, $\eta^2 = .525$, respectively, as was the interaction between these two factors, $F(8, 48) = 2.6$, $p = .018$, $\eta^2 = .304$. This shows that prosodic context affected F2 frequency values so that the vowel was more fronted with increasing prosodic prominence, which also affected frequency change in the course of vowel duration. The significant main effect of dialect, $F(1, 6) = 14.97$, $p = .008$, $\eta^2 = .714$, revealed important differences between the Ohio and Wisconsin productions. Across all prosodic contexts, the Ohio vowels were more fronted than were the Wisconsin vowels (overall F2 means were 2082 Hz for Ohio and 1839 Hz for Wisconsin). This significant difference points to the fact that the Wisconsin vowel /ɛ/, being more centralized, might be undergoing a change in the direction of the Northern Cities Shift.

F1 and F2 for /ɛ/ measured at vowel midpoint

Results of ANOVAs with the within-subject factors prosodic context (U, P, S) and the between-subject factor speaker dialect conducted on midpoints of F1 and F2 revealed a significant main effect of prosodic context both for F1, $F(1.9, 11.3) = 5.7$, $p = .020$, $\eta^2 = .489$, and for F2, $F(1.4, 8.3) = 7.5$, $p = .019$, $\eta^2 = .555$. The main effect of dialect was significant for F2, $F(1, 6) = 12.4$, $p = .013$, $\eta^2 = .673$, but it was not significant for F1. Comparing these results with the results of analyses of formant movement, there is an agreement as to the effects of prosodic context on midpoints of F1 and F2: The vowels were progressively lowered with each higher prosodic position and were more fronted in the highest U-position. However, although both types of analyses also showed dialect differences for F2, the dialect differences for F1 were not significant at vowel midpoint, contrary to what was observed when the vowel was measured at multiple locations. When measured at vowel midpoint, the F1 means were 800 Hz for Ohio and 785 Hz for Wisconsin, thus the difference between them was much smaller than when the overall means came from averaged values across all five measurement points, as reported earlier. Based on the information at vowel midpoint, we can only state that the Ohio /ɛ/ is more fronted *but not lower* than the Wisconsin /ɛ/.

Summary of the production results

The results show that vowel characteristics change as a function of prosodically conditioned variation in predictable ways. Productions of vowels in the most prominent prosodic positions require more articulatory effort than those in less prominent positions in an utterance. This accounts for the graded differences in acoustic vowel characteristics across prosodic contexts, such as differences in duration and the degree of diphthongization (or frequency change). As a result of this general constraint on vowel production, vowels in the most prominent positions are typically longer and exhibit more frequency change in the course of their durations.

We have observed that the diphthongized vowel /e/ was rising and became more fronted with each stronger prominence position for both Ohio and Wisconsin productions. This effect is not surprising in light of findings that stressed (or emphatically produced) vowels reach more extreme frequency values. Vowels in these extreme locations have been shown to exhibit less undershoot of F2 (Moon & Lindblom, 1994) than in less clear speaking styles. As proposed by Lindblom (1990), such hyperarticulated productions may imply more articulatory effort by the speaker and/or may make greater demands on the speech production mechanisms.

Notable differences between the two dialects appeared in the realizations of the short vowel /ɛ/, which was lower and more fronted in Ohio as compared to Wisconsin. The centralized position of the Wisconsin /ɛ/ in the acoustic space gives us an indication that this difference might be related to its participation in the Northern Cities Shift. Although prosodically conditioned variation arises from the same sources, the prosodic effects on the Ohio /ɛ/ are shown to be not the same as on the Wisconsin /ɛ/ due to their different positions in the acoustic space. Thus, although the Wisconsin vowel is also lower and more fronted in the most prominent prosodic context, its position in the acoustic space suggests more perceptual confusions with neighboring short vowels and not mostly with /æ/ as it may be expected for Ohio /ɛ/. These confusions will most likely occur when the listener is faced with an instance of a vowel in a less prominent prosodic context, with shorter duration and less extreme formant frequency values. If this prediction were indeed supported by the perception data, prosodically conditioned variation may provide one explanation of the sources of the Northern Cities Shift.

Crucially, in our view, vowels in most prominent positions exhibit characteristics related to "social affect" in everyday speaking situations. These exaggerated productions have a potential to attract the attention of the listener (e.g., the young language learner), becoming a norm-like "standard" by younger generations of speakers from the same area. In turn, each younger generation, producing more enhanced vowel variants from the already exaggerated ones, prolongs a vowel shift typical for the system in a given dialect area. The emphatic productions are typically more common in women's speech, as confirmed in our present production data, and women thus are most likely to lead the vowel shift. However, the expressiveness of "social affect" is not limited to women's speech, but can also come from individuals playing particular social roles, as with Eckert's "drama queens" (Eckert, 2004).

On the technical side, we assessed in this study formant frequency changes by sampling frequencies at multiple temporal locations in a vowel and not only at vowel midpoint. This measurement technique emphasized the size of the effect of formant movement across prosodic prominence positions. It also revealed interactions between prosodic positions, measurement points, and dialect, which were otherwise impossible to detect at a single measurement at vowel midpoint. Thus, dialect differences could well exist but remain unnoticed as a result of the less robust methodology traditionally used.

CROSS-DIALECT PERCEPTION STUDY

The cross-dialect production study showed that acoustic vowel characteristics vary as a function of prosodic prominence. The question addressed in the perception study is whether listeners are sensitive to this type of variation in making vowel quality decisions. Our claim that prosodic prominence may play an important role in chain shifts will not be supported if this type of variation does not influence listeners' judgments about vowel identity. If there are acoustic differences in vowel characteristics that are perceptually salient, they should produce differences in perceived vowel quality reflected in the choice of phonetic quality and/or in vowel goodness ratings.

METHODS

Stimulus material

The stimuli consisted of all instances of the words *bait* and *bet* recorded in the production study by eight Ohio speakers (four men and four women) and eight Wisconsin speakers (four men and four women). The words were edited out of sentences for a total of 288 tokens (6 sentences × 16 speakers × 3 repetitions). The stimuli were randomized and played to listeners in 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 either dialect.

Listeners

Listeners with no history of hearing and speech disorders were selected from the same regional varieties of American English as the speakers in the production study. There were twenty listeners born and raised in the Columbus, Ohio area (10 men and 10 women, aged 20–46 years) and fifteen listeners (6 men and 9 women, aged 19–36 years) born and raised in the Madison, Wisconsin area. No listener was also a speaker in the production study.

