Settlement patterns and the eastern boundary of the Northern Cities Shift

Aaron J. Dinkin*

Department of Linguistics, University of Toronto, Toronto, ON, Canada

This paper examines the dialectology of eastern New York State. Data are considered from twelve cities and villages bridging the gap between the Inland North dialect region (home to the Northern Cities Shift [NCS]) and the Western New England region. Communities are classified as belonging to the Inland North "core," the Inland North "fringe," or a non–Inland North region. The settlement history of these communities is used to explain the boundaries between the dialect regions; presence of the NCS is found to correlate well with heavy migration from southwestern New England early in a community's history. Looking in detail at the different distributions of the individual sound changes involved in the NCS across the dialect regions established in this paper makes it possible to posit a reconstructed early history for the NCS and make hypotheses about the phonological character of the different sound changes.

1. Introduction and background

1.1. Nature of dialect boundaries

The most thorough study to date of the regional dialect variation in pronunciation of the United States and Canada is the *Atlas of North American English* (*ANAE*: Labov et al., 2006), which divides English-speaking North America into approximately ten major dialect regions (and certain transitional areas) based on the patterns of phonetic and phonological change that are characteristic of each region. *ANAE*'s data is derived, for the most part, from telephone interviews conducted with speakers in urbanized areas with populations greater than fifty thousand people.
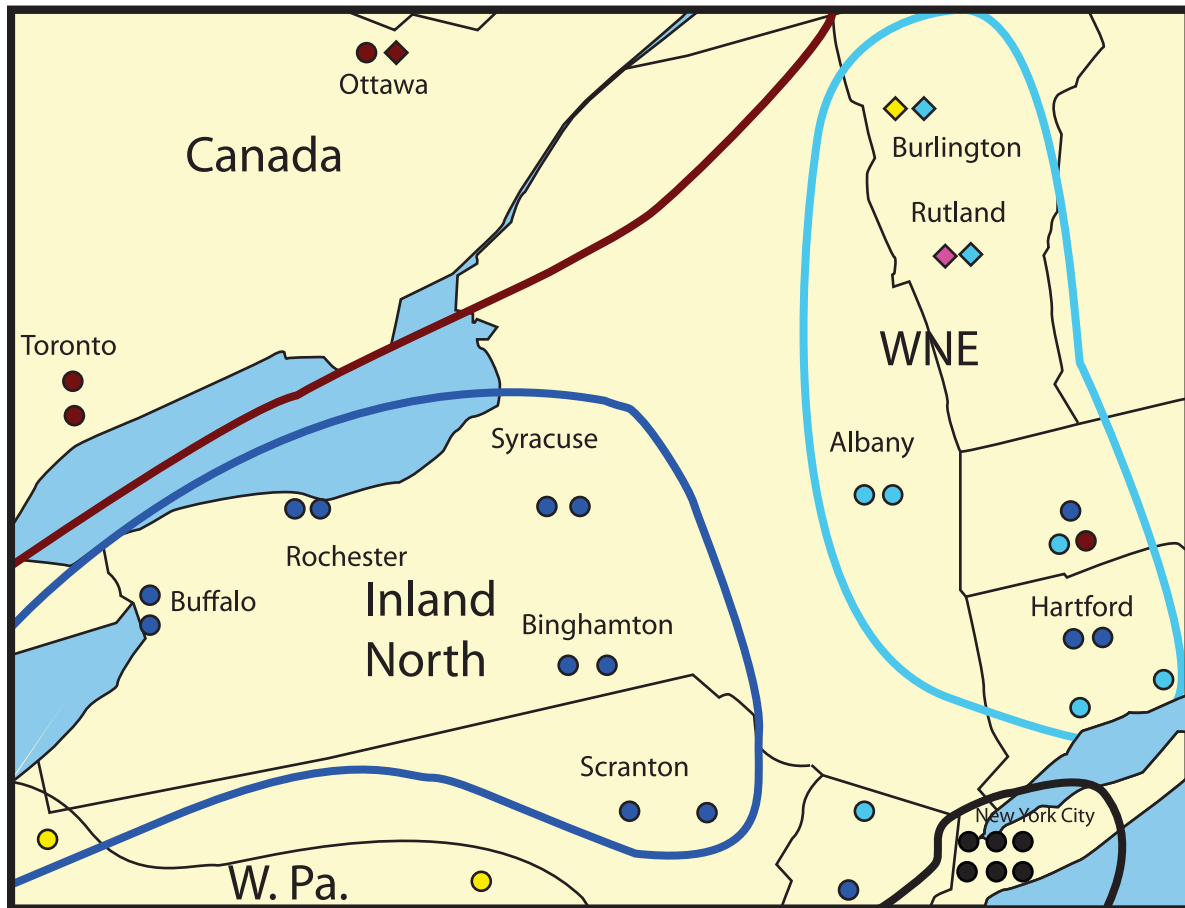
Since the data is taken from speakers in urbanized areas, each of *ANAE*'s dialect regions is defined really only in terms of the major cities it contains, and the boundaries between the regions lie in most cases in less densely populated areas between the cities. This means that—except in the very rare case that two adjacent communities of fifty thousand or more people are located on opposite sides of a dialect boundary, as Windsor, Ontario and Detroit, Michigan are—*ANAE* provides no information about the extent to which speakers in communities near the boundary share the linguistic features of speakers in the major cities, nor even where in the territory between the cities the linguistic boundary actually lies. Thus more detailed studies of communities between the cities that define the dialect regions are necessary in order to determine the nature and location of the boundaries.

There are at least four general possibilities for the nature of boundaries between dialect regions and the linguistic status of communities close to the boundaries:

1. There is a *sharp* boundary line between dialect regions. Communities on each side of the boundary line display all the linguistic features on whose basis the region is defined, to the same extent that communities distant from the border do. *ANAE* itself finds that this is the case between the Inland North and Canadian regions at Detroit and Windsor, and Johnson (2007) finds that this was at one time the case between the dialect regions of Rhode Island and eastern Massachusetts.

2. There is a *gradual* boundary line, with regional features fading out near the boundary. Communities close to the boundary exhibit the characteristic features of one region or the other to a weaker or reduced degree: Either the sound changes are less advanced, or they are present in a smaller fraction of the population than in the core of the dialect region, or both; but each community can still be classified as belonging to one of the two regions.

3. The dialect regions *overlap*, and there are communities in which the characteristic features of both dialects are found—either there are speakers who exhibit the characteristic linguistic features of both regions, or some speakers show the linguistic pattern of one region and some show the other. Bigham (2007) suggests that the area in southern Illinois between the South and the so-called St. Louis Corridor may be such a region.

4. The two dialect regions have a *null* boundary—in other words, they do not meet, and there is some set of communities between the dialect regions that is more linguistically conservative or unmarked

*Address for correspondence: Aaron J. Dinkin, Department of Linguistics, University of Toronto, Sidney Smith Hall, 4th floor, 100 St. George St., Toronto, Ontario, Canada M5S 3G3.
(Email: ajd@post.harvard.edu)

Map 1. New York State, as portrayed in *ANAE*. Map from Dinkin & Labov (2007); used with permission.

and does not exhibit the characteristic features of either region, or constitute a third and perhaps previously undetected dialect region. In a case such as this, the existence of a boundary at all between the two original regions of interest was merely an illusion caused by the lack of data in the intervening area.

Classifying the boundaries between dialect regions can be an important first step to explaining how and under what circumstances linguistic changes spread, or fail to spread, between communities. For instance, Boberg (2000) takes the sharpness of the boundary between Detroit and Windsor as a jumping-off point for arguing that an international boundary can act as a strong obstacle to the spread of phonetic change even in the presence of heavy cross-border communication.

1.2. New York State and the Northern Cities Shift

The state of New York is of particular interest in the study of the nature of dialect boundaries for several reasons. As we see from Map 1, parts of New York

State are within or adjacent to at least four of the dialect regions defined by *ANAE*: Buffalo, Rochester, Syracuse, and Binghamton, in the western and central part of the state, form the easternmost extent of the Inland North region, which extends from here as far west as eastern Wisconsin. Albany is classified as part of the Western New England dialect region; New York City has a dialect region basically to itself; and the Canadian dialect region is adjacent to the far north of New York State, although none of the New York communities sampled by *ANAE* fall within it. Thus the eastern part of New York, between New York City, Albany, Binghamton, Syracuse, and the Canadian border, constitutes an area approximately 150 miles from east to west and three hundred miles from north to south, within which four distinct dialect regions must come together in some way, but which *ANAE* does not sample at all. This is a remarkably large number of dialect boundaries with remarkably little existing data on the points of contact between them. The eastern edge of the Inland North could lie anywhere between Syracuse and Binghamton, on the one hand, and Albany to the east or Ottawa to the
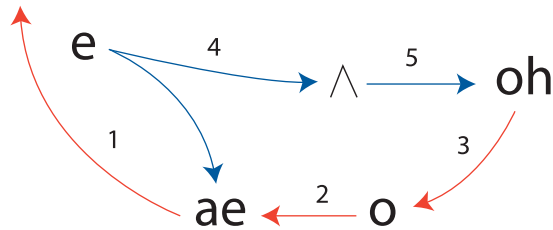
Figure 1. The Northern Cities Shift.

north, on the other hand, and there is no way of knowing what the dialectological status of the communities near the boundary is.

The eastern edge of the Inland North is a topic of interest because of conflicting characterizations in the literature of the relationship between the Inland North and Western New England dialect regions. The Inland North itself is defined as the region subject to the Northern Cities Shift (NCS), a vowel chain shift demonstrated in Figure 1. The key features of the NCS are the raising and fronting of /æ/[1] as in *trap*, the fronting of /o/ as in *lot*, the lowering and backing[2] of /e/ as in *dress*, and the backing of /ʌ/ as in *strut*. Also involved in the shift are the lowering of /oh/ as in *thought* and backing of /i/ as in *kit*; however, these are not used in *ANAE* as diagnostic features for identifying the NCS's geographic distribution, and the analysis in this paper will focus on the first four.

Although all sources agree that Western New England is an essential part of the Inland North's history, predictions differ on whether a boundary exists between them in the present day and, if so, what the nature of that boundary will be. On the basis of lexical data, Kurath (1949) defines a "Hudson Valley" dialect region located between the Inland North and Southwestern New England. However, examining the maps of Kurath & McDavid (1961) fails to reveal any phonological difference between the Hudson Valley and Southwestern New England.[3] Albany, the only city in the *ANAE* data that might be within Kurath's Hudson Valley region, is grouped by *ANAE* with Southwestern New England. If a boundary between the Inland North and Western New England exists, it may pass through Kurath's Hudson Valley, or coincide with one of the Hudson Valley's boundaries. This paper takes *ANAE* as its point of departure and focuses on phonetic and phonological features; lexical variables were not systematically investigated in the study reported herein. But the question of the present-day status of a Hudson Valley dialect area will be raised in a later section.

On the basis of the *ANAE* data, Boberg (2001) divides the Western New England region into Northwestern New England (western Vermont, exhibiting the *caught-cot* merger) and Southwestern New England

(based in the lower Connecticut River valley), with a gradual boundary between them. He concludes that Southwestern New England is a "subtype of the Inland North," (2001:28) differing from the Inland North proper only in phonetic detail, rather than in phonological structure—specifically, "the relative advancement of the Northern Cities Shift (28)." In other words, in Boberg's analysis, Southwestern New England is essentially an eastern extension of the Inland North region, which is just as open to the NCS as communities in the Inland North are; it just happens not to have undergone the shift *yet*. If this is the case, we would not expect to see a sharp discontinuity between the Inland North and Southwestern New England. Rather, if the only difference between them is that the NCS is more advanced in the Inland North proper and less advanced in Southwestern New England, we might expect to find NCS features with an intermediate degree of advancement in the area in the intermediate area between Syracuse and Binghamton on the one hand and Connecticut and Albany on the other.

Boberg is one of *ANAE*'s authors, and the text of *ANAE* echoes the point of his (2001) paper in saying that "the basic configuration underlying the NCS can be found among Western New England speakers" (Labov, Ash, & Boberg, 2006:214). It goes on, however, to present a different interpretation of the phonological status of the Inland North, arguing that the cause of the NCS depended upon the unique settlement history of western and central New York. The argument hinges on the fact that the largest NCS cities in Upstate New York—Buffalo, Rochester, and Syracuse—all lie along the Erie Canal, whose construction spurred the population growth of the region:

> The native-born settlers moving into New York State came from a variety of dialect areas in New England, including Maine, New Hampshire, Providence, and western Connecticut. In addition, the great expansion of New York City after the [Erie] Canal was completed ensured a flow of workers, passengers, and entrepreneurs from outside of New England, up the Hudson River and westward to Buffalo. […] These settlers would have a variety of different and incompatible short-*a* systems: the nasal system of Eastern New England, the continuous nasal pattern of Western New England, the broad-*a* pattern of Boston, and the short-*a* split of New York City. The end result in New York State was none of these, but the general raised short-*a* pattern of the NCS. (*ANAE*: 214)

In other words, as Upstate New York's settlement was driven by the construction of the Erie Canal in the 1820s, the combination of multiple incompatible

phonological treatments of /æ/ from different regional origins gave rise to the NCS's distinctive raising of /æ/. This account makes different predictions about the relationship between the Inland North and Western New England than Boberg (2001) does. Under the *ANAE* interpretation, the NCS is phonologically distinct on a qualitative level from the vowel systems of the dialect regions that contributed to the region's settlement, and it could not have arisen in one of these regions alone. If this is the case, we would expect to see a sharp boundary between the Inland North and the surrounding regions: Communities that share the distinctive Inland North settlement history, driven by the Erie Canal, will share the Inland North phonology and undergo the NCS; communities with a different early settlement history would not be subject to the NCS, and in principle such communities could be arbitrarily close to each other.

