Cross-Dialectal Differences in Dynamic Formant Patterns in American English Vowels

Ewa Jacewicz and Robert Allen Fox

Abstract This chapter provides evidence that vowel inherent spectral change (VISC) can vary systematically across dialects of the same language. The nature and use of VISC in selected "monophthongs" is examined in three distinct dialect regions in the United States. In each dialect area, the dynamic formant pattern is analyzed for five different age groups in order to observe cross-generational change in relation to specific vowel shifts and other vowel changes currently active in each dialect. The dialect regions examined included central Ohio (representing the Midland dialect), southeastern Wisconsin (representing the Inland North whose vowel system is affected by the Northern Cities Shift) and western North Carolina (representing the South whose vowel system is affected by the Southern Vowel Shift). Following a description of these dialect areas, we first introduce principles of chain shifting and the transmission problem, originally developed in the fields of sound change and sociolinguistics. Selective acoustic data are then presented for each dialect region and cross-generational patterns of vowel change are discussed. The chapter concludes that variation in formant trajectories produced between vowel onset and offset (VISC) is central to what differentiates regional variants of American English in the United States. Furthermore, a systematic variation in VISC is found in cross-generational change in acoustic characteristics of vowels within each dialect. The perceptual relevance of this acoustic variation needs to be addressed in future research.

Abbreviations

A0	Child speakers (aged 8 to 12 years)
A1	Youngest adult speakers (aged 19 to 34 years)

E. Jacewicz · R. A. Fox (✉)
Department of Speech and Hearing Science, The Ohio State University, Columbus, USA
e-mail: fox.2@osu.edu

G. S. Morrison and P. F. Assmann (eds.), *Vowel Inherent Spectral Change*,
Modern Acoustics and Signal Processing, DOI: 10.1007/978-3-642-14209-3_8,

A2	Young adult speakers (aged 35 to 50 years)
A3	Older adult speakers (aged 51 to 65 years)
A4	Oldest adult speakers (aged 66 to 88 years)
F1	First formant
F2	Second formant
DARE	Dictionary of American Regional English
NCS	Northern Cities Shift
SVS	Southern Vowel Shift
TL	Trajectory length
VISC	Vowel inherent spectral change
VSL	Vowel section length

1 VISC in Regional Variation in Vowels

There is a long phonetic tradition which views a vowel as a static target, i.e. a linguistic category whose position in the acoustic space (defined as a two-dimensional F1 × F2 plane) can be adequately characterized by the formant values at the vowel's putative steady-state (see Chap. Static and Dynamic Approaches to Vowel Perception). This approach was used in the classic study of the American English vowel system by Peterson and Barney (1952) that brought to light considerable variation in the position of the static target within each vowel category. Years later, this type of acoustic variation became of particular interest to sociolinguists studying regional dialects and language change manifested in vowel shifts and mergers (e.g., Labov 1994; Labov et al. 1972; Thomas 2001). Using the steady-state approach, the regional differences in the overall positions of nominal monophthongs in the F1 by F2 plane became apparent. However, the question of how vowels may differ cross-dialectally in terms of the extent and direction of dynamic formant movements has not been addressed so far.

Vowel inherent spectral change or VISC (Nearey and Assmann 1986) has been found in American, Canadian, and Australian English vowels (e.g., Andruski and Nearey 1992; Hillenbrand et al. 1995; Watson and Harrington 1999). Although the primary focus of work on VISC has involved its perceptual relevance for vowel identification (see Chap. Static and Dynamic Approaches to Vowel Perception, for an overview), these studies brought to light the possibility that VISC may vary systematically across regional varieties of English. If so, VISC may potentially play an important role in synchronic variation in vowels and sound change over time. However, no previous studies have examined whether VISC varies systematically across dialects. This chapter aims to provide some insights into the nature and use of VISC in three distinct regional varieties of English spoken in the United States. In the *Atlas of North American English* (Labov et al. 2006), these regions are labeled Inland North (the Great Lakes region), Midland (south of the Inland North) and the South encompassing several southern states. These three broad dialect regions are defined on the basis of lexical, phonological and grammatical differences (Labov et al. 2006; Wolfram and Schilling-Estes 2006).

A particular striking feature in these regional dialects is that their respective vowel systems are affected by a set of changes termed "vowel shifts" (see Labov 1994; Labov et al. 2006 for extensive discussions and classifications of vowel shifts). According to Labov et al. (2006), the Northern Cities Shift (NCS) is a set of vowel rotations which affects currently northern American English spreading from Buffalo, NY through northernmost Ohio, Michigan, northern Illinois to southeastern Wisconsin and includes the Inland North. A different and much more complex set of changes termed the Southern Vowel Shift (SVS) is found in the South (some parts of this broad region are also defined by another shift called the Back Upglide Shift). Features of the SVS are found in the southeastern states of Virginia, North and South Carolina, Georgia and extend through Alabama, Louisiana, Oklahoma and parts of Texas, including also Kentucky, Tennessee and southern Ohio. Finally, the Midland is regarded as a transitional region between the Inland North and the South, encompassing central Ohio, part of Indiana, Missouri and Kansas. The Midland's vowel system is not reported as undergoing any distinct pattern of shift but there are other ongoing vowel changes in parts of this region.

In this chapter, we present acoustic data sampled from three relatively homogeneous speech communities situated within each of these broad dialect regions. Southeastern Wisconsin (the area between Madison, Milwaukee and Green Bay) was selected as a testing ground for the NCS. Central Ohio (Columbus and adjacent areas) was selected as representative of the Midland dialect. Western North Carolina (including Jackson and Haywood counties in the Appalachian region, an area that was labeled in earlier dialect atlases by Kurath (1949) and Kurath and McDavid (1961) as "Mountain Southern," was selected as the core area of the Southern Vowel Shift. In Labov et al. (2006), this region is now defined as the Inland South, the area in which the Southern Vowel Shift is most developed. The map in Fig. 1 shows the respective locations of these speech communities in each state.

The widespread occurrence of vowel shifts across regions in the United States have stimulated research in dialectology and sociolinguistics, asking questions as to why such vowel changes take place, how they originate and how are they transmitted within a given region to create a dialect or "regional variety" of the language. In our present examination of acoustic characteristics of vowels in these three dialect regions, we aim to illuminate the cross-generational pattern of vowel change with reference to VISC. Although we do not take a particular position with respect to the causal factors in vowel shifts (which involve a complex set of phonetic and social variables), we provide evidence that the extent of VISC changes with each younger generation of speakers from the same speech community. Therefore, cross-generational vowel change within each dialect carries a corresponding change in VISC and this systematic variation may be associated with a particular stage of a given vowel shift. Before we turn to our results, the next section will present a few remarks on the notion of "chain shifting" and its application to American English.

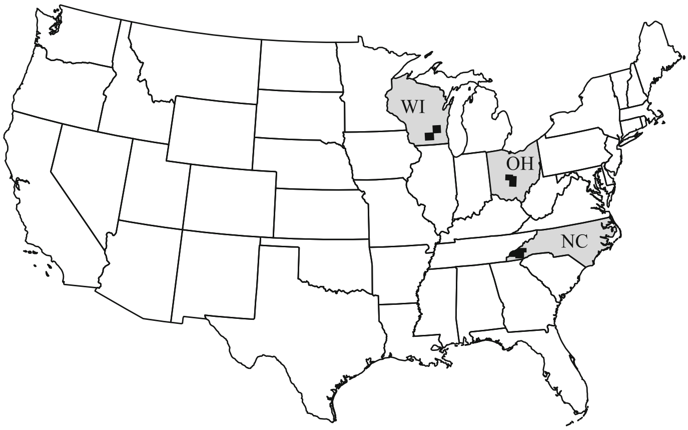

Fig. 1 The three testing areas in southeastern Wisconsin (WI, representing Inland North), central Ohio (OH, representing Midland dialect) and western North Carolina (NC, representing Inland South)

2 Chain Shifting and the Cross-Generational Sound Transmission

Since the earliest reconstructible stages (Proto-Germanic) the rich vowel systems of Germanic languages (including English) have been undergoing continuous change over time. This process has been termed "chain shifting" in descriptive analyses of historical linguists (Stockwell 1978). Based on phonological analyses of historical scripts involving distinctive vowel features such as high/low and front/back, it was found that one or more vowels can change position within the phonological system (a local change) and this change then gradually affects the vowel system as a whole, as other vowels are "pushed" away from the moving vowel(s) or "pulled" into newly opened positions in the vowel system to maintain earlier sets of distinctions. It is in this sense that vowels are thought of as "moving" in the vowel system in chains so that their individual relative positions change over time (i.e., across different generations of speakers) in a specific direction. In the subsequent linguistic phonetic studies, these relative positions have been measured and defined in the two-dimensional F1 by F2 plane, showing that a particular vowel can "rise" in the vowel space (such as from /æ/ to /ɛ/ which involves a decrease of F1 frequency) or "fall" (such as from /u/ to /o/ which involves an increase of F1 frequency).