Procedure

Sound was delivered over high-quality headphones (Sennheiser HD600) at a comfortable listening level (70 dB SPL) to a listener seated in a sound-attenuating booth. The implementation of the experiment was computer-controlled using a program written in Matlab. Listeners were required to give two types of responses. First, they identified the vowel in each stimulus token in a forced-choice task, with the choices being *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. Listeners were instructed to click the mouse on the response box corresponding to the vowel heard. Second, after choosing one of six responses, the listener then rated the goodness of the identified 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. To familiarize themselves with the procedure, listeners were given practice (using 30 randomly selected tokens) prior to the identification test at the beginning of the experimental session. No feedback was provided.

RESULTS

Identification and goodness ratings for /e/

As shown in Figure 3, listeners were sensitive to graded differences in acoustic vowel characteristics across prosodic prominence positions as observed in the production data. Generally, the instances of /e/ produced in positions of higher prominence were classified as /e/ more often than the vowels in the lowest prominence position.

Important dialectal differences were also perceived. Vowels spoken by Ohio speakers obtained higher identification rates by *both* Ohio and Wisconsin listeners. Furthermore, misidentification of /e/ increased with each lower prominence position for Ohio speakers, whereas for Wisconsin speakers, it was high in both the highest and lowest prominence positions. Clearly, the most emphatic productions yielded two distinct identification results: Ohio vowels were perceived almost exclusively as instances of /e/ and Wisconsin vowels were misidentified most often as /i/, which was comparable with the pattern obtained in the lowest prominence position. On average, vowels spoken by women were more often identified as the intended vowels (a “correct” response) than were those spoken by men.

To determine the significance of this variation on vowel identification, the identification data used a repeated-measures ANOVA with the within-subject

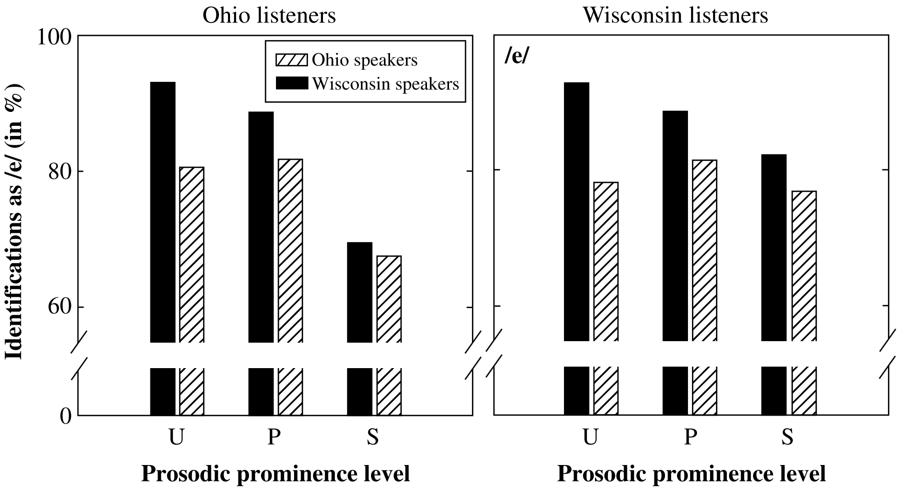


FIGURE 3. Identification responses to /e/ in *bait* across three positions of prosodic prominence (U, P, and S) for Ohio and Wisconsin listeners.

factors being prosodic context, speaker dialect, and speaker gender. The analyses were done separately for Ohio and Wisconsin listeners.[8] For Ohio listeners, all three main factors were significant. A significant prosodic context effect, $F(1, 24) = 12.53$, $p = .001$, $\eta^2 = .410$, resulted from the fact that significantly fewer vowels produced in the S-position were identified as /e/ than in either the U- or P-positions. However, the significant interaction between prosodic context and speaker dialect, $F(2, 31) = 4.72$, $p = .020$, $\eta^2 = .208$, demonstrated that the effect of prosodic context was somewhat different for Ohio compared 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. 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, $F(1, 18) = 13.79$, $p = .002$, $\eta^2 = .434$, resulting from the fact that the Ohio vowels were more often identified as /e/ than the Wisconsin vowels. Speaker gender was also a significant main effect, $F(1, 18) = 22.94$, $p < .001$, $\eta^2 = .560$. In particular, there were more /e/ responses to the *bait* tokens when they were produced by women than by men. The significant interaction between speaker gender and prosodic context, $F(2, 32) = 6.64$, $p = .005$, $\eta^2 = .208$, 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.

For Wisconsin listeners, two main effects were significant. A significant effect of speaker dialect, $F(1, 14) = 14.764$, $p = .002$, $\eta^2 = .513$, resulted from the fact that the Ohio vowels were more often identified as /e/ than were the Wisconsin

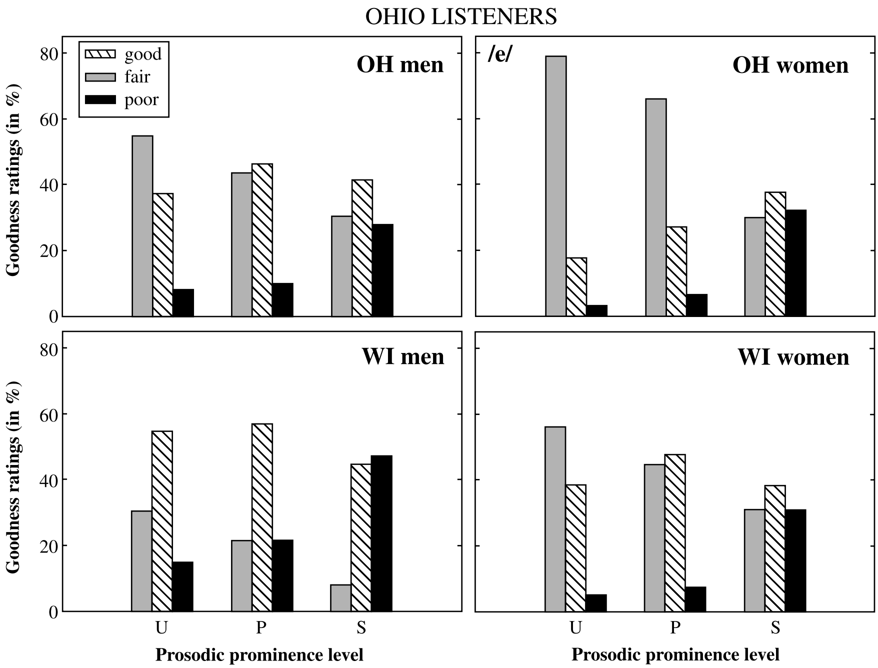


FIGURE 4. Goodness ratings for /e/ by *Ohio listeners* responding to Ohio and Wisconsin productions of *bait* broken down by speaker gender.

vowels. Speaker gender was also a significant main effect, $F(1, 14) = 16.790$, $p = .001$, $\eta^2 = .545$; vowels spoken by women were more often identified as being what speakers intended than vowels spoken by men. A significant interaction between speaker dialect and speaker gender, $F(1, 14) = 5.762$, $p = .031$, $\eta^2 = .292$, showed that Wisconsin men obtained significantly lower "correct" responses than the three remaining groups of speakers.