So, by identifying and examining the linguistic status of communities near the edge of the Inland North—if such an edge exists—we can attempt to determine the nature of the boundary and the phonological relationship between the NCS and Southwestern New England. A gradual transition would suggest that, as Boberg (2001) argues, Southwestern New England's vowels are phonologically no different from the NCS, and that the Hudson Valley should not be distinguished as a separate dialect region; a sharp boundary would suggest that the NCS constitutes a substantive phonological difference between the Inland North and whatever region is adjacent.

2. Methodology

This paper fills in some of the gaps between *ANAE*'s Inland North, Canada, and Western New England dialect regions with data from twelve cities and towns in the eastern half of New York State: Amsterdam, Canton, Cooperstown, Glens Falls, Gloversville, Ogdensburg, Oneonta, Plattsburgh, Poughkeepsie,

Map 2. New York State. Communities sampled in this study are marked with stars. The area marked in light green is the sparsely populated Adirondack State Park. Map based on Dinkin (2009); used with permission.

Map 3. The counties of New York State. Map produced by the U.S. Census Bureau.

Sidney, Utica, and Watertown. Their locations relative to the cities sampled in *ANAE* are visible in Map 2. Map 3 displays New York's county boundaries; Table 1 lists which county each sampled community is in.

In each of these locations, between five and ten interviews were conducted with natives of the community in the summers of 2006–2008, mostly following the Short Sociolinguistic Encounter protocol described by Ash (2002). These are interviews of ten to twenty-five minutes in length for which the researcher recruits subjects by approaching them in publicly accessible places such as parks, swimming pools, cafés, and shops. In the interviews conducted in this study, conversation topics focused on everyday life in the city or village, the local economy, contact between the city or village and surrounding areas, and the subject's own recreational and travel habits. Interviews ended with a few short formal data-elicitation methods, including a written list of approximately fifty words. Limited demographic information (age, occupation, education, residential history) and no personally identifying information was requested. Little to no attempt was made to

balance the sample by gender, age, or socioeconomic class; subjects were recruited strictly by availability.

The Short Sociolinguistic Encounters are supplemented by a smaller set of telephone interviews carried out in a few of the communities in the summer of 2006 and the spring and early autumn of 2008. These followed the Telsur protocol used in *ANAE*: a segment of the telephone directory for a given city or village was randomly selected; those names in each segment that seemed characteristic of the predominant ethnic groups in the community according to the United States Census were called; and two interviews were conducted with the first two speakers called in this manner who were natives of the community and were willing to participate. These interviews were approximately thirty minutes long, and included about ten minutes of spontaneous conversation on the same topics as the Short Sociolinguistic Encounters described above, and twenty minutes of formal elicitation of specific words and minimal pairs.

In a small number of cases, interview subjects were contacted in advance and interviewed at appointed

Community	County	2010 census population	# Interviews
Amsterdam	Montgomery	18620	5+2 telephone
Canton	St. Lawrence	6314	7+2 telephone
Cooperstown	Otsego	1852	5+4 telephone
Glens Falls	Warren	14700	7
Gloversville	Fulton	15665	7+2 telephone
Ogdensburg	St. Lawrence	11128	7+2 telephone
Oneonta	Otsego	13901	9
Plattsburgh	Clinton	19989	7
Poughkeepsie	Dutchess	32736	7
Sidney	Delaware	3900	6+2 telephone
Utica	Oneida	62235	7
Watertown	Jefferson	27023	10

times, rather than approached through the Short Sociolinguistic Encounter method or through random telephone-number selection. In the dataset reported in this paper, these include two telephone interviews in Cooperstown as well as one in-person interview in Poughkeepsie and three in Sidney. Apart from the manner of recruitment of subjects, these preplanned interviews followed the same protocols as described above.

In most communities sampled, the speakers interviewed are distributed relatively evenly across a fairly broad range of ages. This reassures us that the characterization of each community that emerges from the data will not be merely a snapshot of one stage in change in progress, and means that apparent-time analysis of individual communities may be possible. The chief exception to this is Utica, in which the seven speakers interviewed include six born between 1979 and 1989 and only one older outlier born in 1942. Full details on the age and gender distribution of the sample can be found in Table 2.

Vowel charts were plotted for all European-American[4] speakers (and one Turkish-American) who said that they had lived in the community in which they were interviewed from before starting school through adolescence (although many of them had moved away for shorter or longer periods of time after high school). Table 1 lists the number of in-person and telephone interviews from each community analyzed in this paper.

For each analyzed speaker, between about 400 and 600 stressed vowel tokens were measured whenever possible, including all words elicited through formal methods. In the case of a few shorter interviews with more reticent speakers, the number of measurable vowel tokens was between 250 and 400; for a single speaker (Jake V.[5] from Gloversville), only 190 tokens were measurable. However, with the possible

exception of Jake V., even the speakers with the fewest measurable tokens are sampled at least as thoroughly as most speakers in *ANAE*, in which the mean number of tokens measured per speaker was 305.

Formant measurements were extracted according to the methodology described in *ANAE*: F1 and F2 were measured in Praat at a single point in each vowel token selected by hand as representative of the central tendency of the vowel nucleus. (Offglides of diphthongs were not measured.) Mean F1 and F2 values for each vowel phoneme were computed in Plotnik 8, which ignores vowel tokens before sonorants and after obstruent + liquid clusters in calculating means. For inter-speaker comparability, each speaker's vowel measurements were log-mean normalized in Plotnik using the same group norm used in *ANAE*.

3. Results: categorical NCS criteria

3.1. Overall findings

Great variation was found across the full sample of ninety-eight speakers with respect to the presence or absence of the NCS. The most advanced NCS was found in the vowel system of Janet B., a sixty-four-year-old bookstore clerk from Utica, depicted in Figure 2. Janet's /æ/ is extremely high and front, with only three tokens lower than the midline of her vowel space; her mean /e/ is so back, and her mean /o/ so front, that both line up along the center line; and her /ʌ/ is far to the back of the vowel space. Janet B.'s pronunciation of the sentence *A bed, you have your box spring and your mattress* illustrates her low /e/ in *bed*[6], fronted /o/ in *box*, and raised /æ/ in *mattress*, highlighted in Figure 2.
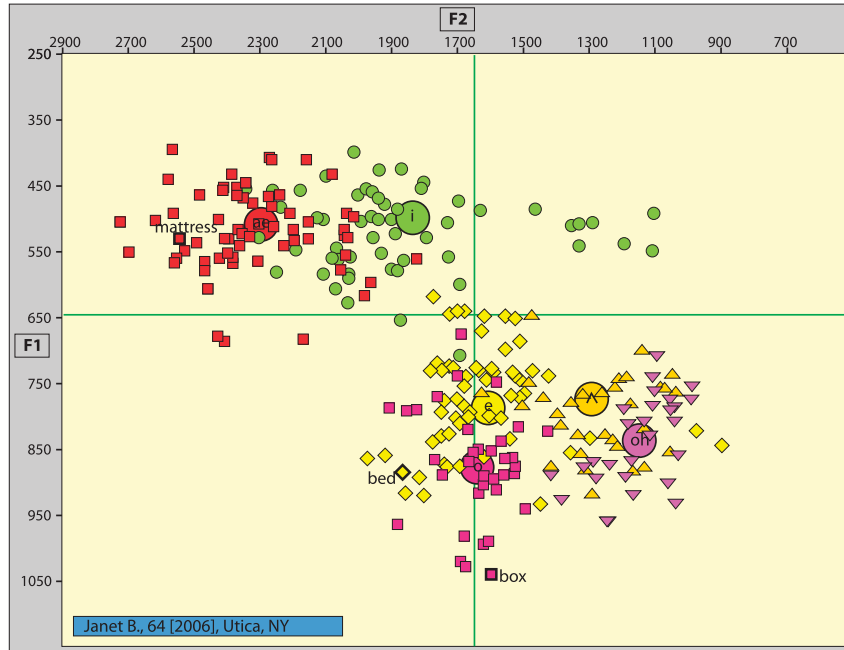
By contrast, Emily R., a twenty-one-year-old college student from Cooperstown, shows no NCS at all: /æ/ remains in low front position, not even on average as far front as /e/; /o/ is some distance back of center; and /e/ and /ʌ/ are about the same distance front and back of center, respectively. Her vowel system is shown in Figure 3, demonstrating her low /æ/ in *hat*, front /e/ in *deck*, and back /o/ in *hot*.

Labov (2007) uses a set of five criteria based on the mean normalized formant values of the NCS vowel phonemes to quantify speakers' degree of participation in the NCS. These criteria are as follows:

- **UD criterion**: /o/ is fronter than /ʌ/.
- **ED criterion**: /e/ is less than 375 Hz fronter than /o/ (i.e. F2 /e/ - F2 /o/ < 375 Hz).
- **EQ criterion**: /æ/ is higher and fronter than /e/.
- **AE1 criterion**: /æ/ is higher than 700 Hz (i.e. F1/æ/ < 700 Hz).
- **O2 criterion**: /o/ is fronter than 1,500 Hz (i.e. F2/o/ > 1,500 Hz).

Table 2. *The speakers described in this paper, classified by age and gender*

	Before 1943	1943–1957	1958–1972	1973–1986	After 1986	Mean year of birth
Amsterdam						
Female		1		3		1970
Male		2			1	
Canton						
Female	1		2		2	1973
Male		1		1	2	
Cooperstown						
Female		3	1	1	3	1967
Male	1					
Glens Falls						
Female				1	1	1975
Male	1		1	2	1	
Gloversville						
Female		2			1	1961
Male	2	1	1		2	
Ogdensburg						
Female	1		1	3	2	1972
Male			1	1		
Oneonta						
Female		1	1	1	2	1974
Male		1	1	1	1	
Plattsburgh						
Female			1	1		1972
Male	1	1		1	2	
Poughkeepsie						
Female		1		1	1	1966
Male	1	1	2			
Sidney						
Female		2	1		2	1964
Male		1	1	1		
Utica						
Female	1			1	2	1979
Male				2	1	
Watertown						
Female			1	3	1	1972
Male		1	4			
Total	9	18	20	24	27	1970

Figure 2. The vowel system of Janet B., a sixty-four-year-old bookstore clerk from Utica.
Red=/æ/; orange=/ʌ/; yellow=/e/; green=/i/; light purple=/oh/; magenta=/o/.
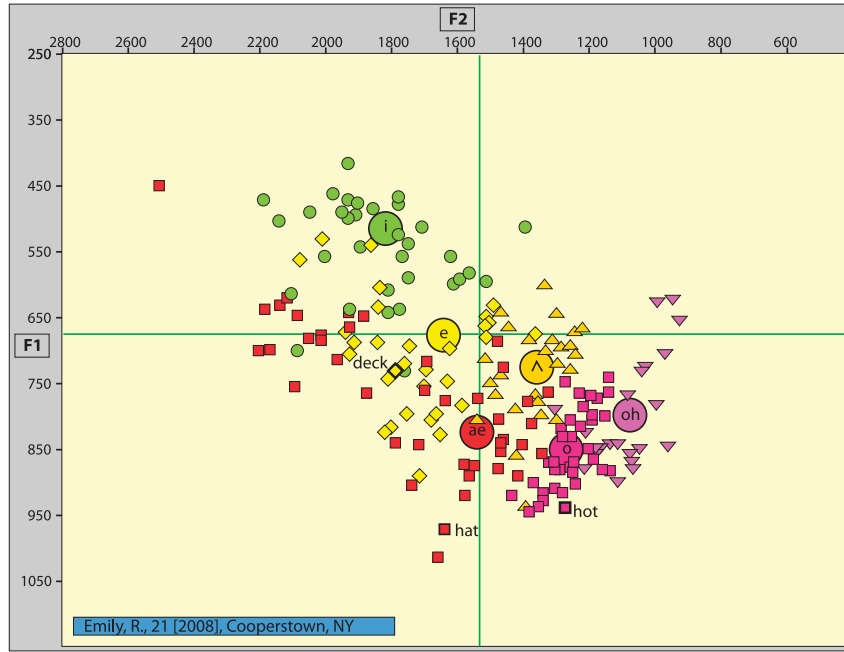The "Janet" audio files are associated with this figure.

These criteria are not entirely unproblematic. They are all categorical summaries of continuously varying quantities—observing that two speakers have /o/ backer than 1,500 Hz obscures the possibility that they might still differ greatly in how front their /o/ phonemes actually are. Several of the criteria compare two phonemes rather than measuring a single one, meaning that two speakers or regions that satisfy a given criterion might satisfy it as a result of different sound changes. And several of the phonemes are involved in multiple criteria, meaning that a speaker with fronted /o/, for instance, might satisfy three criteria on the basis of that one phoneme alone.[7] However, despite these disadvantages, most of these criteria, as used in *ANAE*, succeed in forming clear and consistent dialect boundaries; and as we will see, they do a good job as a first pass at distinguishing Inland North from non–Inland North regions in the current sample. Once we have used these five categorical criteria to establish the basic dialectological layout of the sample, we will move beyond them to dig deeper with more specific and more gradient measurements.

Janet B. easily satisfies all five criteria. Her /o/ is fronter than /ʌ/; /e/ is not only less than 375 Hz fronter than /o/, it is in fact backer than /o/; /æ/ is much higher and fronter than /e/; F1 of /æ/ is 510 Hz, much less than 700; and F2 of /o/ is 1,638 Hz, more than 1,500. On the other hand, Emily R. satisfies none of the five, with /o/ backer than /ʌ/, /e/ fronter

than /o/ by 375 Hz, /æ/ lower and backer than /e/, F1 of /æ/ is 829 Hz, and F2 of /o/ is 1,262 Hz.

