

Dialectal and generational variations in vowels in spontaneous speech

Robert Allen Fox, Ewa Jacewicz

Department of Speech and Hearing Science, The Ohio State University, Columbus, OH, USA.

fox.2@osu.edu, jacewicz.1@osu.edu

Abstract

The pronunciation patterns across dialect regions in the United States are changing. This paper examines acoustic differences between the vowels of old and middle-aged adults produced in spontaneous speech. The new developments in regional vowel systems found recently in citation-form vowels and in read speech were confirmed in the present analysis. While providing additional evidence for the existence of sound change on the basis of spontaneous speech data, this work exemplifies the important challenges facing researchers using spontaneous speech in laboratory analyses such as the loss of fine experimental control or the usefulness of statistical assessment.

Index Terms: dialect variation, sound change, spontaneous speech, vowel acoustics, spectral dynamics.

1. Introduction

Sound change has been studied by historical linguists by means of descriptive comparative analyses for more than two centuries. During the last 50 years, the field has been joined by sociolinguists who addressed the causes of sound change by examination of social factors (e.g., identity, prestige or social networks), typically through observations or data obtained in interviews and phone conversations. More recent phonetic models proposed several possibilities of why sound change occurs, including misperception [1] or variation in hypo- and hyperarticulation [2, 3] which, in turn, can create conditions for miscommunication. In providing descriptive and/or experimental evidence, a key difference between sociolinguists and phoneticians is that the former aim for naturalness in speech samples and the latter for experimental control.

In our recent work, we have constructed a large corpus of recorded speech documenting sound change (in apparent time, across generations of speakers) in three different dialect regions in the US. The data were collected in both highly controlled and naturalistic conditions, thus bridging the two research traditions. Our research focus is on how phonetic innovations (i.e., new pronunciation features) are transmitted and propagated across generations of speakers in a given speech community and how these changes affect the on-going development and differentiation of the vowel systems of regional dialects of American English (AE). The operation of major vowel changes (some currently active, some inactive) have split AE into major regional subsystems. These regional variations in vowels constitute the single most salient feature that differentiates AE dialects. Our research examines the current state of phonetic differences between old and young speakers in order to increase our knowledge of relative importance of acoustic and perceptual cues to the changing pronunciation patterns across dialect regions in the US.

In this paper, we examine cross-generational changes in the acoustic characteristics of six lax vowels involved in the Northern Cities Shift (NCS) [4]: /ɪ, ε, æ, a, ɔ, ʌ/ (contained in words such as *kiss*, *get*, *cat*, *pot*, *thought*, *cut*, respectively). NCS is a positional vowel rotation which affects pronunciation of the English spoken in the Great Lakes region, primarily in

the northern cities like Chicago, Detroit, Cleveland and Buffalo, and the region is termed Inland North. The NCS is initiated by the raising and fronting of /æ/. A series of follow-up positional changes includes lowering of /ɔ/, fronting of /a/, lowering and backing of /ε/, backing of /ʌ/ and some centralization and backing of /ɪ/. The order of the stages may vary across speakers but the general pattern is viewed as a clockwise rotation. Reportedly, the shift has spread farther to the North reaching southeastern parts of Wisconsin [4] although only some of its aspects have been found in this area more recently [5]. A different pronunciation of these six vowels has been reported in central Ohio, the Midland area south of the dialect boundary with the Inland North. In this dialect area, these six vowels have been described as failing to participate in any systematic shift-like rotation [4]. A relatively recent vowel change that has occurred in central Ohio has involved the merging of /a, ɔ/ into a single vowel category. This low back vowel merger causes a loss of the contrast between words such as *caught* and *cot*, particularly among young speakers.