A main effect of prosodic context was not a significant factor for Wisconsin listeners. However, a significant interaction between speaker gender and prosodic context, $F(2, 28) = 4.087$, $p = .028$, $\eta^2 = .226$, revealed that some differences due to the prosodic context were perceived. In particular, female vowels in the U- and P-positions obtained significantly higher identification rates than vowels in all other productions.

The distribution of goodness ratings for the vowels identified as /e/ indicated that differences in prosodic prominence were also perceived in a more subtle way. As shown in Figure 4 for Ohio listeners, vowels in 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. Conversely, the percentage of time the vowel was classified as a "poor" exemplar of /e/

increased in P- and S-positions, respectively. This pattern was consistent across Ohio and Wisconsin productions, although the percentages of "good" exemplars were generally higher for Ohio productions, and considerably more so for women.

A separate analysis was done on the goodness ratings. 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 with the within-subject factors being prosodic context, speaker gender, and speaker dialect. Again, all three main effects were significant for Ohio listeners. There was a significant main effect of prosodic context, $F(1, 24) = 43.67$, $p < .001$, $\eta^2 = .708$, 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 they were significantly higher 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, $F(1, 18) = 41.93$, $p < .001$, $\eta^2 = .700$, and significantly higher ratings for women's vowels than for men's vowels, $F(1, 18) = 27.21$, $p < .001$, $\eta^2 = .602$.

Figure 5 displays goodness ratings for Wisconsin listeners. A similar response pattern was obtained although both Ohio and Wisconsin female productions reached comparable levels of classifications as 'good'. The number of 'good' responses decreased drastically for Wisconsin men, showing that Wisconsin listeners responded more favorably to vowels spoken by either Ohio speakers or by Wisconsin women than by Wisconsin men.

For Wisconsin listeners, all three main effects were also significant. A significant main effect of prosodic context, $F(2, 28) = 22.67$, $p < .001$, $\eta^2 = .618$, was obtained as the mean goodness ratings decreased progressively with each lower prominence position. The effect of speaker dialect was significant and showed that vowels in Ohio productions were rated more highly, $F(1, 14) = 5.46$, $p = .035$, $\eta^2 = .281$. Wisconsin listeners also had significantly higher ratings for women's vowels than for men's vowels, $F(1, 14) = 22.95$, $p < .001$, $\eta^2 = .621$. There was a significant interaction between speaker dialect and speaker gender, $F(1, 14) = 27.47$, $p < .001$, $\eta^2 = .662$, which came from the fact that Wisconsin women's vowels were rated most highly and Wisconsin men's productions obtained the lowest goodness ratings.

Identification and goodness ratings for /ɛ/

Identification of the lax vowel /ɛ/ revealed somewhat different patterns than those found for /e/. The effect of prosodic prominence was evident for the lowest S-position, where the vowel was identified less often as /ɛ/ than in the U- and P-positions. However, no differences between the two latter positions were man-

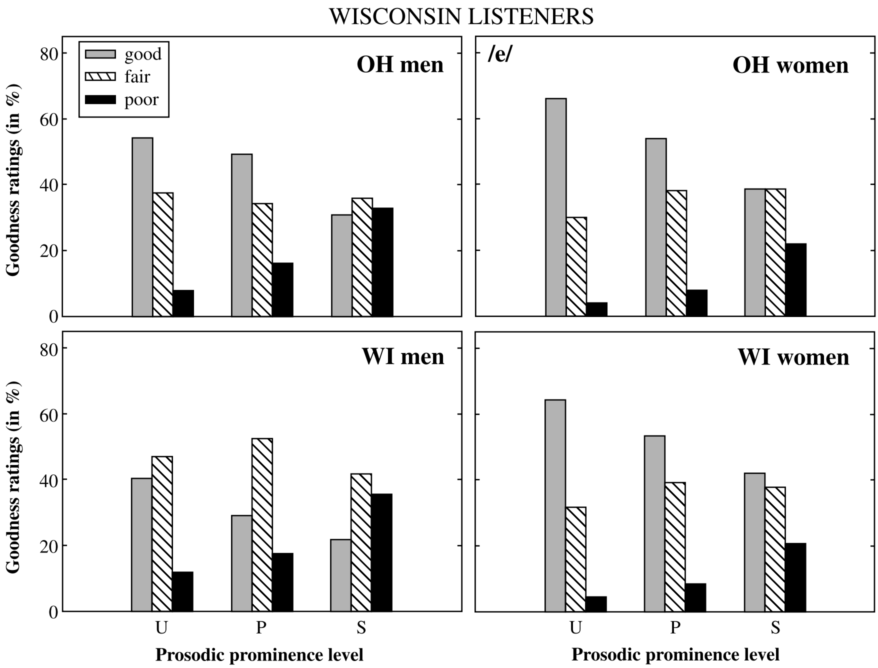


FIGURE 5. Goodness ratings for /e/ by *Wisconsin listeners* responding to Ohio and Wisconsin productions of *bait* broken down by speaker gender.

ifested. Also, as Figure 6 shows, no clear dialectal differences in identification rates across all three prosodic positions were found.

The results of an ANOVA performed on the identification data from Ohio listeners showed that only the main effect of prosodic context reached statistical significance, $F(2, 31) = 12.15$, $p < .001$, $\eta^2 = .403$, indicating that vowels in the S-position were identified significantly less often as /ɛ/ than vowels in either the U- or P-position, which did not differ from each other. A similar effect was obtained for Wisconsin listeners. A significant prosodic context effect was the only main effect reaching significance, $F(2, 28) = 7.65$, $p = .002$, $\eta^2 = .353$, again due to the fact that vowels in the S-position were identified significantly less often as /ɛ/ than vowels in either the U- or P-position, which did not differ from each other. There was also a significant interaction between speaker gender and prosodic context for Wisconsin listeners, $F(2, 28) = 7.784$, $p = .002$, $\eta^2 = .357$, coming from the fact that men's vowels in the S-position were identified significantly less often.