Table 3 lists how many of the ninety-eight speakers in the data set satisfy each of the five criteria, compared to the 446 speakers with fully analyzed vowel systems in the Telsur corpus of *ANAE*. Although the EQ, AE1, and O2 criteria are satisfied by relatively small subsets of the New York State dataset, large majorities satisfy both the ED and UD criteria. Thus, with respect to ED and UD, the New York State speakers in this study overall basically resemble the Inland North speakers from the Telsur corpus. But with respect to the other three criteria the speakers in this study are overall more like the non–Inland North Telsur speakers. It is not expected, of course, that the speakers in this study's dataset will overall resemble the Inland North in all respects, of course; the sampled communities were chosen with the aim of being located on both sides of the eastern border of the Inland North. But it is noteworthy that, instead of being intermediate between Inland North and non–Inland North distributions of all five criteria, they are much closer to the Inland North in two of the five and much closer to non-Inland North communities in the other three. This means that even the communities that are found to be outside the Inland North, and perhaps within the Western New England region, will be likely to show largely Inland North–like ED and UD features. The Telsur corpus contains

Figure 3. The vowel system of Emily R., a twenty-one-year-old college student from Cooperstown. Red=/æ/; orange=/ʌ/; yellow=/e/; green=/i/; purple=/oh/; magenta=/o/. The "Emily" audio files are associated with this figure.

Table 3. *The number of speakers (out of ninety-eight) satisfying the five NCS criteria in this study's New York dataset, compared with ANAE's Inland North region (out of sixty) and the rest of the Telsur corpus (out of 386)*

Criterion	# NYS speakers	*ANAE* Inland North speakers (%)	Other Telsur (%)
UD	82	93	15
ED	82	84	13
EQ	22	66	3
AE1	27	84	17
O2	16	46	5

Table 4. *The NCS scores of speakers in this study's New York data set, compared with ANAE's Inland North region and the rest of the Telsur corpus*

# Criteria	# NYS speakers	*ANAE* Inland North speakers (%)	Other Telsur (%)
5	4	36	1
4	18	26	1
3	15	16	3
2	38	16	9
1	15	5	21
0	8	0	66

thirteen Western New England speakers; nine of them satisfy the UD criterion, but only five the ED criterion. So it is not surprising that a set of speakers straddling the Inland North/Western New England border satisfies UD to a very high degree; but the high rate of ED in the New York State corpus is characteristic of the Inland North but not Western New England.

In addition to how many speakers in the New York corpus satisfy each of the five NCS criteria, we can ask how many speakers satisfy *each number* of criteria— that is, how many speakers satisfy all five criteria, how many satisfy four, and so on. The number of criteria satisfied by any given speaker will be referred to as that speaker's *score* (or *NCS score*). These figures are displayed in Table 4. Whereas a large majority of Telsur speakers outside the Inland North meet none of the five criteria, and a plurality of Telsur Inland North speakers meet all five, in the New York corpus fairly few speakers meet either zero or five; the plurality of them meet exactly two. These results are unsurprising: Table 3 shows that two of the five criteria are met by large majorities of the New York corpus, while the other three are satisfied by relatively small minorities; thus it is expected that the most frequent score in the New York corpus would be two. However, Table 4 shows more clearly than Table 3 how the New York corpus sits in between the Inland North and non–Inland North Telsur subsets with respect to the five criteria.
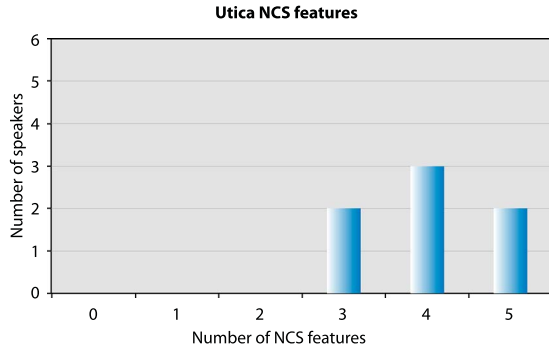
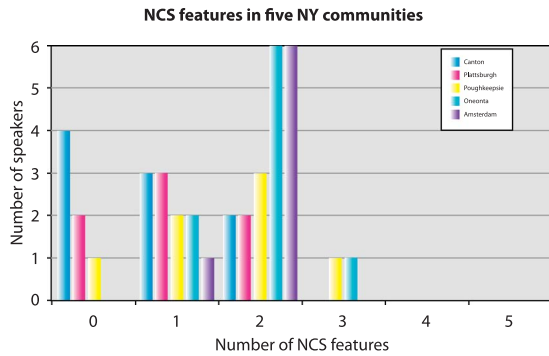**Figure 4.** NCS scores of speakers in Utica.

Figure 6. NCS scores of Telsur speakers in Western New England.

Figure 5. NCS scores of speakers in Amsterdam, Oneonta, Poughkeepsie, Plattsburgh, and Canton.

Figure 7. NCS scores of speakers in Gloversville, Glens Falls, Ogdensburg, and Watertown.

3.2. *Classifying communities*

In order to determine the location and nature of the Inland North/Western New England boundary, it is necessary to look at the twelve sampled cities and villages one at a time rather than in the aggregate, so that they can each individually be assigned to the Inland North, to Western New England, or to some other category. Utica is the easiest city to categorize in this dataset, as seen in Figure 4: Of the seven speakers in Utica, none have scores less than three, and a plurality score four. This places Utica solidly within the Inland North, in which the NCS dominates. This expands the known extent of the core Inland North region eastward by some fifty miles.

As shown in Figure 5, five of the twelve communities can be placed with confidence outside the Inland North region: Amsterdam, Oneonta, Poughkeepsie, Plattsburgh, and Canton. Among thirty-nine speakers sampled in these five communities, only two have a score higher than two; and three of the five communities range down to zero in at least one speaker. But although these five communities are clearly outside the range that would allow them to be categorized as part of the Inland North, neither are they very typical of communities in the Telsur corpus outside the Inland
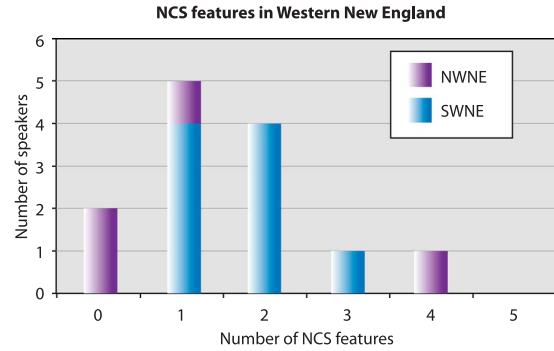
North. Outside the Inland North in the Telsur corpus, fully 87 percent of speakers have scores lower than two; in Amsterdam, Oneonta, and Poughkeepsie, more than half the speakers in this dataset score two or three. Only in Canton do a plurality of speakers meet none of the NCS criteria, and even that plurality is less than a majority. What these five communities overall resemble is *ANAE*'s Western New England region, whose scores are shown in Figure 6: the Western New England data are dominated by speakers meeting one or two criteria, with comparatively few exceptions below one or above two. Amsterdam, Oneonta, Poughkeepsie, and Plattsburgh each individually fit more or less within this profile, and Canton is not far from it. So we can tentatively group these five communities with Western New England, as *ANAE* does Albany.
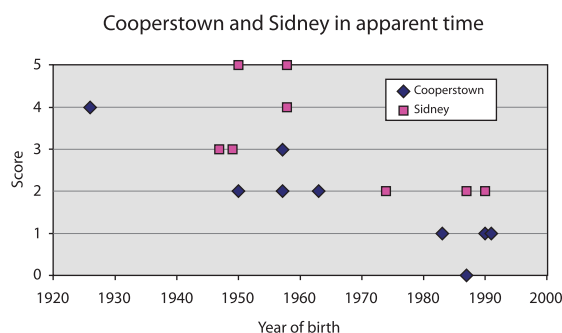
In fact, these five communities can be divided into Northwestern and Southwestern New England, in the same way Boberg (2001) divides the Western New England Telsur speakers. The key feature motivating Boberg's boundary between the two sub-regions is the *caught-cot* merger: the merger is complete or nearly so in Northwestern New England, and largely absent in the sampled cities in Southwestern New England. Although the *caught-cot* merger is not complete in any

Figure 8. NCS scores of speakers in Cooperstown and Sidney. NCS, Northern Cities Shift.

Figure 9. Northern Cities Shift scores in Cooperstown and Sidney versus year of birth.

of the communities sampled in this study, Plattsburgh and Canton each have only one speaker sampled with a secure distinction between /o/ and /oh/, and are the only sampled communities in which more than two speakers are fully merged in perception.[8] Since Plattsburgh and Canton are the northernmost communities appearing in Figure 5 and the only ones with advanced *caught-cot* merger, it makes sense to regard them as more closely affiliated with Northwestern New England, and Amsterdam, Oneonta, and Poughkeepsie as more closely affiliated with Southwestern New England. This distinction is also reflected in the communities' NCS scores, as Figure 5 shows: Plattsburgh and Canton have overall lower scores than Amsterdam, Oneonta, and Poughkeepsie[9] (*t*-test: $p < 0.0005$).

Figure 7 shows the scores of speakers in Gloversville, Glens Falls, Ogdensburg, and Watertown. There are no speakers sampled from any of these cities scoring zero or five. In each of the four cities, speakers' scores range between two and four, with the only exception being a single speaker in Ogdensburg with a score of one. This distribution matches neither the Inland North pattern (dominated by fives and fours with very few speakers below four) nor the Western New England pattern (mostly between zero and two with very few speakers above two); it seems to occupy a position intermediate between the two patterns. Although there appear to be differences between these four cities—Gloversville has a majority of speakers scoring four, and fewer scoring three or two, while Watertown shows a majority of twos and fewer threes and fours—these differences do not reach the level of statistical significance. These four cities are, however, found by *t*-tests to differ at the $p < 0.05$ level from both Utica and the five communities assigned above to the Western New England region. So it appears as if Gloversville, Glens Falls, Ogdensburg, and Watertown constitute an additional coherent set of communities in which the NCS exists but is not as dominant as it is in the Inland North proper; these cities may be

tentatively described as part of the "fringe" of the Inland North. In each of these "fringe" cities, there are speakers in the data who demonstrate the NCS very clearly, but nobody seems to satisfy all five NCS criteria. At the same time, there are also a substantial number of speakers who clearly are not subject to the NCS overall; but even they still mostly satisfy the ED and UD criteria and have /o/ relatively fronted as compared with /e/ and /ʌ/.

Of the ten communities discussed so far, eight have a difference of at most two points between their highest- and lowest-scoring speakers. The other two (Poughkeepsie and Ogdensburg) have all speakers but one within a range of two points, and a single high or low apparent outlier. Cooperstown and Sidney, the remaining two villages in the data, have scores that are a bit more spread out, as shown in Figure 8. Although Cooperstown is dominated by speakers scoring one and two, like some of the communities in Figure 5, it differs from those in that one speaker in Cooperstown has a score as high as four—higher than any speaker interviewed in the communities in Figure 5.[10] Indeed, scores in Cooperstown have a greater range than in any other community sampled in this study, from four all the way down to zero. Meanwhile, Sidney cannot be easily assigned to either the Inland North proper (like Utica) or the "fringe" as defined above: the Inland North proper is dominated by speakers scoring four or five, with relatively few twos and threes; and the fringe, as defined by Figure 7, includes no fives even in Gloversville, the fringe city with the highest average score. Sidney, whose sample in this study is roughly evenly spread out among all the scores between two and five, seems to display a profile unseen elsewhere in this sample.
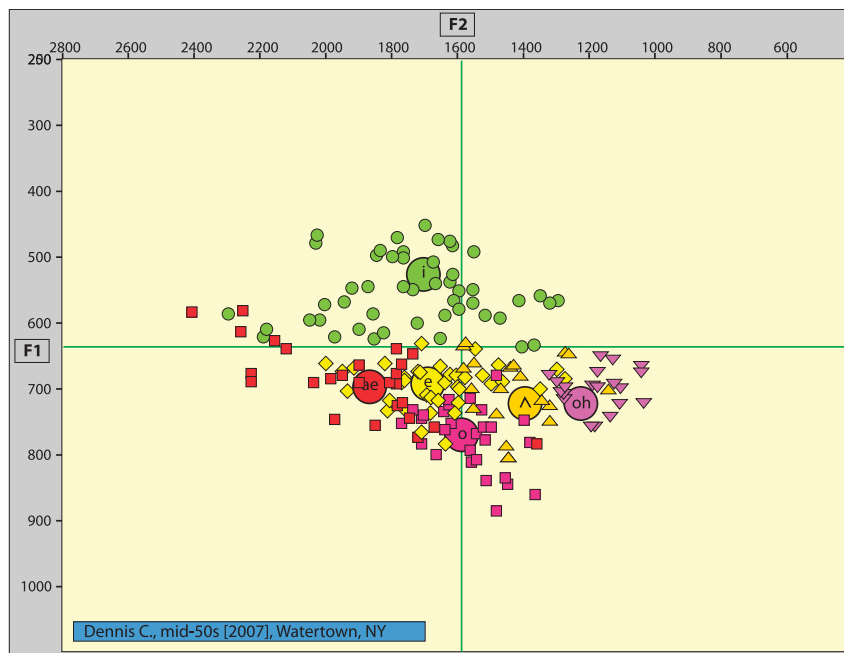
One way to deal with the seemingly irregular behavior of Cooperstown and Sidney would be to declare that Cooperstown belongs to the Western New England dialect region like the cities in Figure 5 and merely has a high-scoring outlier, and that Sidney belongs to an intermediate class between the Inland North proper and the fringe, just as the fringe was