In American English, the sociolinguistic work of Labov and his colleagues (notably Labov et al. 1972, 2006; Labov 1994) has pointed to the operation of three general principles of chain shifting across regions of the United States which are thought to motivate vowel shifts such as the NCS or SVS. These principles date back to Sievers (1876/1881) and have been treated since then as a central type of language change. Accordingly, in chain shifts, Principle I predicts that long vowels rise, Principle II states that short vowels fall and Principle III predicts that back vowels will move to the front. Labov restates these principles in more modern terms in his *Peripherality Hypothesis*, which proposes that in chain shifts, tense vowels (which have longer durations and more extreme articulatory positions) move upward along a peripheral track and lax vowels move downward along a non-peripheral track (see Labov et al. 2006, pp. 15–20, for further details, including an effort to formulate a single principle of chain shifting).

In our present cross-dialect work, we expect to find at least some indication of the operation of these general principles of chain shifting. In particular, we examine positional changes of the three short vowels /ɪ, ɛ, æ/ in the production of five generations of speakers from each dialect area in order to evaluate the operation of Principle II. However, it needs to be emphasized that the principles of vowel change were formulated on the basis that vowels are "static targets" and the only change involves their positional movement in the vowel system. No previous research indicated that vowel change may also involve "intrinsic" spectral properties such as VISC. Therefore, the second aim of our research is to explore how changes in VISC may be manifested across generations of individuals who grew up in the same dialect area as both speakers and listeners of the local dialect.

The logical question arises as to how the vowel changes are transmitted through multiple generations always shifting in the same direction. The transmission problem cannot be confined to a single generation of speakers. Indeed, Stockwell (1978) has identified the "perseverance problem" (how successive generations keep moving vowels in the same directions) as a central issue in sound change. Studying synchronic variation in vowels, sociolinguists overwhelmingly point to the role of social factors in sound transmission such as social class, age, gender, ethnicity, neighborhood and social networks (see Labov 2001 for extensive discussion). In terms of purely phonetic effects, consonantal contexts (usually consonants in the syllable coda) are identified as a primary aspect in segmental conditioning of positional vowel change. In our work, we admit the possibility that cross-generational vowel change (both in terms of positional change and change in VISC) may be, in part, a direct effect of prosodic organization of English language and its specific use of linguistic stress (see Jacewicz et al. 2006, for extensive discussion). Because vowels in stressed syllables are longer and exhibit a greater spectral change compared to reduced vowels in unstressed syllables, they may trigger a shift-like change in that younger generations acquire as their norm those variants which were produced with greater emphasis a generation earlier. To explore these effects, our present acoustic analysis includes both emphatic and nonemphatic vowel variants produced by each successive generation of speakers.

What follows is a report on a few results of a large-scale acoustic study undertaken to examine cross-generational changes in vowels in the three selected dialect regions. The study is conducted in "apparent time" (as opposed to "real time"). From a strictly technical point of view, an apparent-time study is a specific type of cross-sectional study. The apparent-time methodology considers the speech samples of different generations of speakers collected during a single time period to be representative of different "stages" of the dialect and not simply cross-sectional differences as a function of age (see Bailey et al. 1991). Apparent-time studies are a common practice in sociolinguistic research because of obvious problems in conducting a study in a "real time." Real-time studies require collection of samples from the population of interest over an extended stretch of time (usually, decades). These studies would include longitudinal studies (data collected from the same individuals across their life span) and historical studies (e.g., comparing older speech recordings, e.g., from the *DARE* project (*Dictionary of American Regional English*, Cassidy and Hall 1985), to recent speech samples. There are obvious challenges to the "real time" approach, either in terms of the time commitment and/or poor quality of older recordings. However, the results of apparent-time studies prove to be generally reliable if speech communities remain largely the same over the course of time (Chambers 2003).

In each testing location, five generations of speakers were recorded who were born, raised and spent most of their lives in the respective regions. The participants fell into five age groups (A0–A4) whose ages (in years) were: children 8–12 (A0), youngest adults 19–34 (A1), young adults 35–50 (A2), older adults 51–65 (A3) and oldest adults aged 66 and up to the late 80s (A4). There were both males and females in each age group for a grand total of about 400 speakers. The number of speakers varied from 9 to 16 per gender/age/dialect subgroup, depending on availability of subjects within the time frame of the project. Each participant produced a set of single words in citation form, a set of sentences with a variable main sentence stress and a spontaneous talk. The speech material was recorded in the years 2006–2008.

In this chapter, we will report mainly on the results for three vowels /ɪ, ɛ, æ/ which were produced by 198 female speakers (66 per dialect) in a sentential context in stressed and unstressed syllables in a/b_dz/ environment (the words were *bids*, *beds*, *bads*). The vowels in stressed syllables are referred to here as emphatic and those in unstressed syllables as nonemphatic to underscore the fact that such differences are not always related to linguistic stress per se but the emphasis can be produced by a variety of phonetic and paralinguistic factors. Female speakers were chosen to simplify the presentation only and the data from male speakers produce similar patterns. Each speaker read two randomly presented repetitions of a sentence containing the word of interest in each emphasis position for a grand total of 2376 vowel tokens (66 speakers $\times$ 3 vowels $\times$ 2 emphasis positions $\times$ 2 repetitions $\times$ 3 dialects). Our long-term working hypothesis is that the emphatic vowels (rather than nonemphatic) will lead the sound change across generations in a specific direction.

Emphatic vowels are expected to occupy a more peripheral position relative to nonemphatic vowels due to the expansion of the vowel space as a function of emphasis (e.g., Moon and Lindblom 1994). Also, emphatically spoken words "tend to be articulated more forcefully, resulting in longer and more extensive vowel […] gestures" (Agwuele et al. 2008, p. 207) which suggests a greater formant movement in emphatic vowels. In this chapter, we will observe whether and how formant dynamics change cross-generationally in both emphatic and nonemphatic variants in each dialect region. Our present interest focuses on two cross-generational changes: (1) differences in the relative positions of the vowels in the acoustic space and (2) variation in the extent of formant movement (or the amount of VISC). For each dialect, we expect that vowels will change position in the vowel space as a function of speaker generation. What we cannot predict is the extent of the corresponding change in formant dynamics or whether dialects differ in the use of VISC cross-generationally.