A consistent pattern of vowel misidentifications, however, indicates that dialectal differences were nevertheless perceived. The vowels spoken by Ohio speakers were overwhelmingly confused with /æ/ by both Ohio and Wisconsin listeners, whereas Wisconsin vowels were confused more often with /ɪ/, /æ/, and /ʌ/, again by both groups of listeners.

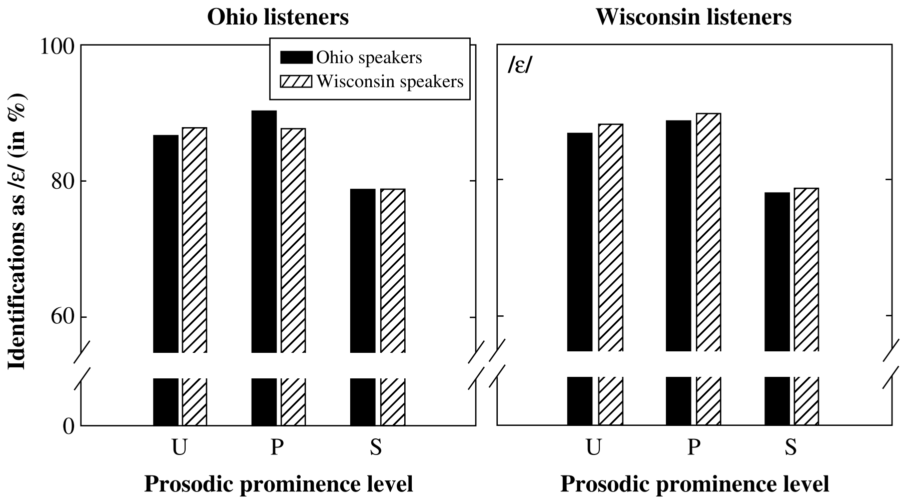


FIGURE 6. Identification responses to /ɛ/ in *bet* across three positions of prosodic prominence (U, P, and S) for Ohio and Wisconsin listeners.

The distribution of goodness ratings for the vowels identified as /ɛ/ provides further insights into the within-category variations. Shown in Figure 7 are goodness ratings for Ohio listeners responding to Ohio and Wisconsin productions of /ɛ/. Ohio vowels were perceived more often as "good" exemplars of /ɛ/ in both the U- and P-positions than were Wisconsin vowels. These ratings were particularly low for Wisconsin men productions. Women's productions of /ɛ/ were generally rated higher (more often as 'good') relative to the vowels spoken by men.

The results of an 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 a significant effect of prosodic context, $F(1, 26) = 86$, $p < .001$, $\eta^2 = 0.827$, and *post hoc* analysis indicated that vowels in the S-position received significantly lower goodness ratings than vowels in either the U- or P-positions (which did not differ from each other significantly). The significant effect of speaker dialect, $F(1, 18) = 9.64$, $p = .006$, $\eta^2 = .349$, showed that Ohio vowels received higher goodness ratings than did Wisconsin vowels. A significant interaction between prosodic context and dialect, $F(2, 30) = 9.03$, $p = .001$, $\eta^2 = .334$, revealed that differences in goodness ratings between the S-position and either the U- or P-position were greater for Ohio vowels than for Wisconsin vowels.

There was also a significant main effect of speaker gender, $F(1, 18) = 33.93$, $p < .001$, $\eta^2 = .653$, 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, $F(2, 31) = 20.8$, $p < .001$, $\eta^2 = .536$, showed that the differences between the S-position and either the U- or P-position were greater for women than for men.

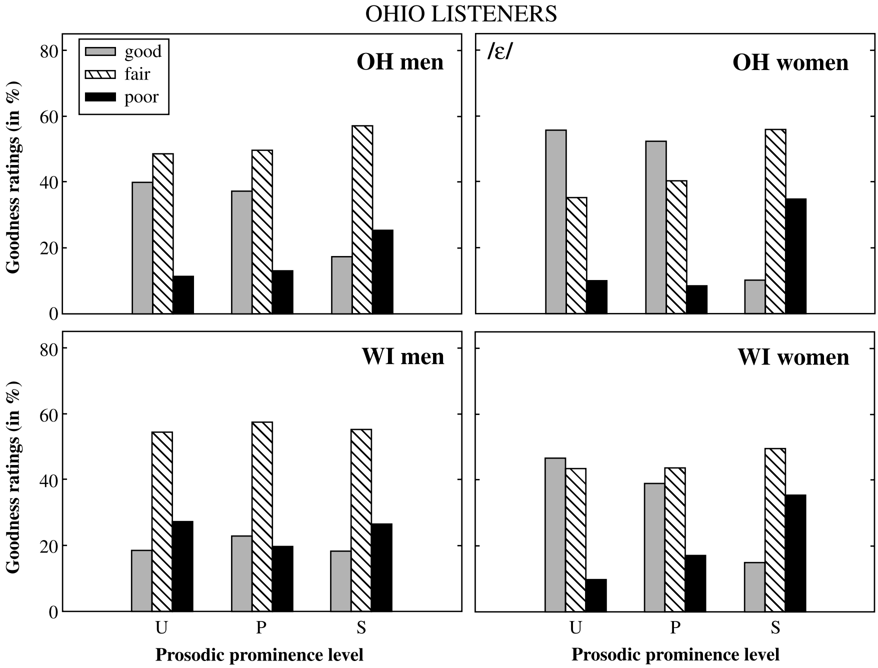


FIGURE 7. Goodness ratings for /ɛ/ by *Ohio listeners* responding to Ohio and Wisconsin productions of *bait* broken down by speaker gender.

As shown in Figure 8, Wisconsin listeners also gave higher goodness ratings to vowels spoken by women, particularly in the U- and P-positions. Perceived dialectal difference was evident for ratings given to male productions. Across all three prosodic contexts, vowels spoken by Wisconsin men, again, received the lowest category ratings of ‘good’.

For Wisconsin listeners, two main effects were significant. A main effect of prosodic context, $F(2, 28) = 45.96$, $p < .001$, $\eta^2 = .766$, showed that vowels in the U- and P-positions were significantly different from vowels in the S-position, but not from one another. A main effect of speaker dialect, $F(1, 14) = 9.17$, $p = .009$, $\eta^2 = .396$, revealed that Ohio vowels were rated more highly than Wisconsin vowels. The effect of speaker gender was not significant, but there were two significant interactions involving gender. A significant interaction between speaker dialect and speaker gender, $F(1, 14) = 7.63$, $p = .015$, $\eta^2 = .353$, came from the fact that Ohio men’s productions were rated most highly and Wisconsin men’s vowels were rated least favorably. The significant interaction between speaker gender and prosodic context, $F(2, 28) = 13.78$, $p < .001$, $\eta^2 = .496$, arose from the fact that women’s vowels in the U- and P-positions were rated most highly compared to the women’s vowels in the S-position.