However, recent finds from our lab pertaining to these two dialect regions have revealed parallel chain-like rotations of lax vowels which operate both on older configurations, i.e., created by the NCS in Wisconsin (WI) and on their more conservative dispersion in central Ohio (OH). These rotations take the general form of the Canadian Shift reported in parts of Canada [6] but their widespread occurrence in different dialect regions in the US suggests a much larger development which we have termed the North American Shift [5, 7]. The shift was found on the basis of detailed acoustic analysis of the lax vowels /ɪ, ε, æ, a/ produced in isolated citation-form *hVd*-words and in read sentences. Cross-generational changes to these vowels involved not only the positional rotations but also their spectral dynamics as measured as the amount of formant frequency change over the course of a vowel's duration. Thus, it appears that the mechanism of sound change utilizes dynamic information in vowels although the causes of these changing pronunciation patterns are not yet well understood.

The present paper takes a further step toward a better understanding of the current changes involving all six lax vowels in central OH and southeastern WI by examining spontaneous speech recordings produced by the speakers from the same corpus. It is of interest whether more natural productions of these vowels compared to those produced in citation form and read speech will evidence the existence of the new vowel rotations and whether there is a corresponding systematic change in spectral dynamics as a function of speaker generation.

Analyzing spontaneous speech has its obvious challenges because the vowels of interest occur in various segmental and prosodic contexts, in mono- and polysyllabic words, and are additionally affected by sudden changes in speech tempo. All these variations can influence the duration of and formant frequencies in vowels. This may result in problems in maintaining fine experimental control which, in turn, may compromise the accuracy of interpretation of sound change. For example, positional changes in the acoustic vowel space may arise from extensive variations in stress rather than from generational changes in pronunciation. Similarly, the generally shorter

vowel durations in spontaneous speech can reduce the extent of formant movement so that vowels will exhibit a lesser amount of spectral change [8]. Consequently, these reductions may not necessarily indicate loss of spectral dynamics in the process of sound change. Interpreting the results of spectral analysis of vowels in spontaneous speech thus involves a careful consideration of all the above factors in a valid assessment of sound change. These and similar concerns guided the analytical approach for which we opted in the current study.

2. Methods

2.1. Speakers

Speakers were 36 adult females who were born and spent most of their lives in either southeastern WI (Madison area) or central Ohio (Columbus area). They represented two generations of long-time residents and their ages ranged from 68-90 y.o. for the old (8 in WI and 7 in OH) and from 35-50 y.o. for the middle-aged speakers (10 in WI and 11 in OH). They were mostly professionals recruited through local community announcements and word of mouth; none of them had any speech disorder. The results for the corresponding male speakers produced similar patterns and will be reported elsewhere.

2.2. Stimulus material and procedure

Each speaker produced a spontaneous 5-10 minutes talk or engaged in a conversation with the experimenter. The topics included family, work, hobby, vacation, local events, retrospective stories, sports or plans for the upcoming weekend. The recordings took place at laboratory facilities at the University of Wisconsin-Madison or The Ohio State University using the same experimental set-up. One female experimenter at each location administered the task and monitored the recordings. A head-mounted Shure SM10A dynamic microphone was used and the speech was recorded and digitized at a 44.1-kHz sampling rate directly onto a hard disk drive.

2.3. Data analysis

Individual words containing at least one exemplar of each vowel were selected from each conversation. The words were stressed although with different magnitude across the speakers and the target vowels occurred in variable consonantal contexts in both mono- and polysyllabic words. Only tokens with vowel duration 50 msec or longer were analyzed acoustically for a total of 882 vowel instances from all 36 speakers. The most frequent vowels in the sample were /æ/ and /ɪ/ (19.8% and 19.1% of all productions, respectively) and the least frequent was /ɔ/ (10.1% of all productions).

Acoustic analysis of each token included word and vowel duration and frequency of the first two formants (F1 and F2) sampled at five equidistant time points in a vowel (20-35-50-65-80%) to estimate formant movement. A measure of formant movement, trajectory length (TL), was then derived from the 5-point measurements to assess generational (and dialectal) differences in the amount of spectral dynamics [8]. TL was first calculated for each of four separate vowel sections, i.e., 20-35%, 35-50%, 50-65%, and 65-80%, where the length of one vowel section (VSL) is,

$$VSL_n = \sqrt{(F1_n - F1_{n+1})^2 + (F2_n - F2_{n+1})^2} \quad (1)$$

The overall formant TL was then defined as a sum of trajectories of four vowel sections,