3 The Midland Vowel Changes in Central Ohio

We begin with a presentation of the data for the Midland dialect in central Ohio, which is not thought to be participating in any known chain shift at present. The panels in Fig. 2 show cross-generational plots across all age groups (A4–A0) for a total of 66 female speakers. The data points in the plots indicate mean F1 and F2 values measured at five equidistant temporal locations corresponding to the 20–35–50–65–80 %-point in the vowel. These multiple measurement points allow us to estimate the extent of formant movement spanning over the central 60 %-section of the vowel to the relative exclusion of immediate effects of surrounding consonants on vowel transitions. Although this approach uses five points (rather than the commonly used onset and offset measurements), it still only provides an estimate of the shape of the actual trajectory. Yet, as we will see, the five-point measurement technique is sufficient enough to capture the basic variation in VISC across dialects and ages (along with the variation due to vowel emphasis) and is relatively easy to implement in analyzing a larger corpus. More details pertaining to this analysis can be found in Fox and Jacewicz (2009). In the current plots, we follow the sociophonetic tradition in displaying the vowels in the F1 by F2 plane in which the axes show values in descending order. Direction of formant movement is indicated by arrows.

Turning to the plots, we see substantial variation in formant dynamics across all three individual vowels each of which is classified as a nominal monophthong in American English. Formant trajectories for the emphatic and nonemphatic versions of all three vowels are relatively close and parallel when produced by oldest speakers (A4). With each successively younger generation, the emphatic and nonemphatic variants of the vowels become progressively more separated one from the other. This represents a change in the basic position of the vowels in the acoustic space. A second difference can be seen in terms of a change in vowel dynamics. As the relative position of the vowel changes across generations, so does the amount of VISC. With each younger generation, each vowel exhibits less formant movement (especially in F2) becoming progressively less diphthongal. The differences between A4 and A0 groups are rather drastic and include also the direction of formant movement. Although emphatic vowels are consistently more peripheral than nonemphatic across all generations, the most salient cross-generational differences seem to lie in the amount of VISC.

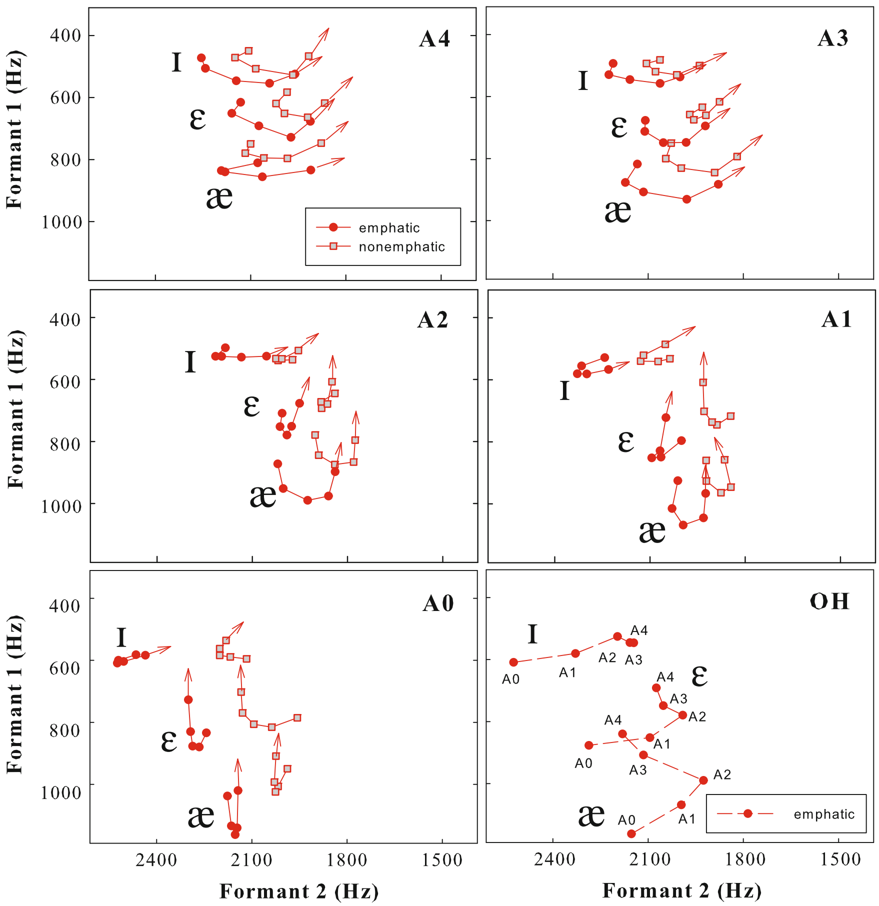

Fig. 2 Means of F1 and F2 measured at five equidistant timepoints in a vowel across five age groups (A4–A0) for the central Ohio dialect. Positional changes of vowel midpoints are shown in the last panel (*bottom right*)

In an initial attempt to quantify variation in the size of the VISC across generations and vowels, we calculated a trajectory length (TL) in the F1 by F2 acoustic space for each vowel. The TL represents a sum of the lengths of the four separate vowel sections between the 20 and 80 %-point, where the length of one vowel section (VSL) is:

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

The assumption is that a longer TL reflects a greater amount of VISC. Shown in Table 1 are the mean TLs for the Ohio speakers. Although there is some variability in the means, there is an overall tendency for the TLs for younger speakers to be shorter than those for older speakers, suggesting that the amount of VISC changes

Table 1 Mean trajectory length (in Hz) for Ohio speakers broken down by vowel, emphasis condition and age group; standard deviations are in parentheses

Vowel	Emphasis condition	Age group				
		A4	A3	A2	A1	A0
/ɪ/	Emphatic	440 (116)	394 (70)	353 (101)	344 (75)	320 (110)
	Nonemphatic	390 (106)	306 (91)	253 (117)	251 (86)	223 (90)
	Combined	415 (111)	350 (95)	303 (119)	303 (94)	272 (110)
/ɛ/	Emphatic	456 (104)	423 (95)	346 (101)	378 (67)	353 (116)
	Nonemphatic	313 (62)	287 (107)	244 (80)	270 (79)	308 (147)
	Combined	38 (111)	355 (121)	295 (104)	324 (91)	330 (131)
/æ/	Emphatic	593 (105)	538 (123)	435 (137)	469 (181)	430 (149)
	Nonemphatic	419 (114)	380 (121)	331 (110)	339 (155)	303 (95)
	Combined	506 (139)	459 (145)	383 (131)	404 (177)	367 (139)

across generations. Also, the mean TLs vary with vowel category: they tend to be longer for the vowel /æ/ and shorter for either /ɪ/ or /ɛ/, which do not seem to differ much from one another. Across all vowels and groups, the mean TLs of the emphatic vowels are longer than of the nonemphatic.

In the final panel of Fig. 2, the formant values at the 50 %-point (often considered as representing the *vowel nucleus* in the sociophonetic literature) of the emphatic vowels are replotted for an immediate display of positional changes across generations. We find a progressive fronting of /ɪ/ with each younger generation. The pattern for /ɛ/ is clearly different in that the vowel first lowers, is then fronted (A1) and shows further fronting when produced by children. The greatest positional change is in /æ/ which not only consistently descends across generations but displays considerable backing and then fronting in young adults and children. Given the ages of the speakers in the A0 group (8–12 year-old girls; mean age for the A0 group was around 10 in all three dialect groups) we would expect that their generally shorter vocal tracts would produce some elevation of both F1 and F2 compared with the adult female speakers. However, these differences may not be as substantial as those between boys and adult males (see Lee et al. 1999; Assmann et al. 2009) and, given the overall direction of the positional differences between the A0 and A1 groups, children seem to continue the path of vowel change seen in the adult cross-generational data.

How do these cross-generational changes relate to the general principles of chain shifting? As pointed out, there are no attested vowel chain shifts in central Ohio. We thus do not expect the operation of Labovian Principle II according to which the non-peripheral (lax) vowels move downward along a non-peripheral track. However, we see specific changes which seem to parallel those in accordance with Principle II. For example, the vowel /æ/ seems to have descended to the bottom of the non-peripheral track and is being pressured to move further. Will it enter the lower peripheral track and, being fronted, will it start rising and breaking to introduce the Northern Cities Shift to this geographic area? Only time can answer this and new data from new generations of Ohioans. Similarly, the vowel

/ɛ/ follows the general direction of /æ/. Does its fronting in children's production signal a reorganization of vowel subsystem in this corner of the vowel space or will it return to its position as produced by children's grandparents? If so, will it develop a greater formant movement as generations ago? Although not related to Principle II, does the fronting of /ɪ/ indicate a more general process of fronting in American English (such as in response to the back vowel fronting) or is this cross-generational change specific to central Ohio? Further work is needed to explain these interesting patterns.