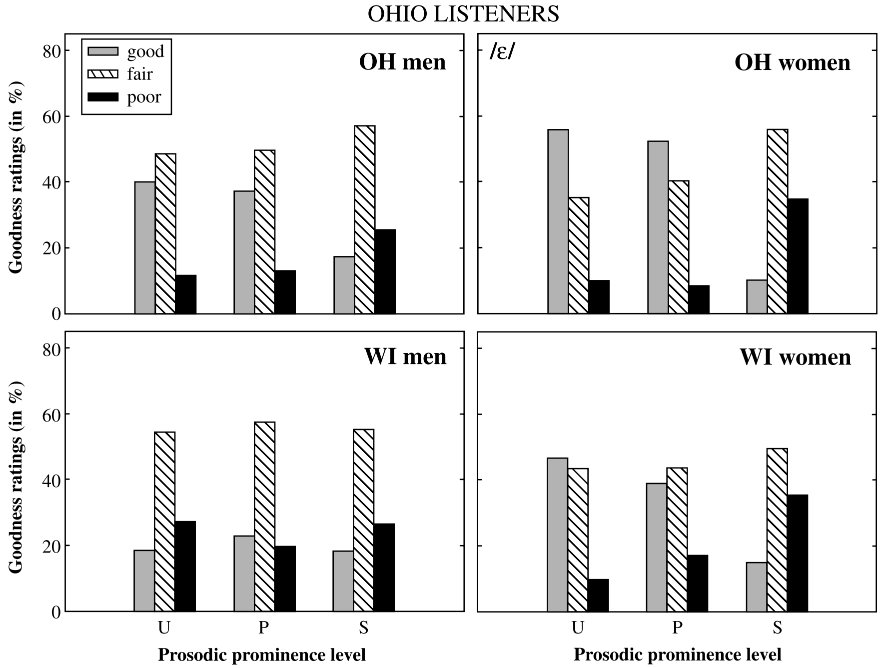


FIGURE 8. Goodness ratings for /ɛ/ by *Wisconsin listeners* responding to Ohio and Wisconsin productions of *bait* broken down by speaker gender.

Summary of the perception results

The perception results showed that listeners were sensitive to vowel characteristics as a function of variation in prosodic prominence observed in the production study. Consistent with our expectations, identification rates for the tense vowel /e/ increased with each stronger prosodic position for Ohio productions. However, the highest prominence position for Wisconsin productions yielded lower identification rates and more confusion with /i/. In searching for an explanation of this pattern, we need to examine the cross-dialect acoustic differences between the two vowels. Although both Ohio and Wisconsin /e/ were raising and became more fronted with each higher prosodic prominence position, there were cross-dialect differences in location of the vowels in the acoustic space and in vowel duration. Ohio /e/, being more fronted, more diphthongal and longer than Wisconsin /e/, produced clearer instances of the category /e/, which were less likely to be confused with any other vowel, particularly in the most prominent positions in a sentence. Compared to Ohio /e/, the Wisconsin variants of /e/ were perceived as being more monophthongal and more 'i-like', as they exhibited less frequency change in the course of their shorter durations. This was true not only for vowels in the lowest prominence position, but also extended to the positions of higher prominence.

Category goodness ratings for the vowels 'correctly' identified as /e/ further showed that listeners responded more favorably to vowels in higher prominence positions, assigning them higher numbers of 'good' category exemplars. Ohio vowels, representing clearer instances of /e/ as already discussed, yielded higher percentages of 'good' responses as compared to Wisconsin vowels. Finally, vowels produced by women, being longer and more diphthongal, obtained higher goodness ratings than vowels spoken by men, regardless of speaker or listener dialect.

The lax monophthongal vowel /ɛ/ exhibits much less frequency change in the course of its duration, even in the most prominent positions, as compared to the diphthongized vowel /e/. We would therefore not expect listeners to rely on information in spectral change in making classification decisions about /ɛ/ to the same extent as in the case of /e/. Rather, we would predict that listeners use multiple cues, which will result in variable confusions with neighboring vowels /æ/, /ɪ/, or /ʌ/. Our present results show that instances of /ɛ/ in the lowest prominence positions yielded an increased number of confusions, and the pattern of confusions was different for Ohio and Wisconsin productions. Ohio vowels were misidentified almost exclusively as /æ/, which is consistent with their more fronted and lower position in the acoustic space relative to Wisconsin /ɛ/. Misidentifications of Wisconsin vowels, however, were more distributed over neighboring categories, being perceived, in descending order, as /ɪ/, /æ/, and /ʌ/. This is again consistent with a more centralized position in the acoustic space of Wisconsin /ɛ/ relative to the Ohio variant. No clear differences in identification rates were found between the U- and P-positions, but the substitution patterns were again consistent with those for the S-position.

Considering category goodness ratings for /ɛ/, the general trend of responses was consistent with what we have observed for /e/. The more prominent positions yielded higher numbers of ratings as 'good' than did the longer Ohio variants and female productions. Again, listeners of either dialect responded more favorably to vowels spoken more clearly, such as those found in higher prosodic positions or when spoken by women rather than by men.

DISCUSSION OF THE RESULTS IN RELATION TO CHAIN SHIFTS

The results of the present cross-dialect study show that listeners respond differently to vowels spoken more clearly (or more emphatically), such as those in positions of increased prosodic prominence in a sentence compared to vowels spoken with lesser emphasis. Although the exact acoustic realizations of a particular vowel category in these prominent positions may vary across regions and dialects, such exaggerated productions provide listeners with targets for potential vowel changes and shifts over time.

How do speakers and listeners make use of prosodic information to carry out vowel change? In this study, we have examined only two vowels from the entire vowel system of two varieties of Midwestern American English, and our interpretation of the results is by necessity inconclusive. Based on results for the long vowel /e/ we can say at present that, acoustically, long vowels rise and become longer when spoken more emphatically. However, perception results helped us to understand that acoustic raising of the Ohio /e/ is not perceived as a raising of the vowel, but rather as a clearer form of it. That is, significant change in acoustic vowel characteristics brought about the best identification of the vowel.