defined as an intermediate class between the Inland North and Western New England. However, we can gain a clearer picture of Cooperstown and Sidney by looking at the speakers from those two villages in a bit more detail, from the perspective of change in apparent time. Figure 9 displays the relationship between NCS score and age.
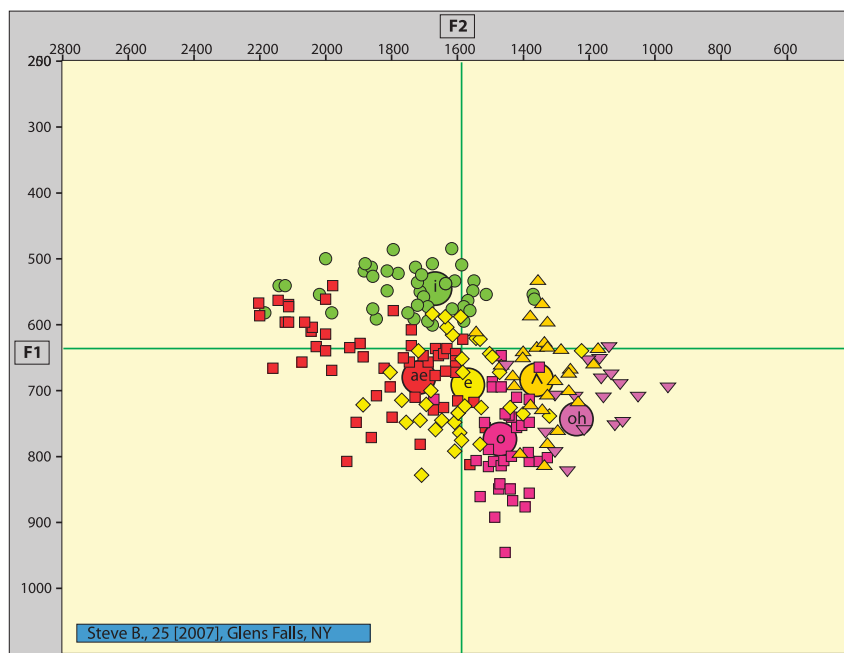
From the apparent-time point of view, the dialectological status of Cooperstown and Sidney becomes much clearer. In Sidney, the three speakers born later than 1970 all have a score of two, while the five older speakers score between three and five. In Cooperstown, the four speakers born after 1980 score one and zero, the four between 1950 and 1965 score two and three (but see note 7), and the one born in 1926 scores four. In both villages the difference between the younger and the older or middle-aged speakers is significant to the $p < 0.02$ level or better; additionally, in Cooperstown, the Pearson correlation between year of birth and score is significant with $p < 0.0005$ and $r^2 \approx 0.83$. So now it becomes clear that Cooperstown and Sidney are both in the process of *retreat* from the NCS. In Sidney, the older speakers fall more or less in the range of the Inland North, reaching scores as high as five but no lower than three; but the younger speakers all score two and would seemingly be at home in a community like Amsterdam or Oneonta, where large majorities of speakers score two. In Cooperstown, the older speakers seem from this data to belong to an Inland North fringe community, like Watertown, with scores between two and four; the

younger speakers all score below two, and in this respect are most similar to places like Canton and Plattsburgh, which were assigned above to the Northwestern New England region. The younger speakers in Cooperstown also agree with Canton and Plattsburgh in showing direct effects of the *caught-cot* merger; although the merger is absent from the older speakers in Cooperstown, all the younger speakers show complete or transitional merger in perception.

To sum up, then, according to the five NCS criteria used by Labov (2007), the twelve cities in this study can be categorized as follows: Utica belongs to the Inland North, fully subject to the NCS. Amsterdam, Oneonta, Poughkeepsie, Plattsburgh, and Canton are not subject to the NCS, although the UD and ED criteria—representing relative frontness of /o/ with respect to /e/ and /∧/—are frequently satisfied in them (unlike most non-NCS communities). These five resemble *ANAE*'s Western New England region to an extent— Amsterdam, Oneonta, and Poughkeepsie grouping with Southwestern New England, and Plattsburgh and Canton with Northwestern New England. Gloversville, Glens Falls, Ogdensburg, and Watertown belong to the "fringe" of the Inland North: the NCS is present in these communities, but inconsistently so. Cooperstown and Sidney are undergoing change in progress away from the NCS: Sidney from a core Inland North community to one more like Amsterdam and Oneonta; and Cooperstown from an Inland North fringe community to one with less conformance to the NCS than any other in this study.[11]

Figure 10. The vowel system of Dennis C., a museum caretaker from Watertown.
Red=/æ/; orange=/∧/; yellow=/e/; green=/i/; purple=/oh/; magenta=/o/.

Figure 11. The vowel system of Steve B., an unemployed roofer from Glens Falls.
Red=/æ/; orange=/∧/; yellow=/e/; green=/i/; purple=/oh/; magenta=/o/.

4. The EQ1 index

4.1. Definition and motivation

The five NCS criteria are a fairly blunt instrument for measuring the participation of a speaker or community in the NCS. This is because they are *categorical* criteria: for instance, the UD criterion is satisfied whenever mean /o/ is fronter than /∧/, regardless of how much fronter it is. In fact, *ANAE* and Labov (2007) do not even appear to take note of whether the F2 difference between /o/ and /∧/ is statistically significant when deciding whether a speaker meets the UD criterion; and for that reason, neither does the data presented above.
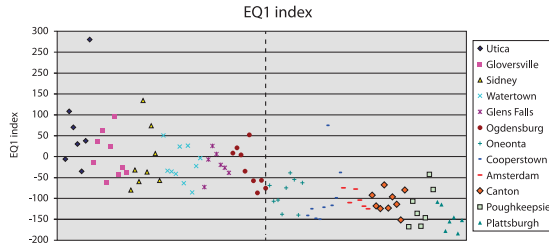
To see why this is important, consider the vowel system of Dennis C., a man in his fifties from Watertown who works as a museum caretaker, presented in Figure 10. Dennis C. easily satisfies the ED, UD, and O2 criteria. However, his mean F1 for /e/ is 697 Hz, and his mean F1 for /æ/ is 701 Hz—meaning he misses satisfying the EQ and AE1 criteria by only 4 Hz and 1 Hz, respectively. It is evident that Dennis C.'s /æ/ is quite raised, and no one would mistake it for a low vowel. It is not raised as far as it could go—some NCS speakers have /æ/ raised as high as /i/ or higher, like Janet B. in Figure 2—but he certainly shows some degree of participation in the NCS raising of /æ/, and the EQ and AE1 criteria give him no credit for it.

Moreover, compare Dennis C. to Steve B., a twenty-five-year-old unemployed roofer from Glens Falls, whose vowels are shown in Figure 11. Steve B. satisfies

both the AE1 and the EQ criteria, but by margins almost as small as those by which Dennis C. fails to satisfy them: Steve's mean /æ/ is 6 Hz higher than /e/ and 14 Hz higher than 700 (i.e., it is 686 Hz). Impressionistically, Steve's vowels look quite similar to Dennis's. Statistically, neither Steve's nor Dennis's /æ/ is significantly different either from /e/ or from 700 Hz, or from each other; for each comparison, a *t*-test finds $p > 0.1$ or worse. But because of the categorical nature of the AE1 and EQ criteria, this similarity between Steve's and Dennis's /æ/ distributions is lost in the data considered above.

To get a more gradient view of communities' different degrees of participation in the NCS, we will use a quantitative version of the EQ criterion—the EQ1 index. This is simply the difference in F1 between mean /e/ and /æ/—positive if /æ/ is higher, and negative if /e/ is higher. For instance, Dennis C.'s EQ1 index is –4; Steve B.'s is +6; Janet B.'s is +280; and Emily R.'s is –150.

The EQ1 index was selected, rather than gradient versions of the other four NCS criteria (i.e., the F2 distance between /e/ and /o/, the F1 value of /æ/, and so on), for several reasons. First, the raising and tensing of /æ/ is often described (by *ANAE*, for example) as the *first* stage of the NCS. If this is the case, the presence of /æ/-tensing will be the most important diagnostic of the NCS: if a community participates in the NCS at all, it ought to show some degree of raising of /æ/. Moreover, if a speaker or community is still in an incipient stage of NCS, they may show a small degree of raising of /æ/ that might escape coarse

Figure 12. EQ1 indices for all ninety-eight speakers. Communities are ordered from left to right by mean EQ1 index; within each community, speakers are ordered by age, with the youngest on the left. The dashed line separates communities with overall high indices, including Inland North core and fringe communities, from those with overall low indices, including those that resemble Western New England. Chart from Dinkin (2009); used with permission.

Table 5. *Mean EQ1 index for each community sampled*

Community	Mean EQ1	SD	*n*
Utica	+69	104	7
Gloversville	+4	53	9
Sidney	−6	74	8
Watertown	−19	43	10
Glens Falls	−19	32	7
Ogdensburg	−25	48	9
Oneonta	−88	36	9
Cooperstown	−96	73	9
Amsterdam	−103	19	7
Canton	−107	26	9
Poughkeepsie	−121	47	7
Plattsburgh	−148	29	7
Telsur Inland North	+22	73	60
Telsur non-IN	−110	55	386

Figure 13. A histogram of the EQ1 indices of speakers in the Telsur Inland North cities and this study's "fringe" cities. Each column along the horizontal axis represents a range of 20 Hz in EQ1 index—so the tallest red column represents twelve Telsur Inland North speakers whose EQ1 indices are between +11 and +30.

measures like the EQ and AE1 criteria but be visible quantitatively.
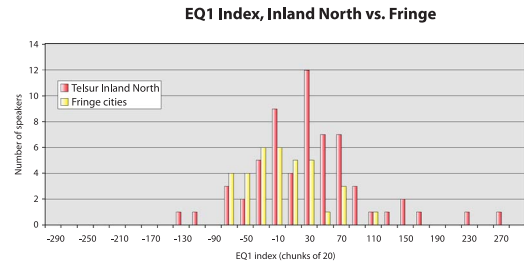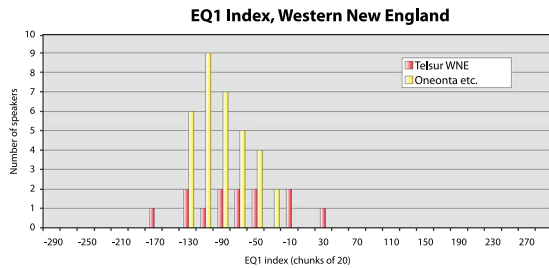
The distance in F1 between /æ/ and /e/ also shows greater variability from community to community than do the quantitative equivalents of the other four NCS criteria. According to an ANOVA analysis, the EQ1 index has an *F* ratio greater than ten—that is to say, the differences in EQ1 index from community to community are overall more than ten times as great as the variation found within the individual communities. The other four quantitative equivalents have *F* ratios between approximately 3 and 8, and therefore the EQ1 index does the best job of distinguishing between the communities sampled.[12] Since the purpose of this study is to group the communities into dialectological categories, it will be most illuminating to focus on the index that makes the sharpest distinctions between communities.

4.2. Results of the EQ1 index

Figure 12 displays the EQ1 indices of all ninety-eight speakers in the data set; Table 5 shows the mean EQ1 index for each community. It is fairly clear from Figure 12 that the twelve communities in the data are divided by the EQ1 index into two sets of six. In the six communities on the left side of Figure 12—Utica, Gloversville, Sidney, Watertown, Glens Falls, and Ogdensburg—all speakers in the data have EQ1 indices greater than or equal to 88. On the right, in Oneonta, Cooperstown, Amsterdam, Canton, Poughkeepsie, and Plattsburgh, all speakers in the data except one have EQ1 indices less than –37. The average of these two limits is –62.5, which can serve as a rough boundary between a "high" range of EQ1 indices, –62 and up, and a "low" range, –63 and below. In the six communities on the left side of Figure 12, only six

speakers have low EQ1 indices; in the six communities on the right, only five speakers have high EQ1 indices. This means a total of only eleven of these ninety-two speakers fall on the "wrong" side of the –62.5 line between the high-EQ1 communities and the low-EQ1 communities. So the distinction between the high- and low-EQ1 communities is a fairly clear one.

Moreover, the two sets of six communities are not only distinct from each other but relatively homogeneous within themselves. An ANOVA analysis reveals that the variation in EQ1 index among Utica, Gloversville, Sidney, Watertown, Glens Falls, and Ogdensburg is not quite sufficient to reach the level of significance ($p \approx 0.051$)[13]. Likewise, *t*-tests find no significant difference between any pair of these six high-EQ1 communities; the pair closest to being significantly different is Utica and Ogdensburg ($p \approx 0.056$). Similarly, ANOVA finds no significant difference ($p > 0.12$) among the six low-EQ1 communities—Oneonta, Cooperstown, Amsterdam,

Figure 14. A histogram comparing the EQ1 indices of the Telsur speakers in Western New England with Oneonta, Amsterdam, Canton, Poughkeepsie, and Plattsburgh collectively.

Canton, Poughkeepsie, and Plattsburgh—although *t*-tests show that Plattsburgh has lower EQ1 indices than both Oneonta and Amsterdam at the $p < 0.01$ level.[14]

It is reassuring that the two sets of six communities into which the EQ1 index partitions the data are similar to the groups into which the communities were classified above according to the five categorical criteria. Oneonta, Amsterdam, Canton, Poughkeepsie, and Plattsburgh, which were grouped as resembling Western New England in the previous section, appear together on the right side of Figure 12; Gloversville, Watertown, Glens Falls, and Ogdensburg, classified as "fringe" Inland North, all appear on the left side of Figure 12. Utica, rather than having overall distinctly higher EQ1 indices than the fringe cities in general, occupies a similar range with only one high outlier, and is not significantly different at the $p < 0.05$ level from any of them.[15]

The fringe cities' EQ1 indices justify identifying them as basically affiliated with the Inland North region, rather than merely being an intermediate category between the Inland North and Western New England that is not more closely associated with either one. Figure 13 shows that, although the mean EQ1 index of the fringe cities is slightly below that of the Telsur Inland North sample, they are well within the general EQ1 distribution of the Inland North overall; in fact, the mean EQ1 index of the fringe cities is −15, only half a standard deviation below the mean of the Telsur Inland North speakers. So the fringe cities can be identified as a set of communities that basically pattern as part of the Inland North, but are slightly less advanced in its key NCS features than the core Inland North region defined in *ANAE*.