4 Northern Cities Shift in Southeastern Wisconsin

While vowel changes in central Ohio do not appear to be part of a chain shift, the NCS reflects an attested chain shift, first reported in late 1960s and widely recognized since Labov et al. (1972). Before presenting our Wisconsin data, we summarize briefly the main characteristics of NCS. The triggering event for the NCS is the raising and fronting of /æ/. This event initiates a series of the following vowel rotations: fronting of /ɔ/, lowering of /ɑ/, lowering and backing of /ɛ/, backing of /ʌ/, and some centralization and backing of /ɪ/. In this way, the vowels exchange their positions like in a chain, operating along the peripheral track (/æ, ɔ, ɑ/) and non-peripheral track (/ɛ, ʌ, ɪ/) until the entire cycle is complete. The exact order of the stages in the chain in the NCS is subject to some inter-speaker variation (see Labov et al. 2006, for further details) but the general rotation is assumed to be clock-wise as shown, schematically, in Fig. 3. Some variation in directionality also occurs within this pattern, as shown by Gordon (2001). Most notable for present purposes is that /ɛ/ can move down rather than back.

Our results for three vowels /ɪ, ɛ, æ/ come from 66 Wisconsin female speakers who produced exactly the same speech material as Ohio speakers and were tested under the same experimental conditions. The recordings were made at the University of Wisconsin-Madison. Figure 4 displays plots across all age groups (A4–A0). The oldest speakers (A4) show the characteristic mark of NCS, i.e., the raising and fronting of /æ/, particularly in its early portion before vowel midpoint (known as Northern breaking). This vowel exhibits an extended formant movement, which is also true for the other two vowels in A4 productions, /ɪ/ and /ɛ/. In the next generation (A3), while /æ/ does not change much in general, we see a

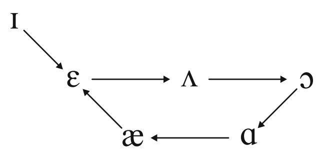

Fig. 3 Schematic rotation of vowels in the Northern Cities Shift indicating a change in their pronunciation. "bid" → "bed" → "bud" → "bawd" → "bod" → "bad" → "bed"/"bid"

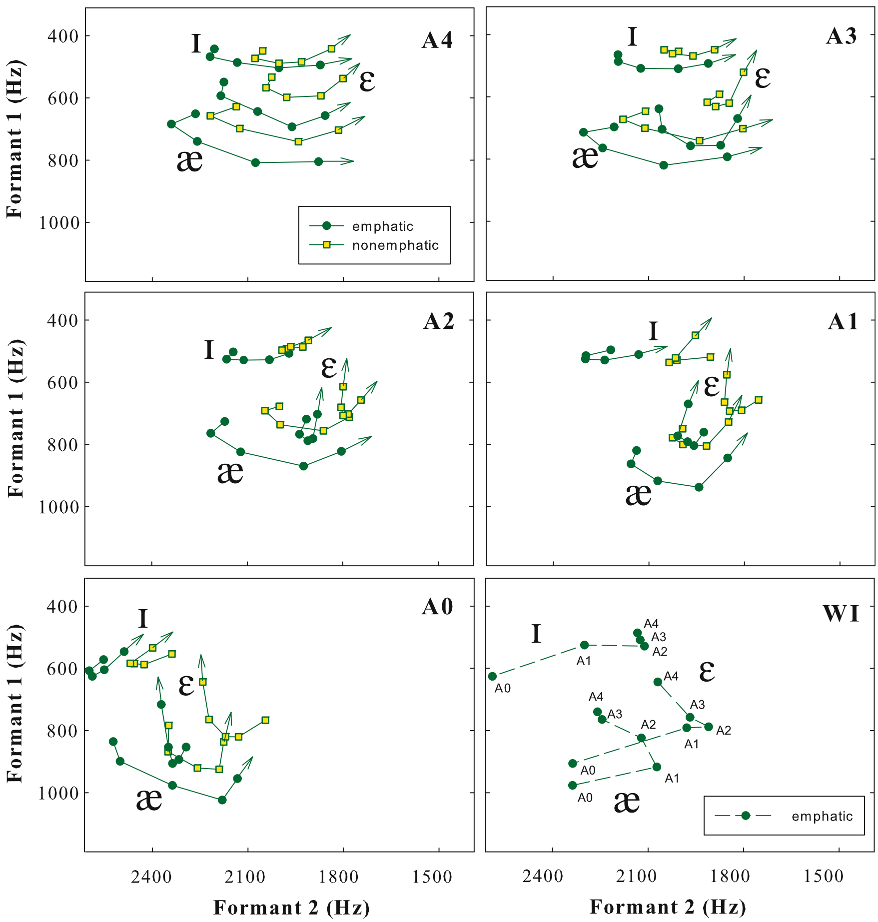

Fig. 4 Means of F1 and F2 for the southeastern Wisconsin dialect along with positional changes of vowel midpoints (*bottom right*)

substantial lowering and backing of /ɛ/, particularly in the stressed position so that the emphatic /ɛ/ tends to overlap with the nonemphatic /æ/. The next generation (A2) introduces a significant change in that /ɛ/ not only moves further downward but it changes its acoustic characteristics. As can be seen, the extensive formant movement found in earlier generations is greatly reduced, especially in emphatic positions. The "monophthongization" of /ɛ/ may signal a later stage of the NCS in this geographic area to be followed only by backing of /ʌ/ and shifting of /ɪ/ in response to the movement of /ɛ/. In the next generation (A1), we see a further decrease in formant movement, particularly in /ɪ/ (which also undergoes fronting) and to some extent in /æ/, which also lowers. Children's vowels (A0) generally maintain the direction shown in young adults (A1) although their vowels have a rather high F2 which can be in part due to their unnormalized formant values.

Table 2 Mean trajectory length (in Hz) for Wisconsin speakers broken down by vowel, emphasis level and age group; standard deviations are in parentheses

Vowel	Emphasis condition	Age group				
		A4	A3	A2	A1	A0
/ɪ/	Emphatic	478 (94)	419 (100)	377 (127)	385 (93)	374 (119)
	Nonemphatic	311 (129)	280 (125)	265 (66)	308 (85)	300 (122)
	Combined	395 (140)	348 (131)	321 (115)	347 (95)	337 (124)
/ɛ/	Emphatic	503 (136)	467 (146)	350 (121)	364 (113)	399 (148)
	Nonemphatic	369 (118)	308 (108)	252 (55)	270 (60)	416 (176)
	Combined	436 (141)	387 (150)	301 (105)	317 (100)	407 (160)
/æ/	Emphatic	694 (111)	658 (156)	611 (121)	516 (124)	607 (216)
	Nonemphatic	562 (196)	518 (159)	423 (171)	362 (104)	421 (155)
	Combined	606 (161)	588 (171)	517 (174)	439 (136)	514 (208)

As was done for the Ohio data, mean TLs were calculated for each vowel across all Wisconsin speakers. These means are shown in Table 2. While pattern of TLs of both /ɪ/ and /ɛ/ is not much different from that seen in the Ohio dialect, the TLs for /æ/ are longer for Wisconsin than Ohio speakers which suggests a greater extent of VISC in Wisconsin /æ/.

In terms of cross-generational changes, unlike in the Ohio data, each Wisconsin vowel showed a somewhat different pattern of variation in TL. For /ɪ/, mean TLs did not differ substantially across the age groups. For /ɛ/, TLs were longest in A4, dropped sharply in A2 but then increased progressively in A1 and A0. Finally, the vowel /æ/ revealed yet another pattern. Mean TLs for the emphatic variant remained approximately the same across all generations, with some drop in A1. However, there were more changes in the nonemphatic variant: the longest mean TLs were in A4 and then decreased progressively across the younger age groups until reaching the shortest values in A1. They increased again in A0.