A different result was obtained for Wisconsin /e/, inviting the question of why the spectral characteristics of the vowel in the highest prominence position, although comparable with those for the Ohio /e/, did not lead to the same perceptual identification. Given the lower identification rate and increased substitutions by /i/, we understand this effect to be a raising of the vowel also in the perceptual domain. That is, clearer production of the vowel under higher prominence led to a better identification of the Ohio variant, but caused listeners to hear it often as /i/ when responding to the Wisconsin vowel. We understand these two outcomes as reflecting internal-systemic relations rather than purely acoustic locations of these vowels. That is, Wisconsin vowels, participating in the Northern Cities Shift, may introduce a slightly different system of phonetic contrasts than the vowels spoken in central Ohio.

The results for the short vowel /ɛ/ provide further support for this interpretation. The distribution of category goodness ratings clearly showed that listeners of either dialect responded well to graded acoustic characteristics of the vowel as a function of differences in prosodic prominence. The dialectal differences between the vowels were most clearly manifested in the substitution pattern and not in identification rates. The obtained substitutions suggest that the Ohio vowel may 'lower' in perceptual classifications, being substituted most often by /æ/, but that it does not 'lower' to the same extent in Wisconsin, where it is also heard as /ɪ/, /æ/, and /ʌ/. Thus, acoustic lowering and perceived lowering may not go hand in hand, but may be linked to systemic differences between the Ohio and Wisconsin vowels. It is noteworthy that the highest prominence position did not produce the clearest perceptual exemplars of /ɛ/ in either dialect. Rather, it resulted in lower identification rates and more confusion with other vowels than what can be expected of vowels spoken with greater emphasis.

Based on the directional trends found in the present results, we predict that the vowel /e/ will raise over time in Wisconsin but not in Ohio, while /ɛ/ will lower in Ohio but will become more centralized in Wisconsin. These trends in the direction of vowel movements echo the general principles of chain shifting discussed at the outset. However, a purely articulatory and impressionistic account of the shifts, as undertaken by Sievers, must be supplemented not only with acoustic evidence from production, but also by evidence of perceptual effects that such changes produce. We can say that the long vowel /e/ will raise to the long vowel /i/ in Wisconsin, most likely due to the other changes taking place in the Wisconsin English, whereas the same vowel will less likely raise to /i/ in a system such as that spoken in central Ohio. Conversely, short /ɛ/ will 'fall' in Ohio English (perhaps because it is already moving downward in central Ohio), but not in Wisconsin, where the Northern Cities Shift affects its centralization. Despite the dialect-specific directions of vowel changes, it is clear that prosody affects acoustic vowel characteristics and has significant effects on how the vowels are perceived.

Finally, the effects of speaker gender on acoustic characteristics of vowels and on the perception of prominence-related changes, observed in this study, cannot be ignored. The statistically significant differences between women's and men's vowels point to the role of women as leading the vowel change by producing longer and clearer vowel variants, which also obtain higher category goodness ratings as compared to men.

SOME BROADER ISSUES

Historical linguists seek to motivate sound changes, whether in the seedbed of articulation or in the social patterns of transmission, including vocalic chain shifts. In this concluding section, we sketch some enduring problems with chain shifts and consider how synchronic prosodic context can contribute to shedding more light on the puzzles.

What motivates chain shifts? One answer is that sound change has fundamentally social motivations, as argued in the work of J. Milroy (particularly 1992). Milroy illustrated this with the case of palatalization of /k/ before front vowels, especially high front ones. Phonetically, this is a quintessentially 'natural' change, whereby coarticulatory conflicts between a velar stop and a following high front vowel often correlate with historical changes, such as affrication, cf. English *cheese* and West Frisian *tsiis* (vs. conservative German *Käse* and Dutch *kaas*, all borrowed from Latin *cāseus*), and other patterns of palatalization just within the documented history of Germanic. Yet it is hardly a universal change even in closely related languages and dialects, such as across Scandinavian or across different varieties of English. Milroy argued (1992:21) that "the proximity of the velar consonant to a front vowel may be a *necessary* condition for palatalization, but as it does not happen in every case, it is not a *sufficient* condition." Milroy gave a compelling account of the social side of this process, which we supplement with a structural motivation connected to social motivations. Similarly, Labov (1994) stressed that there is a role for both basic structural and mechanical motivations. The approach sketched earlier suggests strongly that chain-like synchronic effects are in fact differential according to prosodic prominence, and the data we have analyzed so far support this. Given the statistically significant differences found in the realization of a given vowel in the prosodic syllable position (i.e., the lowest prominence) and utterance position (i.e., the highest prominence), failing to carefully control for prosodic prominence, as well as consonantal context, dramatically reduces our ability to observe the effects of shift in progress. Of course the challenges of incorporating such variables into the analysis of naturalistic data are prohibitively complex and we therefore draw our data from a more controlled setting.

The effects of prosody provide a potential structural springboard, but we have only explored one possible direction of change and have said nothing about actuation, that is, how this potential is realized as being attested chain shifting, something for which an answer surely must be sought in sociolinguistics rather than in phonetics. Certainly, prosodic prominence can warp vowel realizations in ways other than those explored here. Prosodic prominence, cross-linguistically, will surely prove to have a broader range of effects on vowels, and this article has not touched on those, such as Labov's (1994:16) Principle III, that back vowels front. The effects of prosodic prominence merely constrain the possible directions of chain shifting, and those effects make chain shifting highly likely, but which of these paths are ultimately taken up by speakers is presumably ultimately determined by the social soil and climate in which this seed is planted. In contemporary American English, for instance, Principles II (that short vowels lower and back) and III (that back vowels front) are amply attested, but Principle I (that long vowels raise), is not. As we have shown, prosodic prominence correlates with raising effects on long vowels, but that potential is not being transformed into a shift at present, for which, we again stress, explanations must presumably be sought beyond language structure alone.[9]

We have drawn heavily on experimental methodologies used in speech science, but previously seldom employed in sociolinguistics and dialectology. For instance, rather than relying on traditional F1/F2 plots, we investigate the acoustics of vowels by examination of spectral change across prosodic variation and by taking multiple measures over the duration of the vowel. Representations of chain shifts, like those drawn at the outset, and quantitative analyses of chain shifts relying on single measurements of vowel steady states contribute to the traditional illusion that these shifts involve stepwise movement of static vowels over time. The present approach runs directly counter to that one: Prosodic prominence has a profound effect on vowel dynamics, and the dynamic formant trajectories of vowels over time are shaped by higher level phonological processes. Importantly, as we have shown, both speaker and listener are sensitive to this type of variation.