$$TL = \sum_{n=1}^4 VSL_n \quad (2)$$

3. Results

3.1. Positional changes

Traditionally, positional changes of vowels in the acoustic space which indicate sound change have been described on the basis of formant frequency values measured at the vowel's midpoint. Following this tradition, Figure 1 displays mean midpoint (50%) values for the vowels produced by OH speakers referred to as A4 (old speakers) and A2 (middle-aged speakers). Formant frequencies were not normalized because we did not expect considerable differences in the results based on our earlier comparison of normalized and non-normalized values for the present female speakers reported in [5]. The intersecting lines in the graphs indicate the mean 50%-point of F1 and F2 on the continuum between /æ/ and /ɛ/. This approach, adopted with some modifications from [4], uses the relations of F1 and F2 of these two vowels as a qualitative criterion for determining the development of the NCS.

In particular, the first stage of the NCS is manifested as a raising and fronting of /æ/ producing a change in the relative positions of /æ/ and /ɛ/ as NCS progresses. Following this first stage, the remaining vowels in the group undergo systematic chain-like clockwise rotations. In the present approach, we divided the F2 and F1 continua in terms of the calculated acoustic "category boundary" between /æ/ and /ɛ/ for the OH_A4 group even though the NCS is not operative in central OH. However, the æ/ɛ criterion determined on the basis of a conservative configuration between /æ/ and /ɛ/ in OH_A4 speakers helps us to visualize the relative proximity of the two vowels and compare the positional relations of the remaining vowels and changes in A2 adults (relative to OH_A4 speakers) in both dialects.

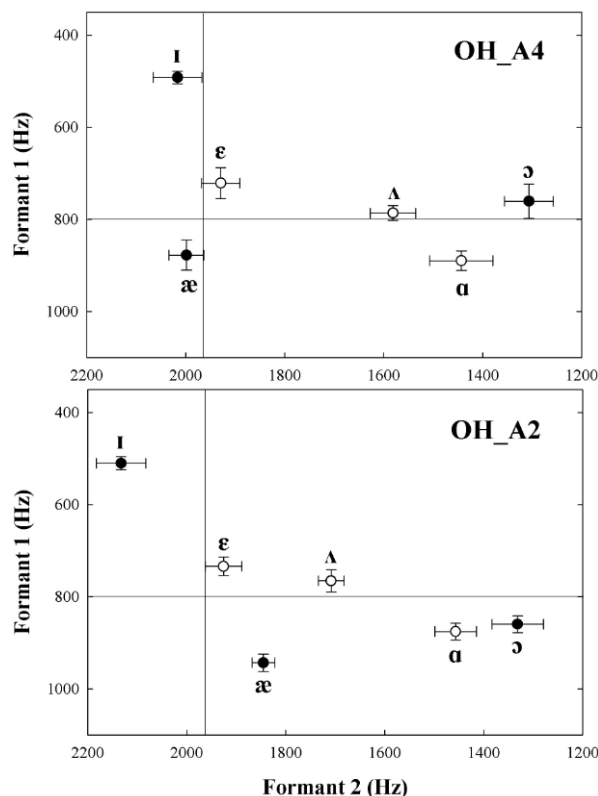


Figure 1: Mean (st.er.) vowel midpoints for OH speakers.

As can be seen in Figure 1, there are several positional changes in OH_A2 speakers relative to OH_A4. In particular, /æ/ is lowered and retracted, /a/ is raised and /ɔ/ is lowered which results in a merger (or near merger) of the two vowels, and both /ʌ/ and /ɪ/ are fronted. As it appears, the /ε/ is the only vowel that does not undergo positional change of some kind. This general anticlockwise vowel rotation in the acoustic space (except of /ɔ/ which descends to merge with the raised /a/) is a new sound change in central OH, reported recently in citation form vowels embedded in the common *hVd*-frame [5]. The present results confirm this pattern in spontaneous speech, which indicates that this sound change is represented in the lexicon of younger generations.