Likewise, the five communities that were classified above as fitting more or less within *ANAE*'s Western New England region in their NCS scores resemble Western New England in EQ1 index as well. Figure 14 demonstrates how the EQ1 indices of Oneonta, Amsterdam, Canton, Poughkeepsie, and Plattsburgh

(mean: −112) match the range of those of the thirteen; Telsur speakers from Western New England (mean: −88); although Western New England appears to have a slightly higher mean, the difference is not significant. From comparing Figures 13 and 14, whose horizontal axes are drawn to the same scale, it is also clear that Western New England and these five cities in New York do *not* lie within the general range of the Inland North. Indeed, the distribution of EQ1 indices in Oneonta, Amsterdam, Canton, Poughkeepsie, and Plattsburgh is typical of non–Inland North communities—the mean EQ1 index of the 373 Telsur speakers outside the Inland North and Western New England is −111, almost exactly the same as these five New York communities.
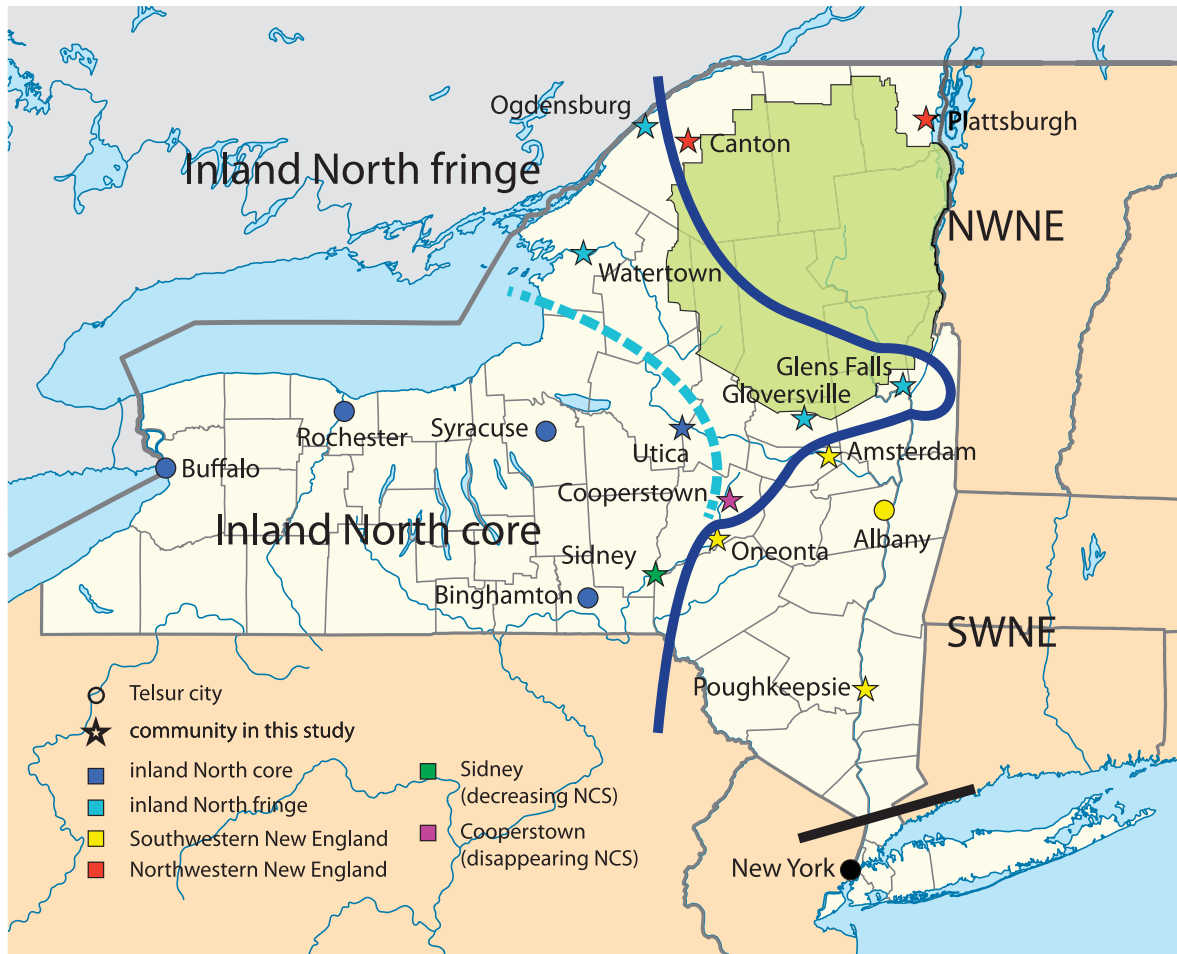
5. Mapping the results

5.1. Summary of classification

To sum up, the NCS scores and EQ1 indices together allow us to categorize the twelve communities in this study's dataset as follows. The Inland North region, where the NCS has a major presence, can be subdivided (in Upstate New York, at least) into "core" and "fringe" areas. In the core, all or nearly all speakers score three or more, while in the fringe, almost all speakers score between two and four, placing the fringe intermediate in score between the Inland North core and Western New England. The fringe agrees with the Inland North core, however, in its distribution of EQ1 indices. In this dataset, Utica is a core Inland North city, and Gloversville, Glens Falls, Watertown, and Ogdensburg are fringe Inland North cities. Sidney appears to have been originally in the Inland North core, but the NCS is weakening there, leaving younger speakers as part of the fringe at best. Oneonta, Amsterdam, and Poughkeepsie pattern more or less with Southwestern New England, and Plattsburgh and Canton with Northwestern New England. Cooperstown appears to be an originally Inland North fringe community which is now retreating from the NCS quite rapidly (apparently as a result of new-dialect formation—cf. Dinkin 2009:§5.5, 2012); it is becoming more like Plattsburgh and Canton than like any other communities in this study, although it is not near the northern border of New York like they are. Map 4 displays the dialect regions of Upstate New York as determined by this analysis.

5.2. The Hudson Valley

ANAE does not identify any other dialect regions between the Inland North and Western New England. The discussion so far in this paper has more or less agreed with that position, finding that the communities
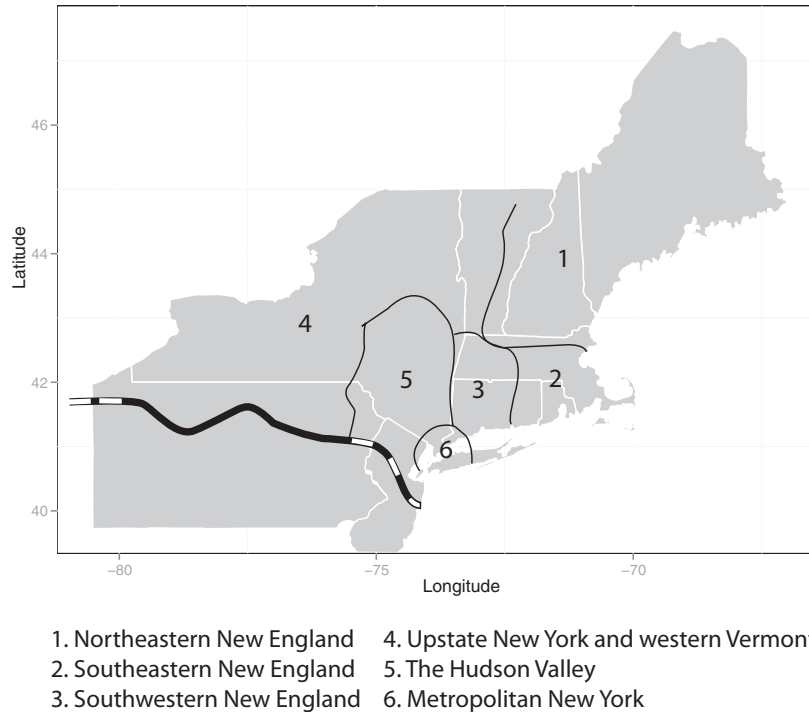
Map 4. The dialect regions determined by this study. The isoglosses indicate the status of the communities *before* the start of the changes in progress in Cooperstown and Sidney: the dark blue line indicates the limit of the Inland North as a whole, and the light blue line separates the Inland North core from the fringe. Map based on Dinkin (2009); used with permission.

not included in the Inland North can be grouped generally with Northwestern or Southwestern New England. However, the southeastern boundary of the combined Inland North fringe and core regions in Map 4, separating Sidney, Cooperstown, Glovers-ville, and Glens Falls on the one hand from Oneonta and Amsterdam on the other hand, seems to correspond roughly to the northeastern boundary of the Hudson Valley region determined by Kurath (1949). Map 5 is a copy of Boberg's (2001) reconstruction of the boundaries Kurath assigns to his Hudson Valley dialect area and adjacent regions. The general location of the boundary between regions 5 and 4 in Map 5 does indeed seem, impressionistically, fairly similar to that of the boundary in Map 4 between the communities associated above with Southwestern New England and those assigned to the Inland North core or fringe. This suggests, of course, that it is the same boundary; the lexical boundary of the 1940s has become a phonetic or phonological boundary by the 2000s.

It is difficult to establish exactly which communities Kurath meant to include in the Hudson Valley region: the map on which Kurath presents these regions' boundaries includes no cities or landmarks other than a few sketchily (and none-too-accurately) drawn rivers. Boberg's version of the map is somewhat better, at least in that it shows the rivers more clearly and accurately; but the relationships shown on it between the dialect boundaries and the rivers must be taken with a grain of salt simply because the boundaries are copied from Kurath's own map and inherit the effects of the latter's lack of detail. Based on the positions of the rivers in Boberg's redrawing, Kurath's Hudson Valley region seems to just barely exclude Glens Falls, Utica, and Sidney, and just barely *include* Gloversville and Cooperstown, as well as Oneonta. But due to the overall lack of clarity of Kurath's map, the precise sets of communities that it appears to include in or exclude from the Hudson Valley region are of less importance than the fact that the Hudson Valley region seems to

1. Northeastern New England   4. Upstate New York and western Vermont
2. Southeastern New England   5. The Hudson Valley
3. Southwestern New England   6. Metropolitan New York

Map 5. Based on Figure 3 from Kurath (1949). Kurath's lexical dialect regions of New York and New England.

correspond fairly well as a general region to the area in southeastern New York excluded from the Inland North fringe and core in Map 4, based on the communities sampled in this study.

These communities, of course, were shown above to be overall relatively similar to the *ANAE* Southwestern New England communities with respect to the NCS features being examined; and so the question of whether a present-day dialect boundary exists between the Hudson Valley and Southwestern New England has yet to be answered. However, the similarity of Kurath's Hudson Valley boundary to the boundary in Map 4 suggests giving Kurath the benefit of the doubt. In that spirit, we will take a cue from Kurath, and identify the region containing Poughkeepsie, Amsterdam, and Oneonta—defined generally as the area of New York State north of the New York City dialect region and southeast of the Inland North fringe, showing NCS scores mostly around two and relatively low EQ1 indices—as the Hudson Valley.

5.3. Boundaries and communication patterns

At first glance, Map 4 seems to indicate that there is a gradual transition between the Inland North and Hudson Valley—from, for example, Utica (Inland North core, full NCS) eastward to Gloversville (Inland North fringe) to Amsterdam (non–Inland North, but with relatively high scores for a non–Inland North city)

to Albany and Western New England proper; or from Binghamton (core) to Sidney (diminishing NCS) to Oneonta (like Amsterdam). Given the observations above that the Hudson Valley appears phonologically similar to Southwestern New England, this is consistent with Boberg's (2001) conclusion that there is no phonological difference between the Inland North and Southwestern New England. Thus the Hudson Valley could be regarded as basically an extension of the Inland North on which the NCS has not had its full effect. But there are irregularities and discontinuities in this picture that suggest that a gradual transition is not the whole story.

Most noticeable is the irregularity in the border itself—the Inland North fringe extends almost all the way to the Vermont border at Glens Falls; but further north or south, at Plattsburgh, Albany, or Poughkeepsie, the NCS is not found anywhere near so far east. Now, there's no reason at all for a gradual transition between the Inland North, the Hudson Valley, and Western New England to imply that the outer boundary of the NCS must be at a uniform distance from the edge of New York State at every latitude; but it still seems in need of some explanation that at Glens Falls the fringe extends so much further from the Inland North core than it seems to anywhere else. If the Inland North fringe, as Boberg's analysis might suggest, is merely the advancing expansion of the NCS toward the Western New England territory that is open to it, then we would expect the

fringe to extend furthest from the core along the major routes of communication and travel between the Inland North core and the Hudson Valley—much as, in Illinois, NCS features are expanding outside of the Inland North via the communities along Interstate 55 between Chicago and St. Louis (Labov, 2007).

In the case of the Inland North fringe, the chief route of east–west communication and travel is either the New York State Thruway (Interstate 90) or—if we allow for the eastward spread of NCS-like features earlier than the NCS was first reported—the Erie Canal and Mohawk River, and the railroads that follow the Canal. However, neither the Thruway nor the Erie Canal and Mohawk River quite follow the direction of the eastern extent of the Inland North fringe. While the Thruway, Canal, and Mohawk, heading east from Utica, pass through Amsterdam and Albany, the Inland North fringe bypasses both Amsterdam and Albany and includes Gloversville and Glens Falls, both of which are some distance to the north. Moreover, although dialect diffusion is frequently found to affect larger cities sooner than smaller cities (the "cascade" model; cf. Labov, 2003), Amsterdam is in fact (slightly) larger than Gloversville, so population differences cannot account for the NCS having reached Gloversville but not Amsterdam. Thus the eastern edge of the Inland North does not support the hypothesis of NCS features simply spreading from the west across a dialect continuum that is phonologically open to it, since the non-NCS city of Amsterdam is both more populous and more conveniently connected to the Inland North core than the Inland North fringe city of Gloversville is.