How do these cross-generational vowel changes relate to the NCS? If we track the movement of vowel nuclei only (the 50 % point) as displayed in the last panel of Fig. 4, we see the raised /æ/ in A4 and its progressive lowering and backing across next generations along with an unexpected fronting in children. While the raised /æ/ is clearly a mark of the NCS, its subsequent downward movement is not. However, in terms of the *Peripherality Hypothesis*, the vowel could have reached its highest possible position in the peripheral track and entered a non-peripheral track which would explain its lowering and some backing in accord with Principle II. This, of course, is not the expected progression of the vowel in the chain, which should continue rising to approximate the positions of /ɛ/ and /ɪ/ with each younger generation. It could be the case that such vowel rotations can be observed in other consonantal contexts since contextual variation is a possible conditioning factor. The cross-generational data from the present speakers in this particular consonantal context do not indicate a continuous rising of /æ/ in this part of Wisconsin. The lowering and backing of /ɛ/ and /ɪ/ across A4, A3 and A2 groups is entirely in agreement with the direction of the NCS (and the Principle II). However, younger generations (A1 and A0) introduce a directional change to this pattern (i.e., fronting). It is unclear how this change relates to the NCS and this issue can only be solved by data from future generations.

Operation of the NCS has not been definitively linked to changes in formant dynamics aside from observations about the ongliding of /æ/. Yet, based on the present results, the cross-generational changes in VISC are apparent. Some changes are more abrupt than others such as those in the emphatic /ɛ/ and to a lesser extent in /ɪ/ introduced by A2 speakers. Namely, in the process of lowering and backing, the extent of formant movement of /ɛ/ undergoes a drastic reduction in comparison to A4 along with a change in directionally of movement. A relatively smaller reduction in the movement of /ɪ/ prepares the stage for a greater change in F2 the next generation (A1) along with its progressive fronting. Is such variation in VISC linked to the particular stages of the shift? Certainly, formant values measured traditionally at the vowel's midpoint cannot reveal such potentially complex relationships and we do not know whether sudden changes in formant dynamics precondition stages in vowel shifting. The general principles of chain shifting were formulated on the basis of the assumption that the nominal monophthongs do not exhibit formant movement.

5 Southern Vowel Shift in Western North Carolina (Inland South)

The third set of cross-generational data included in this chapter pertains to the operation of the chain shift called the SVS which defines the dialects of the South. As indicated earlier, our participants come from the Appalachian area in western North Carolina which is identified as a center of the most advanced features of the southern vowel system (Labov et al. 2006). The present data come from 66 female speakers who were born and raised in the area, produced the same speech material and were tested under the same experimental conditions. The recordings took place at Western Carolina University in Cullowhee, NC. In this section, we first present cross-generational data for the vowels /ɪ, ɛ, æ/ in order to be consistent with the previous reports for Ohio and Wisconsin. We then discuss the operation of the SVS in a greater detail in a set of vowels produced in citation form.

Figure 5 displays plots across all age groups (A4–A0) for 66 female speakers. The trajectory shapes of vowels produced by A4, A3 and A2 speakers indicate a type of formant movement typical of the "Southern drawl" or breaking, where the vowel "breaks" into two parts (Sledd 1966). That is, the vowel in *bit* may be pronounced as [bi:jɪt] giving an auditory impression of two vowels in a sequence connected by the glide [j]. Note the proximity of /ɪ/ and /ɛ/ in the production of A4 and A3 speakers and a continuous separation of all vowels (including their emphatic and nonemphatic variants) with each younger generation. The last panel (bottom right) tracks the positional changes of vowel midpoints, indicating that the

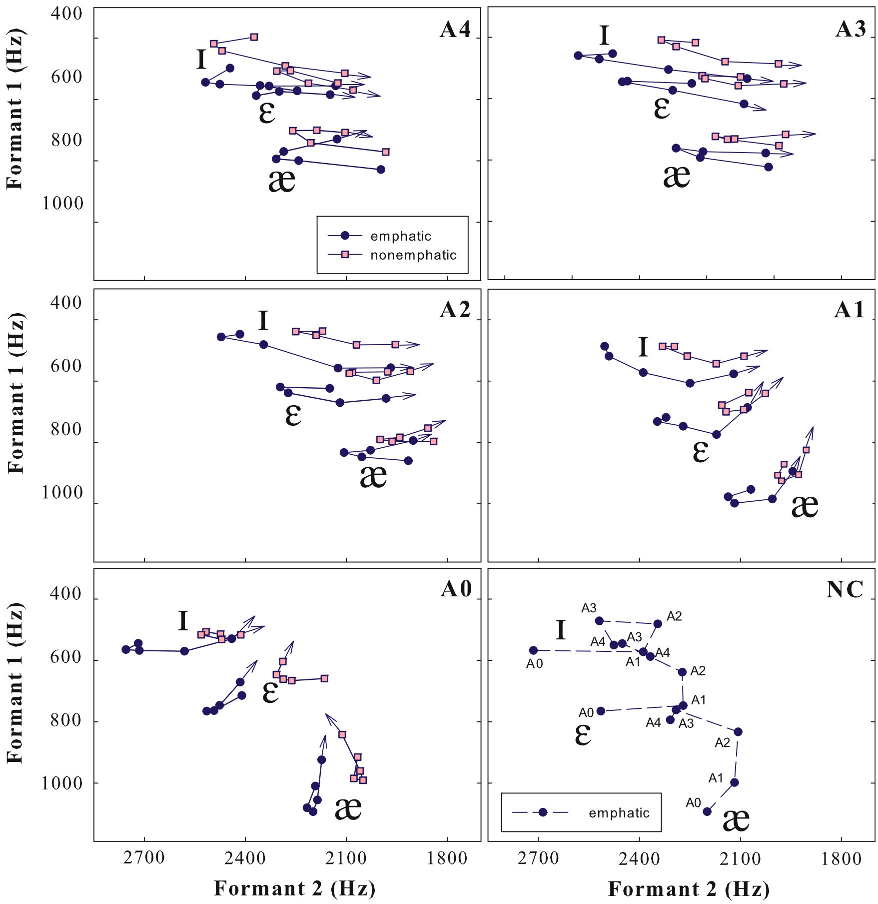

Fig. 5 Means of F1 and F2 for the western North Carolina (Inland South) dialect along with positional changes of vowel midpoints (*bottom right*)

vowels /ɛ/ and /æ/ begin to descend in the acoustic space in the production of A2 speakers while /ɪ/ shows a less decisive pattern. Table 3 shows a summary of TL values across age groups. Notable differences from both Ohio and Wisconsin data include generally longer TLs for North Carolina speakers, indicating greater formant movement. Also, it is the A3 group (and not A4) that shows the greatest extent of VISC.