Figure 2 shows the corresponding results for WI speakers. The æ/ε boundary lines in both WI graphs are for OH_A4 speakers and are redrawn from Figure 1. Compared to OH_A4, almost all WI_A4 vowels are fronted (except for /ʌ/) and both /æ/ and /ε/ are raised. A set of rather drastic positional changes can be seen in WI_A2, including retraction of all 6 vowels, lowering of /ε/ and /ɔ/ and some raising of /a/. There are elements of the NCS in these rotations: the raised and fronted /æ/ in WI_A4 (Stage 1 of the NCS) is followed by a set of changes in WI_A2 including lowering of /ɔ/ (Stage 3), lowering and backing of /ε/ (Stage 4) and backing of /ʌ/ (Stage 5). However, the crucial element of the NCS chain is missing, which is the fronting and raising of /a/ in order to take up the position of /æ/. On the contrary, /a/ seems to be backing and raising to approximate the position of /ɔ/, thus the rotation is in the opposite direction. The final Stage 6 of the NCS, the lowering of /ɪ/ to approximate the position of /ε/ is also absent in WI_A2.

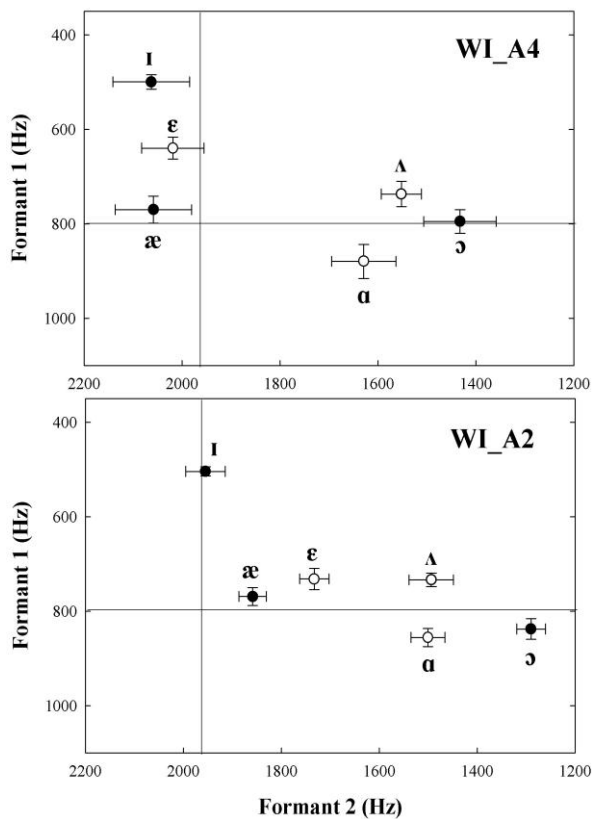


Figure 2: Mean (st.er.) vowel midpoints for WI speakers.

The results for WI speakers indicate that the NCS is not currently active in southeastern WI but there are some new

developments that are entering this vowel system. For example, there is an indication that /a/ and /ɔ/ will merge in the future and /æ/ will lower in the vowel space (although the latter has been seen primarily in younger generations, i.e., in young adults and children [5]). The current spontaneous speech data give us a better understanding how an older vowel system (A4) is being reorganized (A2) to set a stage for future developments.

3.2. Spectral dynamics

In general, we expected a reduction of the amount of spectral change (i.e., reduced formant movement) in A2 speakers compared with A4. This expectation is based on our previous analyses of vowels in citation form and read speech [5, 7]. Figures 4 and 5 display formant movement from samples taken at all 5 time points in the vowel; each vowel symbol is placed close to the 80%-point, indicating the direction of formant movement.