Another aspect of the relationship between Gloversville, Amsterdam, and Albany seems to call into question the importance of present-day communication patterns in determining the boundary of the Inland North. Gloversville and Amsterdam are quite close together—less than fifteen miles apart by road, with three or four sparsely-populated towns in between them—and yet the difference between them in this dataset is fairly stark: Gloversville has the highest mean NCS score of any Inland North fringe city, while Amsterdam has no speakers scoring above two; and the two cities' EQ1 indices do not overlap at all (Gloversville's lowest is –61 and Amsterdam's highest is −75). Even more important than the two cities' mere proximity is their regional orientation, as reflected by the interview subjects' responses to questions about their local travel habits. Gloversville and Amsterdam are both regionally affiliated with the Albany area: residents of both cities watch television channels that broadcast out of Albany[16], read newspapers from Schenectady (which is midway between Albany and Amsterdam on the Thruway), and travel to Albany and Schenectady to go shopping. Each of

the twelve in-person interview subjects in Amsterdam and Gloversville reported frequent trips to Albany, Schenectady, or both.[17] But although Amsterdam and Gloversville are both part of the greater metropolitan area of Albany and subject to Albany's regional influence, Amsterdam is part of the same general Hudson Valley dialect group as Albany and Gloversville is not. Similarly, Oneonta appears to be regionally more oriented toward Binghamton than toward Albany, and receives Binghamton and Utica television stations, but does not appear to be subject to the NCS as Binghamton and Utica are. So the present-day regional affiliations and communication patterns of small and medium-sized Upstate New York cities are not a good predictor of which are included in the Inland North region and which are not; the spreading of the NCS does not seem to be effectively determined merely by channels of communication.

6. Settlement of the communities in the sample

6.1. Historians' descriptions of settlement patterns

Kurath (1949:1) states that "there can be no doubting the fact that the major speech areas of the Eastern States coincide in the main with settlement areas and the most prominent speech boundaries run along the seams of these settlement areas"; a striking example of this in Kurath's (1939) work is the linguistic and settlement boundary between Eastern and Western New England. As for the topic of the current study, *ANAE* and Boberg (2001) contend that the settlement history of upstate New York is important in explaining the origin of the NCS: Boberg (2001) focuses on the role of Western New England as a "staging ground" for the Anglophone settlement of the Inland North to explain the phonological similarity between the two regions, and *ANAE* argues that the tensing of /æ/ was made possible by the settlement boom drawn into central and western New York by the construction of the Erie Canal. This suggests that the early settlement history of the twelve communities in this study's sample could illuminate the distribution of dialect boundaries.

What is now New York State was founded in the 1620s as a Dutch colony named New Netherland, and only came under English control in 1664. During the New Netherland period, the Dutch founded towns along the Hudson River that still exist today, including not only New York City (then called New Amsterdam) but as far north as Schenectady and Albany (then called Beverwyck). Even after the English gained control of the colony and changed its name and that of its chief city to New York, the Dutch population remained mobile and new towns were founded by Dutch settlers and their descendants. Poughkeepsie is

one such: it was first settled by Dutch families in the 1680s, and Dutch was the main language of Dutchess County[18], of which it is the seat, until almost the 1770s (Platt, [1905] 1987).

Amsterdam, although founded much later than the period of Dutch colonial dominance, was another community subject to Dutch influence and settled by the descendants of Dutch settlers, as its name suggests. Amsterdam was founded in the late eighteenth century (Farquhar & Haefner, 2006) and Dutch families such as the Vedders and Hagamans were leaders of the community for several decades (Donlon, 1980). At the time when the name of the town was changed from Veddersburg to Amsterdam in 1804, out of recognition of the strong Dutch influence in the community, "the hamlet had acquired a considerable population, with an almost equal proportion of Dutch and Yankees" (Frothingham, 1892b:184).

How, then, does Gloversville differ from Amsterdam? After the American Revolutionary War in the 1770s, the location that would become Gloversville was basically depopulated. The settlement which led to the present-day city was not composed of descendants of the original Dutch New Netherland settlers, but rather by westward migrants from New England: Frothingham (1892a:334) writes: "The immigration was largely of Anglo-Saxon elements. The Dutch and Germans of the Mohawk Valley were already dwelling upon richer lands.… The New Englander, however… was naturally restless." In particular, "among the early settlers the Connecticut influence seems to have been strongest. A large element of the population came from the neighborhood of Hartford, and especially from West Hartford" (335). So the difference between Gloversville and Amsterdam is in their sources of settlement: While Amsterdam, like Albany and Poughkeepsie, had from its earliest days a large and influential Dutch population, Gloversville had very little influence from the early Dutch settlements of New York; its population was derived mostly from New England in general and Connecticut in particular. This supports Boberg's conclusion that settlement from Western New England supplied the necessary preconditions for the NCS in the Inland North—in Gloversville, the city settled from Western New England, the NCS is present; in Amsterdam, with little or less Western New England settlement history and substantial New Netherland Dutch influence, the NCS is absent.

This pattern can be tested on the other communities in the sample; the exposition here will be abbreviated in some respects, but fuller discussion of the best data I could find on the settlement history of each community can be found in Chapter 3 of Dinkin (2009).

The area that would become Glens Falls was first settled in 1763 and 1783 by a community of Quakers who had originally migrated from Connecticut (Brown, 1963; Hyde, 1936); and migration to Glens Falls from Connecticut, and through it to elsewhere in the Inland North, continued from 1784 until nearly 1850 (Glens Falls Historical Association, 1978). Utica's population was "in main part from New England" (Roberts, 1911:261); Utica was originally part of the town of Whitestown, and according to Ryan (1983:19), "almost 90% of the pioneer families of Whitestown came from Connecticut or Massachusetts." Utica is the county seat of Oneida County, whose boundaries, according to Durant (1878), were apparently drawn so as to include New England Yankee settlements and exclude Dutch settlements; and the first landowners in Watertown came "mostly from Oneida County" (Hough, 1854:250) and thus were part of this New England settlement pattern themselves. The available information on Ogdensburg is slightly less detailed; Merriam (1907:3) writes that it was settled by a "tide of emigration from New England," though she does not state specifically what part of New England the migration originated from.

Thus Ogdensburg, Watertown, Utica, Glens Falls, and Gloversville—all of the cities categorized in this paper as linguistically part of the Inland North core or fringe—appear to have been settled predominantly from New England. The communities that are undergoing change in apparent time away from the NCS appear to be the same. With respect to Cooperstown, Cooper (1886[1838]:20) himself writes "During the summer of 1787, many more emigrants arrived, principally from Connecticut, and most of the land on the patent was taken up." The History of Delaware County (1880) lists the names of several pioneering settlers of Sidney, and the majority of those whose geographic origins can be traced originate from Connecticut; moreover, Murray (1898:48) says that "The great mass of the early settlers in Delaware county were from New England." So, all of the communities in which the NCS is found in this study derived their early settlement primarily from New England, and in at least most cases from Southwestern New England.

Among the non-NCS communities, Poughkeepsie and Amsterdam are both discussed above; Poughkeepsie was settled by Dutch families and Amsterdam was at least half Dutch in its early population. Plattsburgh's early settlers, as listed by Hurd (1880), seem to have been principally from Long Island. Of Oneonta, Campbell (1906:36) writes "The first settlers were mostly German Palatines from Schoharie and the Mohawk"; the web site of the city of Oneonta[19] agrees with Campbell but adds the Dutch, saying "The first settlers to make this area their home were Palatine Germans and Dutchmen from the Schoharie and Mohawk Valleys." Neither source lists New England as a major origin for the settlers of

Oneonta, although a few of the individual Oneonta pioneers listed by Campbell have New England origins (just as one of the principal pioneers of Sidney listed in the *History of Delaware County* has a Dutch background).

Canton is the fifth non-NCS community in this study, but unlike the four discussed in the preceding paragraph, it in fact was settled from New England. In particular, Hough (1853:279) writes of Canton that "in 1802, the town began to settle rapidly[…] most of them with families, and from Vermont." It is unsurprising to find that Canton was settled from Vermont, of course; Canton has been assigned in this paper to the Northwestern New England dialect region, which to date has been described (Boberg, 2001) as consisting essentially of western Vermont. So, unlike the preponderance of Inland North communities, which were clearly settled from Southwestern New England, Canton's settlement was derived principally from Northwestern New England.

Therefore we can conclude that, insofar as clear reports of settlement history can be found, all the communities where the NCS is observed in this dataset were settled principally by populations of Southwestern New England origin, while the other communities were not. Although the status of Ogdensburg is slightly unclear from the historical data, the account given by Merriam (1907) is certainly at least consistent with this conclusion, if not directly supporting it.

6.2. Settlement history and the Hudson Valley

The patterns of settlement further justify identifying Poughkeepsie, Amsterdam, and Oneonta in this study with Kurath's Hudson Valley region. Where the communities of the Inland North all drew settlers from Western New England, Amsterdam, Oneonta, and Poughkeepsie instead all drew settlers from the original Dutch New Netherland population. Meanwhile, although Kurath draws a dialect boundary between Southwestern New England and Northwestern New England, Boberg (2001) argues that that boundary is not justified by Kurath's data. As Boberg implies, if that boundary is ignored in Map 5, all the areas ultimately settled from Southwestern New England—southwestern and northwestern New England as well as northern, central, and western New York—are united in a single dialect region, while the Hudson Valley area is separate. Kurath likewise describes westward migration from Western New England as having set the stage for the linguistic status of Upstate New York. So it makes sense to interpret Kurath's Hudson Valley region as constituting "the region not settled by Western New Englanders," and in particular the region in which Dutch influence was stronger than New England influence. Thus in the first half of the twentieth century as well,

the dialect regions were found to correlate well with settlement patterns, and the Hudson Valley was considered to be a linguistic region distinct from Southwestern New England. In this light, let us examine the present-day relationships between the Inland North, the Hudson Valley, and Southwestern New England.

7. Absence of the NCS in Southwestern New England

7.1. The problem

The fact that the distribution of the NCS in central and eastern New York State appears to be determined by settlement from Southwestern New England seems to support Boberg's (2001) general argument that Southwestern New England shares the same phonological system as the Inland North, and the settlement of the Inland North from Southwestern New England is the source of the phonological preconditions for the NCS. Despite Boberg's contention that Southwestern New England is *phonologically* identical to the Inland North, however, the difference between Southwestern New England and the Inland North is clear: According to the criteria used in this paper, Southwestern New England does not really show the NCS. This is an apparent paradox: If settlement from southwestern New England determines whether a community in central or eastern New York is subject to the NCS, why is present-day Southwestern New England itself *not* subject to the NCS?

A possible response to this paradox is that Southwestern New England *is* subject to the NCS, but to a lesser degree than the Inland North proper; this is the position Boberg takes. It is true to an extent, in that the seven Telsur speakers in southwestern New England proper (i.e., Connecticut and western Massachusetts) show higher NCS scores than the rest of the Telsur corpus outside of the Inland North ($p \approx 0.005$): three of them score 1, three 2, and one 3, whereas outside the Inland North in general, 66 percent of speakers score 0. Moreover, it seems that the NCS did not occur simultaneously in every community subject to it; in Ogdensburg the NCS seems to be still in progress even after it has apparently gone to completion in the other communities in this study. Perhaps the NCS originated in central or western New York, and then spread northward and eastward into the communities that now constitute the Inland North fringe. Under this scenario, even if Southwestern New England is in principle open to the NCS, the eastward spreading of the full NCS was never able to *reach* southwestern New England, which shares no geographical borders with the Inland North core or fringe. This scenario appears to be supported by the presence of one Telsur

speaker with an NCS score of four in Rutland, Vermont: The nearest community to Rutland of more than fourteen thousand people is Glens Falls, some fifty miles to the southwest, which is the easternmost known point of the Inland North fringe. So, according to this scenario, the reason the NCS has not expanded into Southwestern New England is because the Inland North does not come very near Southwestern New England; but where the Inland North fringe approaches *Northwestern* New England (which, like the Inland North, was originally settled from Southwestern New England), the NCS has been able to make a bit of eastward progress into Rutland.

But this is not a fully satisfactory resolution to the paradox, for two reasons. First, if the NCS can spread into Northwestern New England after all, we are left with the question of why Ogdensburg displays the NCS and Canton does not; Ogdensburg may have been principally settled from Southwestern New England, but Canton was definitely settled from Vermont. Second, and more important, the seven Telsur speakers in Southwestern New England show approximately the same distribution of NCS scores as the speakers from the three Hudson Valley cities in the current sample. As in the Hudson Valley, the two most frequently satisfied NCS criteria in Southwestern New England are UD and ED, representing frontness of /o/ relative to /ʌ/ and /e/. The mean EQ1 index of Telsur's Connecticut and western Massachusetts speakers is –80, perhaps slightly higher than the mean –102 of Poughkeepsie, Amsterdam, and Oneonta but certainly not to a statistically significant degree ($p > 0.34$). So, not only does Southwestern New England not display the NCS to the same degree that places that were *settled* from Southwestern New England do, but it is very similar (using the measures employed in this paper) to places that were *not* settled from Southwestern New England. So, why should the Hudson Valley, which was *not* settled from Southwestern New England, bear a closer linguistic resemblance to Southwestern New England than those that were? Or to put it another way, if Southwestern New England is in principle open to the NCS, what is it that makes Southwestern New England different from the Hudson Valley, which shows no evidence of being open to the spread of the NCS? If the Hudson Valley were as open to the NCS as Southwestern New England is supposed to be, then surely there would be more evidence of it in Amsterdam, for example.

7.2. The distribution of individual NCS features

As mentioned above, large majorities of Hudson Valley speakers in this study satisfy the ED and UD criteria (nineteen out of twenty-three for both ED and UD),

Table 6. *Mean /o/ F2 in various sets of communities*

Speakers	/o/ mean F2	n
Telsur Inland North	1495	60
Inland North fringe	1461	35
Telsur Southwestern New England	1418	7
Hudson Valley	1411	23
Other Telsur /o/ ~ /oh/ distinct	1337	243
Other Telsur /o/ ~ /oh/ merged	1252	130

"Other Telsur" indicates all communities outside *ANAE*'s Inland North and Western New England regions. "Distinct" and "merged" indicate communities respectively outside and inside the green isogloss of *ANAE*'s Map 9.1, which indicates the areas of completed *caught-cot* merger.