A trend common in all three dialects (including the Inland South) was that TLs for emphatic variants were significantly longer than for nonemphatic; also, TLs were shorter for the younger than for the older speakers. However, of greatest interest in examining the North Carolina vowels is not only the extent to which these three front vowels have changed cross-generationally; the operation of the SVS involves two additional essential vowels, /i/ and /aɪ/. Unfortunately, the North Carolina speakers were not asked to produce the vowel /i/ in sentences as variation

Table 3 Mean trajectory length (in Hz) for North Carolina speakers broken down by vowel, emphasis condition and age group; standard deviations are in parentheses

Vowel	Emphasis condition	Age group				
		A4	A3	A2	A1	A0
/ɪ/	Emphatic	623 (175)	717 (166)	694 (135)	587 (124)	541 (210)
	Nonemphatic	607 (169)	520 (138)	437 (156)	372 (209)	375 (232)
	Combined	616 (156)	619 (180)	565 (194)	480 (200)	458 (234)
/ɛ/	Emphatic	557 (178)	730 (232)	614 (179)	504 (152)	449 (190)
	Nonemphatic	474 (126)	441 (122)	422 (145)	337 (128)	293 (102)
	Combined	515 (155)	586 (234)	518 (188)	421 (161)	371 (170)
/æ/	Emphatic	662 (213)	713 (255)	552 (186)	443 (176)	460 (166)
	Nonemphatic	489 (167)	495 (180)	392 (155)	323 (71)	388 (188)
	Combined	576 (206)	604 (244)	472 (187)	383 (144)	424 (178)

in this vowel was not of interest to the study as a whole (i.e., the vowel is not a part of NCS-rotations nor are there any changes to /i/ in central Ohio). For this reason, we now turn to a different set of acoustic measurements taken from citation form words in the *hVd* context, which were produced by all participants of the study.

The SVS is a more complex chain shift than the NCS and is manifested somewhat differently in different parts of the South. Generally, the triggering event of the SVS (Stage 1) is the deletion of the offglide in /aɪ/ and its monophthongization so the vowel is pronounced basically as an /a/. Stage 2 is the centralization and lowering of /e/ and fronting and raising of /ɛ/, the so called reversal of the front/back locations of /e/ and /ɛ/. Stage 3 is a parallel lowering of /i/ and fronting and raising of /ɪ/ which results in another reversal of the front/back locations of /i/ and /ɪ/. The upper left panel in Fig. 6 captures Stage 3, the most advanced stage of the SVS, in the production of A4 speakers. For clarity of presentation, vowel /e/ is not included in the plots in Fig. 6 as it has considerable overlap with /ɪ, ɛ, æ/.

We will focus our discussion on the subsequent cross-generational reorganization of the vowel subsystem affected by the SVS. First, we find the oldest speakers producing a monophthongal version of the diphthong /aɪ/. Their vowels /i/ and /ɪ/ are in close proximity which may be a reflection of their earlier reversal. A4 speakers also produce the raised and fronted /ɛ/ and /æ/. Beginning with the A3 generation, there is progressive re-diphthongization of the monophthongal /aɪ/, progressive fronting of /i/ and lowering and backing of /ɪ, ɛ, æ/. Finally, the young adults (A1) have a very fronted /i/, a clear diphthong /aɪ/, and no breaking in /ɛ/ and /æ/ although both vowels still show a considerable amount of formant movement. The vowel /ɪ/, on the other hand, does remain diphthongized and is undergoing fronting as if it were "pulled" by /i/. The children's data show a continuation of this trend: both /ɛ/ and /æ/ are slightly lowered and less diphthongized and /ɪ/ still shows a substantial amount of spectral change. However, both /i/ and /ɪ/ are separated one from another. Will /ɪ/ undergo a further reduction in VISC in the next generation from this area and will the existing subsystem be a subject to further reorganization? Only time will allow us to answer these questions.

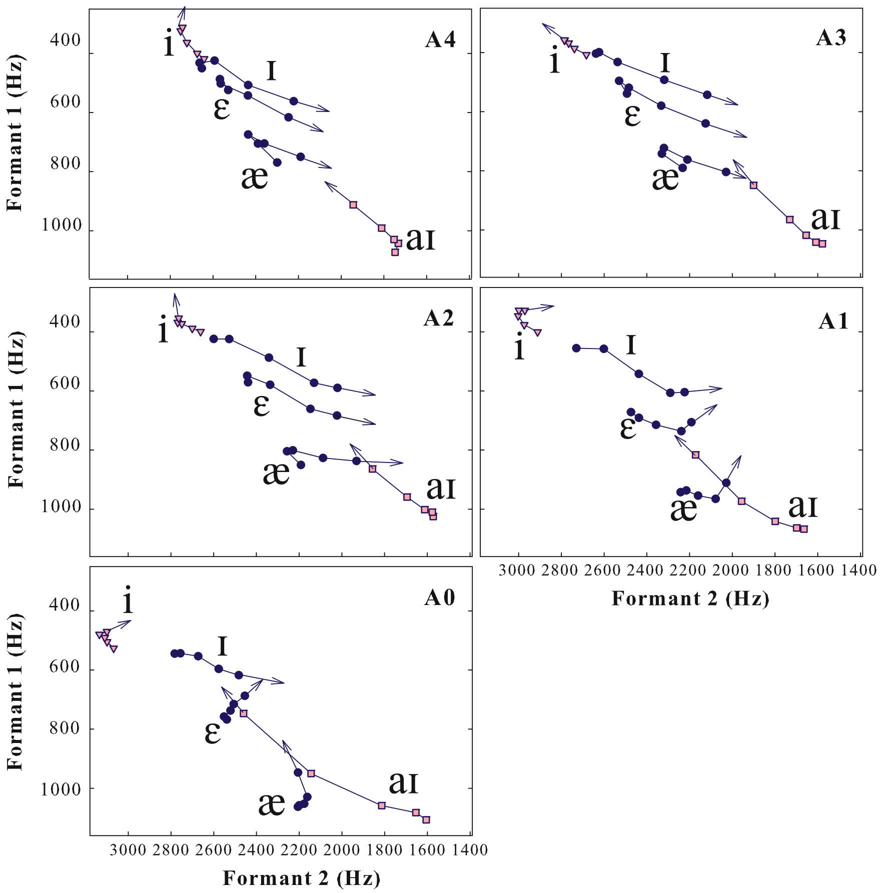

Fig. 6 Means of F1 and F2 for subset of vowels (produced in citation form *hVd* words) in the Southern Vowel Shift in western North Carolina

The complexity of changes to the vowel subsystem affected by the SVS can only be appreciated when we examine the cross-generational changes in VISC. Clearly, formant dynamics change across generations of North Carolina speakers and the vowel system of the youngest ones does not look like the system of speakers from older generations. These data provide new support for the view that the systemic reorganization occurs gradually, generation by generation, and particular vowels change their acoustic characteristics in terms of both formant frequencies and the extent of spectral change.

6 Cross-Dialectal Variation in Formant Dynamics of /æ/

The final set of observations about the dialect-specific use of formant dynamics pertains to vowel /æ/. The case of /æ/ is of interest to us for the following reason. We found a progressive lowering and backing of the Ohio variant with each younger generation which also corresponds to a reduction in its formant movement. Lowering and backing was also found in Wisconsin /æ/ whose formant movement (at least in the emphatic variant) did not generally change across generations. How do these patterns compare to the third variant of /æ/, that spoken in the South, which is generally considered to demonstrate "Southern raising?" In our data, southern /æ/ also lowers with each younger generation following the path of the two other dialects. If it is the case that /æ/ lowers cross-dialectally in accord with Principle II, what would be the corresponding change in the extent of VISC for each of the three regional variants?

The plots in Fig. 7 show the three emphatic variants of /æ/, one for each dialect, across all five age groups redrawn from Figs. 2, 4, and 5. As can be seen, there are positional differences among the three variants of /æ/ which also show three distinct patterns of formant movement. Wisconsin /æ/ is the most raised and the most fronted, Ohio /æ/ is the lowest, and North Carolina /æ/ is positioned in between these two. All three dialect variants are raised in the production of the oldest generation and descend in the acoustic space in a steady fashion with each younger generation. Of particular interest are the cross-generational changes in the nature and amount of formant movement.

All three variants are heavily diphthongized in A4 speakers and we can see three distinct patterns of /æ/-production: the Wisconsin variant has its nucleus raised due to the NCS, the North Carolina variant indicates the Southern breaking (notice the difference in the directionality of formant movement between the northern and the southern variant) and the Ohio /æ/ does not show any raising of its nucleus (and no influence of the NCS) although its direction of movement is as in the Wisconsin variant. This general pattern is maintained in the productions of the three older generations and then changes rather abruptly in young adults (A1). In particular, Wisconsin /æ/ has a smaller extent of formant movement and its nucleus is not raised as much as in the previous generations, North Carolina /æ/ changed its directionality so that the Southern breaking is no longer present, and Ohio /æ/ shows a spectral change in F1 rather than F2, which is opposite to what we see in A4, A3 and to some extent in A2 speakers. Finally, while Wisconsin children's /æ/ still shows the NCS influence, the North Carolina and Ohio variants changed greatly so that the only spectral change is in F1.