How can we establish that the individual elements of shift-like changes are directly connected? Such questions go to the core question for any theory of phonology: How and to what extent do sounds behave as parts of a system? Preliminary work on this largely unstudied parameter of variation leads us to hypothesize that these chain-like vocalic changes are indeed connected, but in a far different way than assumed in research until now. There are some temporal differences in these shifts, so that NCS is widely believed to begin with /æ/ raising and Southerners seem to have begun to front /u:/ earlier in time than /o:/. Still, from the synchronic perspective of a speaker in the midst of such a shift, we argue that these are in some sense neither necessarily "pull" nor "drag" relations, but instead, the prosodic system conditions them as *a group* under greater or lesser emphasis. That is, prosodic emphasis warps the vowel space in particular ways in dialects undergoing shift and, in that specific sense, the changes are fundamentally related: They are brought about by exactly the same factor. While we do not yet have anything to say about how or why particular dialects 'choose' or evolve different patterns of vocalic shift under prosodic prominence, our analysis predicts that the set of possible chain shifts in human language corresponds directly to the set of possible ways speakers can distort their vowel systems under prosodic prominence. If this line of thought continues to be borne out by additional results, then it recasts the question of interconnectedness, although the resolution of that issue lies beyond the scope of this article.

The core question posed at the outset was how shifts endure over generations and why they recur. The perseverance problem, the enduring nature of shifts over generations (or the appearance that parallel shifts are repeated centuries later), finds a more simple solution under this analysis: The basic prosodically driven asymmetries of vowel realizations reflect an underlying principle, one presumably present in the grammar of speakers of Germanic languages (and those of other families) for millennia, that accounts for the underlying unity of shifts, for why "vowel shifts fall into a few limited types" (Kiparsky, 2003:335). The particular instantiation of this overarching principle—that is, how the vowel space is acoustically warped under stress—varies somewhat, generation by generation and dialect by dialect, yielding the distinct but related patterns attested.

By virtue of learners' focus on prominent realizations of vowels, and thus on the distortions of the vowel space wrought by prosodic prominence, children may interpret as normal vowel realizations sounding closer to where more emphatic vowels were a generation earlier. If this is borne out in future research, the principles of vowel shifting would thus be shown to derive from more general principles of articulatory strengthening and weakening.

We argue that chain shifts in vowel systems are at least in part a direct effect of higher-level prosodic organization. In Germanic, stressed syllables show distinct phonetic strengthening processes and unstressed syllables show equally distinct phonetic weakening processes; this includes the realizations of vowels in those syllables. Indeed, the development of glides in utterance-level stressed tense mid vowels and the variability of effects of prosodic prominence on lax mid vowels correlate tightly with attested chain shifts documented by generations of scholars.

NOTES

1. The literature on chain shifts and related debates is prohibitively large and has been reviewed repeatedly in a number of important recent works. We refer readers to Labov (1994, 2001) and Gordon (2002), as well as Thomas (2001), on regional variation in American English vowels more generally.

2. That is "dass die Articulationen eines Lautes um so energischer und sicherer vollzogen werden, je stärker derselbe zum Bewusstsein kommt, d.h. je grösser seine Intensität oder seine Quantität ist. Dies erklärt beim langen Vocal sowohl eine Steigerung der specifischen Zungenarticulation. . . . Beim kurzen Vocal dagegen, der nur einen momentanen Zungenschlag erfordert, wird gar leicht das eigentliche Mass der Entfernung von der Indifferenzlage nicht erreicht, d.h. es wird eine Wandelung der Vocale mit stärkeren specifischen Articulationen zu Lauten von neutralerer Articulation angebahnt (sowohl was Zungen- als was Lippenthätigkeit betrifft)." This quote, with a different translation, can be found in Labov (1994:221).

3. There is, though, one issue from speech science that sociolinguistics has wrestled with in particular, that of normalization across speakers, especially normalizing values across genders (see Labov, 1994:55–56, and elsewhere).
4. See, though, Labov (1994:252–253), for a brief discussion of a more dynamic view of vowels.
5. By the same token, this kind of work may offer an opportunity to develop Stockwell's (1978) ideas about the role of glides in chain shifts.
6. We assume, as do many specialists in language change, that internal and external motivations for change often complement one another in cases of 'successful' language change. That is, in contrast to older discussions of internal *versus* external motivations of language change, those changes that reach completion often reflect clear structural and clear social motivations.
7. This is inferred from the vowel charts given for 23 of his speakers from central Ohio, who were born between 1892 and 1982.
8. Average intelligibility of individual speakers varied from 81.0% to 90.2% for Ohio speakers and from 64.1% to 88.9% for Wisconsin speakers when responses were averaged across all prosodic contexts. The intelligibility was higher when responses were averaged across the tokens in the U-position only and varied from 83.8% to 95.7% for Ohio speakers and from 57.1% to 92.4% for Wisconsin speakers. Since the productions of one Wisconsin male speaker obtained lower intelligibility scores (64.1% across all prosodic contexts and 57.1% across U-position) than productions of the other speakers, we conducted ANOVAs of arcsine transformed percent-correct responses for both Ohio and Wisconsin responses. The same pattern of significant main effects and interactions were obtained as reported for all speakers' productions after excluding the responses to the one Wisconsin male speaker. Comparing the results, we concluded that there was no need to exclude this speaker from the data set.

For a comparison, average intelligibility of individual speakers in the Hillenbrand et al. (2001) study ranged from 88.7% to 98.0% for responses to vowels averaged across a variety of consonant contexts. However, the consonant–vowel–consonant (CVC) syllables used there were read as isolated tokens and our testing stimuli occurred in a sentence context and further varied according to differences in prosodic prominence.
9. While chains involving long/tense vowel raising have not been reported to date, of course, our perceptual results may suggest the first indications of /e/ raising.

REFERENCES

Anderson, Bridget L. (2003). *An acoustical study of southeastern Michigan Appalachian and African American southern migrant vowel systems*. Doctoral dissertation, University of Michigan.

Ash, Sharon. (2003). A national survey of North American dialects. In D. Preston (ed.), *Needed research in American dialects*. Durham: Duke University Press. 57–74. [*Publications of the American Dialect Society*, 88.]

Beckman, Mary E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.

Beckman, Mary E., & Pierrehumbert, Janet B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook* 3:255–309.

Bradlow, Ann R. (2002). Confluent talker- and listener-related forces in clear speech production. In C. Gussenhoven & N. Warner (eds.), *Papers in laboratory phonology VII*. Berlin: Mouton de Gruyter. 241–273.

de Jong, Kenneth J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97:491–504.

______ (2004). Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics* 32:493–516.

Eckert, Penelope. (2004). Preadolescent drama queens and the acceleration of phonological change. *The 78th Linguistic Society of America Annual Meeting, Boston*.