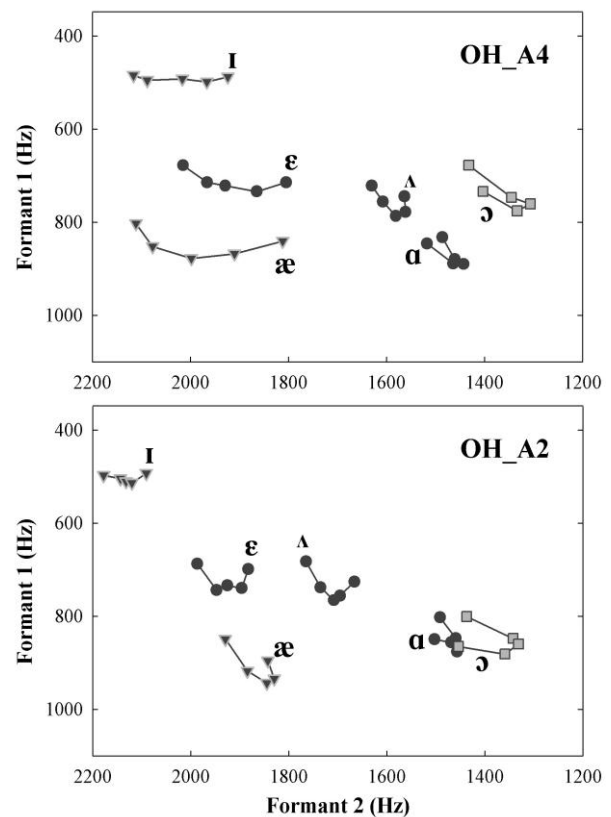


Figure 3: Mean dynamic formant pattern from 5 measurement points in each vowel for OH speakers.

We infer from the plots that there is a reduction in spectral dynamics in the front vowels /ɪ, ε, æ/ but the patterns for the remaining three vowels are more variable. Notably, there is a change in the direction of formant movement in /ʌ/ in both OH_A2 and WI_A2 groups. Although we can only speculate at present, this change may have taken place as a consequence of the retraction of /ε/ and reduction of its dynamics in A2 speakers to maintain a greater contrast between the two vowels. This possibility needs to be explored separately, however.

In general, WI_A2 speakers showed a greater reduction of formant movement in front vowels than did OH_A2 speakers. Measured as a TL difference between the A4 and A2 productions, the most drastic change was found for the vowel /æ/ in WI_A2 (a reduction from TL=491 Hz in A4 to TL=274 Hz in

A2). This change was much smaller in OH_A2 (TLs=397 and 323 Hz, respectively), most likely because the /æ/ was less “diphthongized” in OH_A4 speakers compared to WI_A4.

Spectral reductions in /ɪ, ε/ were comparatively smaller both in OH and WI and the greatest TL difference was for WI /ε/ (TL=385 Hz in A4 and TL=322 Hz in A2). Interestingly, both /ɑ/ and /ɔ/ had exactly the same amount of formant movement (TLs=351 Hz) in OH_A2, providing additional evidence for the low back merger. A rather unexpected increase (rather than reduction) in formant movement was found in /ʌ/ in both OH_A2 and WI_A2 groups along with a change in its direction. The increase in OH was from TL=254 Hz (A4) to TL=329 Hz (A2) and it was smaller in WI, from TL=293 Hz (A4) to TL=319 Hz (A2).

It needs to be pointed out that the systematic variations in spectral dynamics –mostly reductions as a function of speaker generation– are relatively small and occur in a shorter time course of vowels in spontaneous speech relative to read or citation form speech, even if the target words are in some ways “stressed” by the speaker. Thus, it is unclear whether statistical significance can be reasonably expected and further research and work with spontaneous speech, including appropriate statistical modeling, will provide more insights as to the effectiveness of statistical analyses in assessing subtle differences. Using a common repeated-measures ANOVA model in the present sample, the main effect of age group on TL was significant only for the vowel /æ/, where the TL difference was largest [F(1, 31)=4.95, p=0.034].