Clearly, these three dialect variants differ in the nature and the extent of VISC although the common trend for /æ/ is to lower in the acoustic space with each younger generation in accord with Principle II. This also includes the southern variant which is assumed to be raised in the South. It is unclear why emphatic variants in both Ohio and North Carolina change their dynamics across generations while the Wisconsin vowel does not. The principles of chain shifting seem to have

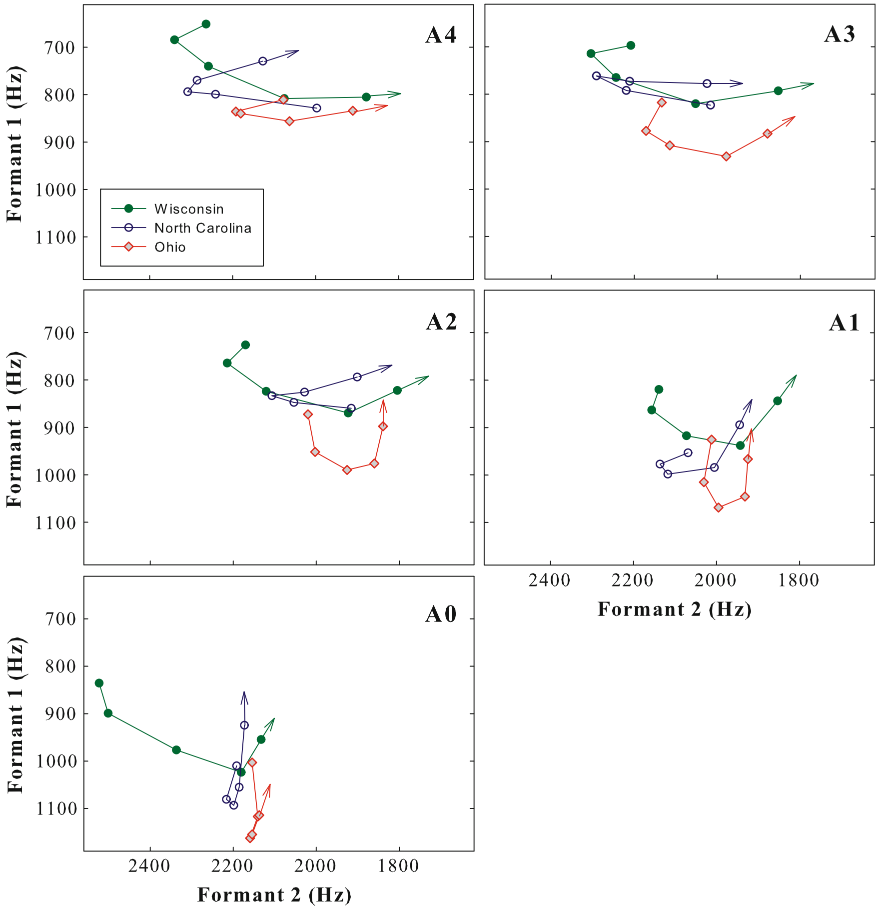

Fig. 7 Means of F1 and F2 for the emphatic variants of /æ/ across dialects and age groups

reached their explanatory limit in this case and a deeper study of cross-generational variation in VISC in each of the three vowel systems is most likely to increase our understanding of such patterns.

7 Summary of Regional Variation in Formant Dynamics

The cross-dialectal and cross-generational data presented in this chapter underscore the richness of acoustic information between the onset and offset of a vowel. Changes in the shapes of the formant trajectories during the production of a vowel are precisely what differentiates dialects and generations of speakers across regions in the United States. Vowels in American English change their acoustic characteristics over time as new generations of speakers are born and acquire vowel systems from available linguistic input. While sociolinguists have focused on the directionality of vowel movement in the acoustic space (as determined by the formant values at the syllabic nucleus), the variation in dynamic vowel characteristics has not been addressed so far as a dialect feature.

The use of multi-point measurement of the whole trajectory, as in the present approach, helps us realize that even the most monophthongal vowel is rarely static, i.e., its formant shapes do not resemble straight lines which would be plotted in a F1 by F2 plane as a single data point. Rather, spectral change includes at least minimal variation in either F1 or F2 or a combination of both. Some additional variation in VISC comes from phonetic sources such as prosodic effects which may affect the shape of the whole trajectory as shown in the present data. Obviously, consonantal environment introduces the most predictable type of changes. However, there is also another powerful source of variation in VISC, that of the dialect-specific use of the dynamic information in a vowel. A good example here is the cross-dialectal variation in the trajectories of /æ/ whose shape and direction of movement may signal Southern breaking, reflect the first link in the NCS, or follow yet other patterns, as in central Ohio. What we also find across the dialects is that the extent of VISC changes cross-generationally. This change is most likely related to dialect-specific chain shifts or other vowel changes which we find in regional vowel systems. Although the principles of chain shifting have been formulated on the basis of positional changes of vowel nucleus, there seems to be a correspondence between a change in the extent of VISC and the positional change of a vowel related to the reorganization of a given subsystem.

As hypothesized earlier (Jacewicz et al. 2006), phonetic stress may play a role in chain shifting. Emphatic vowels have a potential to lead the vowel change due to their enhanced acoustic characteristics on which children may focus in the process of cross-generational sound transmission. The present data suggest that emphatic variants indeed "pull" vowels in specific directions over time, although this possibility must be explored further in future studies. The present data clearly show a substantial difference in TL values between the emphatic and nonemphatic variants. For all vowels and all age groups, the emphatic vowels have longer TLs than nonemphatic vowels and the only exception is the vowel /ɛ/ in Wisconsin children. Although longer TLs for the emphatic variants were expected, the cross-generational changes in the extent of VISC in these emphatic variants were not. As already pointed out, the present data suggest a type of coordinated systemic changes associated with reorganization of a vowel subsystem which include both positional changes and changes in the extent of VISC, which are most readily observed in the emphatic variants. However, more vowels (including back vowels) need to be studied and more work needs to be done to understand this complex relationship.

8 Dialectal Spectro-Temporal Variation and Vowel Perception

We used the TL measure in this chapter to compare the total trajectory change (restricted to the 5-point measurement) across dialects and ages. Although, indisputably, a greater number of measurement points can produce a more accurate, veridical representation of the formant trajectory, a 5-point measurement system can still produce a good estimate of the actual trajectory length and a reasonable characterization of the trajectory shape. However, the TL measure fails to account for the change in the directionally of movement. For example, while Wisconsin and North Carolina /æ/ in A4 group have comparable TL values (606 and 576 Hz, respectively), their formant trajectories move in opposite directions. A similar problem occurs when the angle of formant movement changes in the course of vowel duration.

Also, while the TL measure provides information about the cumulative size of the formant changes, it fails to account for the speed of these changes as it has no true temporal component. However, there may be important dynamic differences across dialects and speaker age that relate to how quickly (or slowly) these spectral changes are made. The spectral rate of change measure, also included in Fox and Jacewicz (2009), has been shown to be quite effective in addressing the cross-dialectal variation in VISC when restricted to a single age group. However, caution and further modeling is needed when applying the spectral rate of change measure to cross-generational data because vowel inherent duration may be confounded by several factors including articulation rate, aging effects and dialect differences.