Table 7. *Mean /e/ F2 in various sets of communities*

Speakers	/e/ mean F2	n
Utica + Telsur New York Inland North	1625	15
Inland North fringe	1644	35
Hudson Valley	1717	23
Telsur Inland North w/o New York State	1755	52
Telsur Southwestern New England	1780	7
Other Telsur	1850	373

while at most two satisfy any of the other NCS criteria. Of the seven Telsur speakers in Southwestern New England, six satisfy UD, while three satisfy ED and no more than two satisfy any other criterion. Now, the ED and UD criteria each combine measurements of two distinct features of the NCS: the fronting of /o/ and the backing of /e/ or /ʌ/. These pairs of features, however, are in principle independent of each other: outside of the overall chain-shift structure of the NCS, there is no direct causal relationship between the fronting of one low vowel and the backing of one or two mid vowels; and thus saying that a community outside the Inland North satisfies (for example) the ED criterion obscures the question of whether that community has a fronted /o/, a backed /e/, or both. Since it is in the ED criterion that the Hudson Valley resembles the Inland North and differs from Southwestern New England, let us decompose ED and look at /o/ and /e/ separately.

Table 6 displays the mean F2 of /o/ in each of several subsets of this study's data and the Telsur corpus. The key finding here is that although /o/ is backer in southwestern New England and the Hudson Valley than in the Inland North, it is nevertheless a great deal fronter than the average /o/ outside of the

Inland North, even when regions where the *caught-cot* merger dominates are excluded.[20] This means that, compared with the rest of North American English, the Hudson Valley and Southwestern New England have a fronted /o/, though not quite to the same extent that the Inland North region does. While /o/ is backer in the Hudson Valley than in the Inland North fringe ($p \approx 0.013$), Southwestern New England's /o/ is very close to the Hudson Valley's but does not reach the level of significant difference from the Inland North fringe ($p > 0.21$).

Table 7 displays the mean F2 of /e/ in each of several sets of communities. While Table 6 treats all the Telsur Inland North communities as a set, Table 7 separates the four New York Inland North cities in the; Telsur corpus (Binghamton, Buffalo, Rochester, and Syracuse) from the rest of the Inland North and groups them with Utica, the only core Inland North community sampled in the current study. The purpose of this is to emphasize one of the most striking results on Table 7: The backing of /e/ is a great deal more advanced in the New York portion of the Inland North, core and fringe, than in the rest of the Inland North. Indeed, even the Hudson Valley cities, which are not subject to the NCS as a whole, have /e/ at least as backed as the Inland North communities outside of New York State (the difference between them is not statistically significant), and substantially backer than the rest of the Telsur corpus as a whole ($p < 10^{-6}$). For F2 of /e/, unlike /o/, the seven southwestern New England speakers are markedly different from the Inland North fringe ($p < 0.001$), and the Hudson Valley appears to sit between the Inland North fringe and Southwestern New England.[21]

So, to sum up, the relationships between Southwestern New England, the Hudson Valley, and the Inland North differ with respect to three key aspects of the NCS. In the raising of /æ/ over /e/, as shown in Figures 13 and 14, Southwestern New England and the Hudson Valley are relatively close to each other (mean EQ1 indices –80 and –102, respectively), and much farther from the Inland North fringe (mean EQ1 index –14). In the fronting of /o/, the Hudson Valley is significantly different from the Inland North fringe while Southwestern New England is not; and in the backing of /e/, the Hudson Valley is more similar to the Inland North than southwestern New England is.

The answer to the question asked above about why the NCS does not spread into the Hudson Valley may be that it *does*—but only partially: The backing of /e/ and fronting of /o/ are NCS features that are robustly present in the Hudson Valley, while the raising of /æ/ does not extend much beyond the Inland North. Labov (2007) argues that it is easier for changes in individual

phonemes to expand past their original isoglosses than for an entire chain shift to spread in the same manner as it originally occurred. So the various NCS features do not show uniform behavior across the eastern boundary of the Inland North. To explain these different behaviors, let us consider the relative chronology of the different phases of the NCS.

7.3. The origin of the NCS

There is disagreement in the literature about the earliest stages of the NCS. As mentioned above, Labov and his collaborators (as exemplified in, e.g., *ANAE*) usually describe the raising and tensing of /æ/ as the first stage of the NCS, creating a pull chain in which /o/ is fronted in order to fill the space left in the low front position by the raising of /æ/. Other researchers (e.g., Thomas, 2001; McCarthy, 2010), on the other hand, argue that the fronting of /o/ took place very early, preceding /æ/-raising; this argument tends to be based on analysis of archival data and early recordings, finding /o/-fronting more reliably than /æ/-raising in the speech of Inland North natives born in the late nineteenth or early twentieth centuries. The earliest study known to me that finds in Upstate New York phonetic features now associated with the NCS (Thomas, 1935, 1936) describes both /æ/ as raised and /o/ as fronted; an earlier study of the city of Ithaca (Emerson, 1891) describes /æ/ as low front and /o/ as low back.

The contention that /o/-fronting preceded /æ/-raising is consistent with the behavior of /o/ observed in the current study. Southwestern New England is the origin of the settlement of the Inland North, and it resembles the Inland North in that its /o/ is markedly fronter than the /o/ of non–Inland North communities in the Telsur corpus. It does not resemble the Inland North with respect to /æ/. This suggests that the fronting of /o/ could have begun early in the history of the NCS, before the present-day Inland North diverged from Southwestern New England speech; thus when the settlers of the Inland North region migrated westward, they already carried with them a somewhat fronted /o/. The Hudson Valley communities that were not settled by New Englanders did not necessarily, under this scenario, already have a fronted /o/, but the fronting of /o/ would have spread to them at a later date from both directions.

The backing of /e/ is a much newer change; in fact, apparent-time data shows that it is still in progress in the Inland North fringe overall while raising of /æ/ and fronting of /o/ are not. This change apparently originated in the New York State component of the Inland North after it had already diverged from Southwestern New England, unlike fronting of /o/;

for this reason, Southwestern New England's /e/ is much less backed than in New York's Inland North communities, while its /o/ is comparable with at least the Inland North fringe. Like /o/-fronting, /e/-backing appears to have spread from the Inland North to the Hudson Valley; and then it must have advanced from there to Southwestern New England as well. Thus those two regions have an /e/ that is substantially backer than North American English as a whole, but still not as backed as in the Inland North in New York State.

According to this approach, the raising of /æ/ would (like /e/-backing) have originated in the Inland North after it had diverged from Southwestern New England; but then, unlike /e/-backing, /æ/-raising never expanded into the Hudson Valley or, for the most part, New England beyond it. The raising of /æ/ may also have allowed the Inland North to develop a fronter /o/ than its Western New England predecessor system, by opening up additional phonetic space for /o/ to move forward into. But why should the raising of /æ/ fail to spread while the backing of /e/ and fronting of /o/ apparently spread easily into the Hudson Valley?

A hint at the answer to this can be found in the allophonic structure of /æ/ in the different regions. *ANAE* describes several allophonic configurations that /æ/ can take. Chief among these are the so-called nasal system, in which /æ/'s prenasal allophone is raised and fronted in a phonologically discrete manner and does not overlap phonetically with the non-prenasal allophone, and the continuous system, in which prenasal tokens of /æ/ are on average higher and fronter than other tokens, but gradiently as part of a single phonological distribution of /æ/. In Dinkin (2011a), I find that continuous /æ/ systems are very rare in the Hudson Valley but quite frequent in the Inland North; and I argue on phonological grounds that the nasal /æ/ system may be able to prevent the development of general NCS raising of /æ/—the raised prenasal allophone in effect prevents non-prenasal /æ/ from moving into its phonetic space. If this is the case, the allophonic patterns may explain why /æ/-raising, unlike other NCS features, did not spread effectively from the Inland North to the Hudson Valley: The Hudson Valley's nasal /æ/ system blocked acquisition of the Inland North's raised /æ/, but there was no structural difference sufficient to block the backing of /e/ or the fronting of /o/.

This account can be unified with *ANAE*'s hypothesis about the origin of the NCS. To review, *ANAE* argues that the tensing of /æ/ originated when the construction of the Erie Canal drew settlers from a variety of dialect regions, with a variety of phonological /æ/ patterns, into the same area. This account at face value does not fully account for the distribution of /æ/-tensing in New York State. For example, Amsterdam is located along the Erie Canal and was founded and settled in the same general time frame as the NCS communities in this study; but the presence of the Erie Canal was not sufficient to cause the NCS there. Combining the Erie Canal explanation with this study's findings of southwestern New England–origin settlement yields a consistent dialectological picture.

Under such a combined explanation, the general raising of /æ/ under the NCS would have been not *merely* the result of a koineization of multiple incompatible /æ/ systems in one place. Rather, given that there is little evidence for /æ/ having become substantially raised this early, it may be that the result of multiple incompatible /æ/ systems coming into contact, in communities founded by Southwestern New Englanders but subject to increased migration thereafter as part of the Erie Canal population boom, was the continuous /æ/ system. Communities in the Hudson Valley, having a different mix of contributing /æ/ systems and no underlying Southwestern New England substrate, may instead have ended up with /æ/ systems with discrete allophones, such as the nasal system. The persistence of such an /æ/ system in the Hudson Valley would then have prevented the NCS raising of /æ/ from spreading east from the Inland North, even while other NCS vowel changes did so.

Not all the communities in Upstate New York in which substantial /æ/-raising is found are located on the Erie Canal or directly benefited from the population boom it caused, but as far as we can tell, all were founded by Southwestern New Englanders. These communities not on the Canal form the Inland North fringe in this paper. They would have started with the same Southwestern New England–derived /æ/ system that was the substrate for the development of general tensing in the Erie Canal Inland North cities. By virtue of being in Upstate New York, many of them along major trade routes that connected to the Erie Canal, they would have been *in more or less regular linguistic contact* with the Erie Canal communities that were developing the NCS. Thus, in this model, the Inland North fringe did acquire the NCS in a manner consistent with the cascade model of diffusion—it originated in the larger cities of the Inland North core, and spread with some delay to the smaller and more distant communities of the Inland North fringe—while it was blocked from the Hudson Valley by phonological incompatibility.

A possible fault in this speculative scenario is that it suggests Southwestern New England ought to have had a continuous /æ/ system in the period when the Inland North was beginning to be settled, and the Hudson Valley ought to have had a nasal system or other discrete-allophonic system by the time the NCS

was beginning to spread; and there is no direct evidence that either of these was the case. The present-day Southwestern New England speakers in the Telsur corpus have /æ/ systems as nasal as the Hudson Valley's. Nor is there evidence in the data of Kurath & McDavid (1961) that Southwestern New England and the Hudson Valley had different /æ/ systems in the first half of the twentieth century. However, it is not implausible to suppose that the koineization scenario posited by *ANAE* might have led to a continuous /æ/ system even if none of the source dialects had a continuous system, inasmuch as the result of koineization tends to be phonologically unmarked; in earlier work (Dinkin, 2006) I provide a sketch of how something like this could have happened, though without reference to the continuous system per se.

The settlement of the Hudson Valley, as discussed above, was in large part derived from non-English-speaking populations—either the original Dutch settlers of New Netherland, or more recent Dutch and German immigrants. Indeed, Dutch and German remained principal languages in Hudson Valley communities into the early nineteenth century (cf. Platt, [1905] 1987 on Poughkeepsie; Campbell, 1906 on Oneonta); at the time of the Erie Canal's construction, English was a relatively new language to much of the Hudson Valley. An anonymous reviewer of this paper astutely asks whether Dutch and German substrates in the English phonology of those communities might have the effect of preventing NCS raising of /æ/. It is not clear to me what effect such a substrate might have had on /æ/, though it might be worth noting that Grand Rapids, Michigan—which, according to U.S. Census data, appears to have a greater proportion of Dutch ancestry among its population than any other major city in the country—has in the Telsur data as raised and continuous an /æ/ as any other Inland North core community. For now, however, it seems to be sufficient to note that the Hudson Valley and Inland North communities, in the early nineteenth century, were emerging from very different linguistic backgrounds, and it is easy to believe that it is as a result of this that they ended up with the structurally different /æ/ systems that leads the modern Hudson Valley to have resisted the NCS raising of /æ/ while accepting other NCS vowel shifts.

This scenario reconciles the two accounts of the "initial stages" of the NCS. As Thomas (2001) and McCarthy (2010) argue, the fronting of /o/ was the first step in the NCS, in the sense that it began earlier than any of the other sound changes thought of as being part of the NCS, before the divergence of the Inland North from Southwestern New England. On the other hand, as *ANAE* argues, the tensing of /æ/

was the triggering event of the NCS in the sense that that appears to be the change which uniquely distinguishes the NCS and the Inland North from the surrounding regions and their phonological systems.

It also, of course, resolves the conflicts between the accounts of the nature of the relationship between the Inland North and Western New England given by Boberg (2001) and in *ANAE*. Like *ANAE*, this paper contends that the NCS raising of /æ/ is a unique phonological feature that is distinct from the phonology of Southwestern New England, and could not have happened in an area that did not have the demographic history of New York State. However, Southwestern New England is essential to the history of the NCS, to the extent that communities in central and northern New York that were not settled from Southwestern New England did not develop it, even if in other respects they resemble the communities that did. Where Boberg's analysis seems to predict a gradual boundary between the Inland North and Southwestern New England, and the *ANAE* analysis seems to predict a sharp boundary, this paper predicts a null boundary: the Inland North and Southwestern New England do not actually meet, but are separated by the Hudson Valley. However, few phonological differences are observed between the Hudson Valley and Southwestern New England, none of them very large or statistically very robust; from that point of view, the Hudson Valley can be considered to be dialectologically united with Southwestern New England in the present day.[22] In that respect, the key feature distinguishing the Inland North from the Hudson Valley/Southwestern New England region is the tensing of /æ/, and the boundary appears to be more or less gradual: Between the Inland North core and Hudson Valley is the Inland North fringe, where /æ/ is certainly higher than in the Hudson Valley, but less homogeneously so than in the Inland North core.