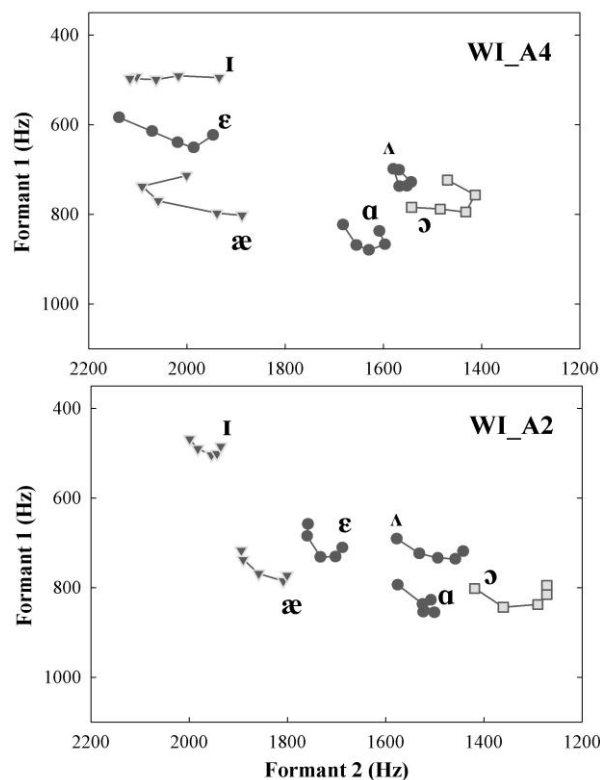


Figure 4: Mean dynamic formant pattern from 5 measurement points in each vowel for WI speakers.

4. Summary and conclusions

In this paper we have presented results for cross-generational changes in the characteristics of six lax vowels produced dur-

ing spontaneous speech. The new pattern of shift-like vowel rotations in central OH found earlier in citation-form vowels and in read speech was confirmed in the present analysis. A corresponding finding for southeastern WI confirmed that the NCS is not an active chain shift in younger generations. The elements of the NCS found in citation form and read speech were also evident in spontaneous speech.

The experimental results further showed that sound change is manifested not only in positional vowel rotations but also in changes in spectral dynamics. Consistent with previous findings, the spontaneous speech data showed a reduction in formant movement in younger generation for three front vowels and these reductions correspond to positional vowel changes in these speakers. There was a greater variability in the remaining three vowels, however.

While providing evidence for the existence of sound change on the basis of spontaneous speech data, this work exemplifies important challenges in future research as the field gradually transitions to the use of spontaneous speech in laboratory analyses. With the loss of experimental control of stress and of the effects of consonantal context on formant movement, we are facing fundamental questions regarding interpretation of acoustic cues in speech segments or the usefulness of statistical assessment of differences which are expected to be more subtle than in read speech. As the work with spontaneous speech progresses and methodology undergoes further refinement, it may become more advantageous to refocus our attention and turn to larger constituents such as syllables or words as fundamental units of phonetic analysis to better characterize dynamic information in speech. These and similar questions need to be addressed in future work.

5. Acknowledgements

This work was supported by NIDCD/NIH research grant R01DC006871. We thank Joseph Salmons for his contributions to this research.

6. References

- [1] Ohala, J. J. “The phonetics of sound change”, in C. Jones [Ed], Historical linguistics: Problems and perspectives, 237-278, Longman, 1993.
- [2] Lindblom, B. “Explaining phonetic variation: A sketch of the H & H theory”, in W. Hardcastle and A. Marchal [Eds], Speech production and speech modeling, 403-439, Kluwer, 1990.
- [3] Lindblom, B. Guion, S., Hura, S., Moon, S.-J. and Willerman, R., “Is sound change adaptive?” *Rivista di Linguistica*, 7:5-37, 1995.
- [4] Labov, W., Ash, S., and Boberg, C., “Atlas of North American English: Phonetics, phonology and sound change”, Mouton de Gruyter, 2006.
- [5] Jacewicz, E., Fox, R. A., and Salmons, J., “Cross-generational vowel change in American English”, *Language Variation and Change*, 23:45-86, 2011.
- [6] Boberg, C., “The Canadian shift in Montreal”, *Language Variation and Change*, 17:133-154, 2005.
- [7] Jacewicz, E., Fox, R. A., and Salmons, J., “Vowel change across three age groups of speakers in three regional varieties of American English”, *Journal of Phonetics*, 39: 683-693, 2011.
- [8] Fox, R., A., and Jacewicz, E. “Cross-dialectal variation in formant dynamics of American English vowels”, *Journal of the Acoustical Society of America*, 126: 2603-2618, 2009.