As it turns out, vowel duration is also subject to regional variation and systematic differences in vowel duration have been found for the three dialect regions studied here (Jacewicz et al. 2007). Moreover, the dialect differences in articulation rate also proved to be significant. In particular, articulation rate (excluding pauses) in the Wisconsin speech is faster than in North Carolina speech for both young and older adults apart from the aging effects (Jacewicz et al. 2009a). Given that vowels are significantly longer in North Carolina speech and articulation rate is slower, one would assume a straightforward relationship between these two. That this is not the case and the slower articulation rate does not imply a change "across the board" in temporal properties of segments has been shown in a study of stop closure voicing for these two dialects (Jacewicz et al. 2009b). Namely, the stop closure duration was found to be longer in Wisconsin and not in North Carolina speech. This would suggest that vowel duration is dialect-specific as is the nature of formant movement along with positional relations among the vowels. Further modeling is needed to relate vowel-inherent duration to temporal variation as a function of dialectal and cross-generational changes in vowels.

The set of data presented in this chapter lends support to the conclusion reached earlier by Cox (1999) who studied vowel change in Australian English. She points to the

> Changeable nature of language and the fact that specifications of formant structures are only valid for a particular dialect at a particular time in history. The systemic nature of vowel change is clearly documented as well as the close interdependent relationships between the monophthongs and diphthongs. Change in one class can be seen to affect the other in a parallel fashion in this dialect of English (p. 20).

In our view, the mechanism of language change utilizes the variation in VISC in ways which are not yet well understood. However, many questions arise as to the perceptual relevance of this type of synchronic variation in formant dynamics to vowel identification. Although the present apparent time data show cross-generational changes in vowel characteristics, we need to bear in mind that the speakers used in this study have been born and lived in the same speech community most of their lives. This suggests that both adults and children, despite differences in their respective productions, have been exposed as listeners to diverse dynamic cues in vowels including a variety of shapes of their formant trajectories. How will this experience with dialect-specific features affect their ability to make vowel identification decisions? Will spectral information between vowel onset and offset be largely ignored, will it be helpful or will it be critical in identifying specific vowels? How sensitive are these listeners to cross-generational changes in VISC and do they perceive such variation at all?

A set of related questions can be asked about listeners growing up in one dialect area who have never been in contact with features of another dialect. Will formant dynamics of vowels from another dialect influence their perceptual response in a listening task or will the attunement to their own dialect guide their identification choices? What will be the confusion pattern and what will it tell us about listeners' use of spectral dynamics in vowels from their non-native dialect? Will they manifest sensitivity to cross-generational variation in VISC in their non-native dialect? These and other questions await answers in future research.

Acknowledgments This work was supported by the grant R01 DC006871 from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health. We thank Joseph Salmons for his contributions to this research and his comments on this chapter. Special thanks go to the personnel of the Speech Perception and Acoustics Labs at Ohio State as well as our collaborators at the University of Wisconsin-Madison (Dilara Tepeli) and Western Carolina University (Janaye Houghton) for their help with recordings, data collection and analysis. We would also like to thank Peter Assmann, Geoffrey Stewart Morrison and Michael Kiefte for their comments on earlier versions of this chapter.

References

Agwuele, A., Sussman, H.M., Lindblom, B.: The effects of speaking rate on consonant vowel coarticulation. Phonetica **65**, 194–209 (2008). doi:10.1159/000192792

Andruski, J.E., Nearey, T.M.: On the sufficiency of compound target specification of isolated vowels in /bVb/ syllables. J. Acoust. Soc. Am. **91**, 390–410 (1992). doi:10.1121/1.402781

Assmann, P.F., Nearey, T.M., Bharadwaj, S.V.: Developmental study of vowel-inherent spectral change. J. Acoust. Soc. Am. **125**, 2696 (2009)

Bailey, G., Wikle, T., Tillery, J., Sand, L.: The apparent time construct. Lang. Var. Change **3**, 241–264 (1991)

Cassidy, F.G., Hall, J.H.: The Dictionary of American Regional English. Harvard University Press, Cambridge (1985)

Chambers, J.K.: Sociolinguistic Theory, 2nd edn. Blackwell, Oxford (2003)

Cox, F.: Vowel change in Australian English. Phonetica **56**, 1–27 (1999). doi:10.1159/000028438

Fox, R.A., Jacewicz, E.: Cross-dialectal variation in formant dynamics of American English vowels. J. Acoust. Soc. Am. **126**, 2603–2618 (2009). doi:10.1121/1.3212921

Gordon, M.: Small-town Values and Big-city Vowels: A Study of the Northern Cities Shift in Michigan. Duke University Press, Durham (2001)

Hillenbrand, J.M., Getty, L.A., Clark, M.J., Wheeler, K.: Acoustic characteristics of American English vowels. J. Acoust. Soc. Am. **97**, 3099–3111 (1995). doi:10.1121/1.411872

Jacewicz, E., Fox, R.A., Salmons, J.: Prosodic prominence effects on vowels in chain shifts. Lang. Var. Change **18**, 285–316 (2006)

Jacewicz, E., Fox, R.A., Salmons, J.: Vowel duration in three American English dialects. Am. Speech **82**, 367–385 (2007). doi:10.1215/00031283-2007-024

Jacewicz, E., Fox, R.A., O'Neill, C., Salmons, J.: Articulation rate across dialect, age, and gender. Lang. Var. Change **21**, 233–256 (2009a). doi:10.1017/S0954394509990093

Jacewicz, E., Fox, R.A., Lyle, S.: Variation in stop consonant voicing in two regional varieties of American English. J. Int. Phonetic Assoc **39**, 313–334 (2009b). doi:10.1017/S0025100309990156

Kurath, H.: A Word Geography of the Eastern United States. University of Michigan Press, Ann Arbor (1949)

Kurath, H., McDavid, R.I.: The Pronunciation of English in the Atlantic States. University of Michigan Press, Ann Arbor (1961)

Labov, W.: Principles of Linguistic Change. 1: Internal Factors. Blackwell, Oxford (1994)

Labov, W: Principles of Linguistic Change. 2: Social Factors. Blackwell, Oxford (2001)

Labov, W., Ash, S., Boberg, C.: Atlas of North American English: Phonetics, Phonology, and Sound Change. Mouton de Gruyter, Berlin (2006)

Labov, W., Jaeger, M., Steiner, R.: A Quantitative Study of Sound Change in Progress. U.S. Regional Survey, Philadelphia (1972)

Lee, S., Potamianos, A., Narayanan, S.: Acoustics of children's speech: Developmental changes of temporal and spectral parameters. J. Acoust. Soc. Am. **105**, 1455–1468 (1999). doi:10.1121/1.426686

Moon, S.-J., Lindblom, B.: Interaction between duration, context, and speaking style in English stressed vowels. J. Acoust. Soc. Am. **96**, 40–55 (1994). doi:10.1121/1.410492

Nearey, T.M., Assmann, P.F.: Modeling the role of vowel inherent spectral change in vowel identification. J. Acoust. Soc. Am. **80**, 1297–1308 (1986). doi:10.1121/1.394433

Peterson, G., Barney, H.L.: Control methods used in a study of the vowels. J. Acoust. Soc. Am. **24**, 175–184 (1952). doi:10.1121/1.1906875

Sievers, E.: Grundzüge der Phonetik. Breitkopf and Härtel, Leipzig (1st Edn., 1876) (1881)

Sledd, J.: Breaking, umlaut, and the southern drawl. Language **42**, 18–41 (1966)

Stockwell, R.: Perseverance in the English vowel shift. In: Fisiak, J. (ed.) Recent Developments in Historical Phonology, pp. 337–348. Mouton, The Hague (1978)

Thomas, E.R.: An Acoustic Analysis of Vowel Variation in New World English. Publication of the American Dialect Society, p. 85 (2001)

Watson, C., Harrington, J.: Acoustic evidence of dynamic formant trajectories in Australian English vowels. J. Acoust. Soc. Am. **106**, 458–468 (1999). doi:10.1121/1.427069

Wolfram, W., Schilling-Estes, N.: American English. Blackwell, Malden (2006)