Erickson, Donna. (1998). Effects of contrastive emphasis on jaw opening. *Phonetica* 55:147–169.

______ (2002). Articulation of extreme formant patterns of emphasized vowels. *Phonetica* 59:134–149.

Erickson, Donna, Fujimura, Osamu, & Pardo, Brian. (1998). Articulatory correlates of prosodic control: Emotion and emphasis. *Language and Speech* 41:399–417.

Ferguson, Sarah H., & Kewley-Port, Diane. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America* 112:259–271.

Fougeron, Cécile, & Keating, Patricia. (1997). Articulatory strengthening at edges of prosodic domain. *Journal of the Acoustical Society of America* 106:3728–3740.

Fox, Robert A. (1983). Perceptual structure of monophthongs and diphthongs in English. *Language and Speech* 26:21–60.

Fox, Robert A., Jacewicz, Ewa, & Salmons, Joseph. (2003). Prosodic domain-initial effects on the acoustic structure of vowels. *Journal of the Acoustical Society of America* 114(4):2396.

Fox, Robert A., Jacewicz, Ewa, & Salmons, Joseph. (forthcoming). Prosodically induced phonetic variations in vowels: A source of language change? In *Festschrift*. Chicago: University of Chicago Press.

Fry, Dennis B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27:765–768.

Gay, Thomas. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America* 63:223–230.

Gordon, Matthew. (2002). Investigating chain shifts and mergers. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 244–266.

Harris, Katherine S. (1978). Vowel duration change and its underlying physiological mechanisms. *Language and Speech* 21:354–361.

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

Hillenbrand, James M., Clark, Michael J., & Nearey, Terrance M. (2001). Effects of consonantal environment on vowel formant patterns. *Journal of the Acoustical Society of America* 109:748–763.

Hillenbrand, James M., & Nearey, Terrance M. (1999). Identification of resynthesized /hVd/ syllables: Effects of formant contour. *Journal of the Acoustical Society of America* 105:3509–3523.

Jacewicz, Ewa, Fox, Robert A., & Salmons, Joseph. (forthcoming). Prosodic conditioning, vowel dynamics and sound change. In *Variation in Phonetics and Phonology*. Berlin: Mouton de Gruyter.

Jacewicz, Ewa, Fujimura, Osamu, & Fox, Robert A. (2003). Dynamics in diphthong perception. In M. J. Sole, D. Recasens, & J. Romero (eds.), *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona, Spain*. 993–996.

Jacewicz, Ewa, & Salmons, Joseph. (2003). *The prosodic basis of vocalic chain shifts in Germanic*. Paper presented at the Germanic Linguistics Annual Conference, Buffalo, NY.

Jacewicz, Ewa, Salmons, Joseph, & Fox, Robert A. (2004a). *Prosodic domain effects and vocalic chain shifts*. Paper presented at the Laboratory Phonology Conference (LabPhon 9), University of Illinois, Chicago, IL.

Jacewicz, Ewa, Salmons, Joseph, & Fox, Robert A. (2004b). *Prosodic conditioning, vowel dynamics and sound change*. Paper presented at the Variation in Phonetics and Phonology Conference (Var-Phon 1), University of Potsdam, Germany.

Kiparsky, Paul. (2003). The phonological basis of sound change. In B. D. Joseph & R. D. Janda (eds.), *The handbook of historical linguistics*. Oxford: Blackwell. 313–342.

Krause, Jean C., & Braida, Louis D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America* 112:2165–2172.

——— (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America* 115:362–378.

Labov, William. (1994). *Principles of linguistic change. 1: Internal factors*. Oxford: Blackwell.

——— (2001). *Principles of linguistic change. 2: Social factors*. Oxford: Blackwell.

Labov, William, Ash, Sharon, & Boberg, Charles. (2006). *Atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter.

Labov, William, Yaeger, Malcah, & Steiner, Richard. (1972). *A quantitative study of sound change in progress*. Philadelphia, PA: U.S. Regional Survey.

Lindblom, Björn. (1990). Explaining phonetic variation: A sketch of the H & H theory. In W. J. Hardcastle & A. Marchal (eds.), *Speech production and speech modeling*. Dordrecht: Kluwer. 403–439.

Lindblom, Björn, & Studdert-Kennedy, Michael. (1967). On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America* 42:830–843.

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

Milroy, James. (1992). *Language variation and change*. Oxford: Blackwell.

Moon, Seung-Jae, & Lindblom, Björn. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* 96:40–55.

Nearey, Terrance M. (1989). Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America* 85:2088–2113.

Picheny, Michael A., Durlach, Nathaniel I., & Braida, Louis D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research* 29:434–446.
Rump, H. H., & Collier, R. (1996). Focus conditions and the prominence of pitch-accented syllables. Doctoral dissertation, Eindhoven University of Technology, Eindhoven, The Netherlands.
Sluijter, Agnes M. C., & Van Heuven, Vincent J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 8:185–190.
Salmons, Joseph, & Jacewicz, Ewa. (2004). *Vocalic chain shifts as an effect of synchronic prosody*. Paper presented at the Linguistic Society of America Annual Meeting. Boston, MA.
Shattuck-Hufnagel, Stefanie, & Turk, Alice. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25:193–247.
Sievers, Eduard. (1881). *Grundzüge der phonetik*. Leipzig: Breitkopf & Härtel. [1st ed., 1876.]
SPSS, Inc. (1997). SPSS Advanced Statistics 7.5. Chicago, IL.
Stockwell, Robert. (1978). Perseverance in the English Vowel Shift. In J. Fisiak (ed.), *Recent developments in historical phonology*. The Hague: Mouton. 337–348.
Strange, Winifred. (1989). Dynamic specification of coarticulated vowels spoken in sentence context. *Journal of the Acoustical Society of America* 85:2135–2153.
Terken, Jacques. (1991). Fundamental frequency and perceived prominence of accented syllables. *Journal of the Acoustical Society of America* 95:3662–3665.
Thomas, Erik. (2001). *An acoustic analysis of vowel variation in New World English*. Durham, NC: Duke University Press. [*Publications of the American Dialect Society*, 85.]
______ (2002). Instrumental phonetics. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 168–200.
Trager, G. L., & Smith, H. L. (1951). *An outline of English structure*. Washington, DC: American Council of Learned Societies.
Weinreich, Uriel, Labov, William, & Herzog, Marvin. (1968). Empirical foundations for a theory of language change. In W. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics*. Austin, TX: University of Texas Press. 95–188.