8. Conclusion and future directions for research

To sum up, the key dialectological findings of this paper are as follows:

- The NCS is found in communities a great deal further north and east in New York State than previously observed; however, it is less frequent and less complete in these communities (the Inland North "fringe") than in the previously studied Inland North core communities.
- At least in central, eastern, and northern New York, the NCS is only present in communities whose early settlers were predominantly migrants from Southwestern New England. The persistence of the early-nineteenth-century settlement patterns in the present-day linguistic boundaries is striking; however:

- In Upstate New York, even communities which do not have the NCS show some features typically associated with it, such as backing of /e/ and modest fronting of /o/, although to a lesser extent than in New York's Inland North fringe or core communities. However, substantial raising of /æ/, the most distinctive hallmark of the NCS, is not present in such communities.

These findings are interpreted as indicating that the fronting of /o/ originated in Southwestern New England and was brought into Upstate New York by the settlers of the Inland North, but the raising of /æ/ originated later, within the Inland North. As suggested by *ANAE*, the preconditions for this NCS /æ/-raising are a result of the population and economic growth of the region brought by the construction of the Erie Canal. The raising of /æ/ failed to successfully spread beyond the Inland North fringe, while other NCS features such as fronting of /o/ and backing of /e/ succeeded in expanding southeastward into the region designated here as the Hudson Valley. The key difference in phonological structure between the Hudson Valley and the Inland North that prevented the eastward spread of /æ/-raising, according to my findings in Dinkin (2011a), is the difference between "nasal" and "continuous" patterns of /æ/ allophony, as defined in *ANAE*.

These findings also reaffirm the primacy of early settlement history, rather than present-day communication patterns, in determining the location of the boundaries of major dialect features, with only minor alteration around the edges in places like Cooperstown and Sidney.

This paper raises the question of whether the dialect boundary between the Hudson Valley and Southwestern New England persists to the present day and manifests noticeable phonetic or phonological differences between the two regions. This question will have to wait for a more substantial corpus of phonetic data from Southwestern New England, in order for meaningful statistical comparisons with the Hudson Valley to be made. Another area this paper suggests would benefit from additional data collection is St. Lawrence County, the area surrounding Ogdensburg and Canton. Data from more communities in that area (perhaps with clearer records of settlement history than are available for Ogdensburg) would help isolate which factors really determine the presence or absence of the NCS in far northern New York—a question left somewhat vague in this paper. Finally, more in-depth sociolinguistic studies of Sidney and Cooperstown could illuminate the motivations and trajectory of those villages' retreat from the NCS in greater detail than has been possible here.

Notes

[1] I use the notation of *ANAE* for phonemic vowel classes.

[2] See Gordon (2000:16–17, 80–83) for discussion of the two different trajectories that the movement of /e/ can take in the NCS.

[3] There is no clear sign of the NCS in Kurath & McDavid's maps either. However, it is fairly clear from Thomas (1935, 1936) that the NCS already existed in New York state by the time the LAMSAS fieldwork reported by Kurath & McDavid had begun; to speculate on why no sign of it appears in LAMSAS data would be beyond the scope of this paper.

[4] Only two nonwhite speakers were interviewed in the course of this project—both African-American women from Poughkeepsie. They are not analyzed in this paper due to lack of a baseline of comparison. Apart from Poughkeepsie, all cities and towns sampled in this study were, as of the 2000 Census, less than 13 percent African-American by population and more than 79 percent white.

[5] All names used in this paper for individual speakers are pseudonyms.

[6] This particular token of /e/ exhibits lowering but not backing, as shown in Figure 2.

[7] Of course, if /o/-fronting is considered to be one of the most important components of the NCS, this over-counting might be considered a feature rather than a bug.

[8] See Dinkin (2011b) for more detail on the *caught-cot* merger in this sample.

[9] Of course, the two factors here distinguishing Northwestern from Southwestern New England—*caught-cot* merger and a lower rate of satisfying NCS criteria—are not independent. Several NCS criteria have to do with the frontness of /o/; a speaker who merges /o/ with /oh/ is more likely to have /o/ backed than one who makes the distinction.

[10] Anecdotally, some middle-aged natives of Cooperstown spoken to in the course of this research who declined to participate in a recorded interview seemed, impressionistically, to exhibit relatively strong NCS features. Although these speakers are, obviously, not included in the data presented in this paper, they suggest that the speaker from Cooperstown scoring four in Figure 8 is not merely an outlier.

[11] There is weak evidence pointing toward change in apparent time in a few of the communities in the sample apart from Cooperstown and Sidney: Ogdensburg toward increasing NCS score, and Plattsburgh, Poughkeepsie, and Oneonta toward decreasing score. In Plattsburgh, the trend clearly seems to be the result of movement toward the *caught-cot* merger rather than an actual change in NCS status; in the others the seeming trend is neither statistically robust nor

categorical, so little can be said about them, and the discussion is omitted for the sake of brevity.

[12] All of these F ratios are statistically significant at the $p < 0.001$ level or better—that is to say, there are significant differences between communities in all five gradient NCS criteria.

[13] Obviously p-values just barely over 0.05 do not demonstrate that there is no real difference between communities. They do, however, indicate that if there is a real difference between communities, it is likely to be a relatively small difference compared with those that do achieve significance on datasets of similar size.

[14] The standard used for significance here is $p < 0.01$ instead of $p < 0.05$ because fifteen t-tests must be carried out to search for significant differences among six communities; a large number of t-tests increases the probability of p being <0.05 accidentally.

[15] By contrast, the range of NCS scores for Utica is higher than even the highest-scoring fringe city—from three to five rather than from two to four—and is different at the $p < 0.02$ level from both Ogdensburg and Watertown.

[16] For example, the Time Warner Cable web site at *timewarnercable.com* lists almost the exact same set of channels available in Gloversville as in Amsterdam; all of the broadcast channels listed are licensed to Schenectady, Albany, or points even further east, except for one local station licensed to Gloversville.

[17] By contrast, all but two said they very rarely or never go to Utica, the next closest larger city and the nearest known Inland North core community.

[18] Despite the spelling, the name "Dutchess" has nothing to do with the Dutch; the county was named by the English in honor of the Duchess of York.

[19] *http://www.oneonta.ny.us/oneonta/historic.asp*, viewed on 21 December, 2008.

[20] The difference between Southwestern New England and the /o/ ~ /oh/–distinct *ANAE* regions is significant at $p < 0.05$; between the Hudson Valley and the distinct regions, $p < 10^{-4}$.

[21] The Hudson Valley and the Inland North fringe differ at $p \approx 0.006$, and the Hudson Valley and Southwestern New England at ($p < 0.05$).

[22] Labov (2007) reports a variant of the New York City /æ/ system in Albany, and Dinkin & Labov (2007; see also Dinkin, 2009) found the same in one of three speakers from Schenectady. A few of the speakers from Poughkeepsie in the present study (cf. Dinkin, 2009; Labov, 2010) appear to display the diffused New York City /æ/ system as well. This /æ/ system is absent in Oneonta and Amsterdam and therefore is not sufficient to constitute a phonological difference between Southwestern New England and the wider Hudson Valley region as defined in this paper, although it may hint at a lower-level dialect division within the Hudson Valley.

References

Ash, Sharon. 2002. The distribution of a phonemic split in the mid-Atlantic region: Yet more on short A. *U. Penn Working Papers in Linguistics* 8(3): 1–15.

Bigham, Douglas. 2007. Vowel variation in southern Illinois. *American Dialect Society (ADS)*. Anaheim, CA.

Boberg, Charles. 2000. Geolinguistic diffusion and the U.S. Canada border. *Language Variation and Change* 12(1): 1–24.

Boberg, Charles. 2001. The phonological status of Western New England. *American Speech* 76: 3–29.

Brown, William H. (ed.). 1963. *History of Warren County, New York*. Glens Falls, NY: Board of Supervisors of Warren County.

Campbell, Dudley M. 1906. *A history of Oneonta*. Oneonta, NY: G.W. Fairchild.

Cooper, James F. 1886 [1838]. The chronicles of Cooperstown. In Shaw, S.M. (ed.), *A centennial offering: Being a brief history of Cooperstown*, 9–61. Cooperstown, NY: The Freeman's Journal.

Dinkin, Aaron J. 2006. Unnatural classes and phonological generalization in dialect formation. *New Ways of Analyzing Variation (NWAV) 35*. Columbus, OH.

Dinkin, Aaron J. 2009. *Dialect boundaries and phonological change in Upstate New York*. Philadelphia, PA: University of Pennsylvania dissertation.

Dinkin, Aaron J. 2011a. Nasal short-*a* systems vs. the Northern Cities Shift. *Penn Working Papers in Linguistics* 17(2): 71–80.

Dinkin, Aaron J. 2011b. Weakening resistance: Progress toward the low back merger in New York State. *Language Variation and Change* 23(3): 315–345.

Dinkin, Aaron J. 2012. Cooperstown, New York as a site of new-dialect formation. *New Ways of Analyzing Variation (NWAV) 41*. Bloomington, IN.

Dinkin, Aaron & William Labov. 2007. Bridging the gap: Dialect boundaries and regional allegiance in Upstate New York. *Penn Linguistics Colloquium (PLC)*, 31, Philadelphia.

Durant, Samuel W. 1878. *History of Oneida County, New York*. Philadelphia: Everts & Fariss.

Donlon, Hugh P. 1980. *Amsterdam, New York: Annals of a mill town in the Mohawk Valley*. Amsterdam, NY: Donlon Associates.

Emerson, Oliver F. 1891. *The Ithaca dialect: A study of present English*. Boston: J. S. Cushing & Co.

Farquhar, Kelly Y. & Scott G. Haefner. 2006. *Amsterdam*. Charleston, SC: Arcadia.

Frothingham, Washington (ed.). 1892a. *History of Fulton County*. Syracuse, NY: D. Mason.

Frothingham, Washington (ed.). 1892b. *History of Montgomery County*. Syracuse, NY: D. Mason.

Glens Falls Historical Association. 1978. *Bridging the years: Glens Falls, New York, 1763–1978*. Glens Falls, NY: Glens Falls Historical Association.

Gordon, Matthew J. 2000. *Small-town value and big-city vowels: A study of the Northern Cities Shift in Michigan. Publications of the American Dialect Society* 84. Durham, NC: Duke University Press.

History of Delaware County. 1880. New York: W.W. Munsell. http://www.dcnyhistory.org/munsell.html (8 January, 2009).

Hough, Franklin B. 1853. *A history of St. Lawrence and Franklin Counties, New York, from the earliest period to the present time*. Albany, NY: Little & Co.

Hough, Franklin B. 1854. *A history of Jefferson County in the state of New York*. Albany, NY: Joel Munsell.

Hurd, Duane H. 1880. *History of Clinton and Franklin Counties, New York*. Philadelphia: J.W. Lewis.

Hyde, Louis F. 1936. *History of Glens Falls*. Glens Falls, NY: Glens Falls Post Company.

Johnson, Daniel E. 2007. *Stability and change along a dialect boundary: The low vowels of southeastern New England*. Philadelphia, PA: University of Pennsylvania dissertation.

Kurath, Hans. 1939. *Handbook of the linguistic geography of New England*. Providence, RI: Brown University.

Kurath, Hans. 1949. *A word geography of the eastern United States*. Ann Arbor, MI: University of Michigan Press.

Kurath, Hans & Raven I. McDavid Jr. 1961. *The pronunciation of English in the Atlantic states*. Ann Arbor, MI: University of Michigan Press.

Labov, William. 2003. Pursuing the cascade model. In Britain, D. Cheshire & J. Cheshire (eds), *Social dialectology: In honour of Peter Trudgill*, 9–22. Philadelphia: John Benjamins.

Labov, William. 2007. Transmission and diffusion. *Language* 83(2): 344–387.

Labov, William. 2010. *Principles of linguistic change, volume 3: Cognitive and cultural factors*. Malden, MA.: Wiley/Blackwell.

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

McCarthy, Corrine. 2010. The Northern Cities Shift in real time: Evidence from Chicago. *Penn Working Papers in Linguistics* 15(2): 101–110.

Merriam, Nellie. 1907. The first settlement of Ogdensburg. In Swe-Kat-Si-Chapter (ed.), *Reminiscences of Ogdensburg, 1749–1907: Daughters of the American Revolution*, 1–17. New York: Silver, Burdett & Co.

Murray, David (ed.). 1898. *Delaware County, New York: History of the century, 1797–1897*. Delhi, NY: W. Clark. http://www.dcnyhistory.org/murray17971897.html (8 January, 2009).

Platt, Edmund. 1987 [1905]. *The Eagle's history of Poughkeepsie: From the earliest settlements 1683 to 1905*. Poughkeepsie, NY: Dutchess County Historical Society.

Roberts, Ellis H. 1911. Utica: Its history and progress. In Larned, J. N., *A History of Buffalo, vol. II.*, 257–291. New York: The Progress of the Empire State Company.

Ryan, Mary P. 1983. *Cradle of the middle class: The family in Oneida County, New York, 1790–1865*. New York: Cambridge University Press.

Thomas, Charles K. 1935. Pronunciation in Upstate New York (III). *American Speech* 10(4): 292–297.

Thomas, Charles K. 1936. Pronunciation in Upstate New York (IV). *American Speech* 11(1): 68–77.

Thomas, Erik R. 2001. *An acoustic analysis of vowel variation in New World English. Publications of the American Dialect Society* 85. Durham, NC: Duke University Press.