Fall 12-22-2009

The permeability of dialect boundaries: a case study of the region surrounding Erie, Pennsylvania

Keelan Evanini
University of Pennsylvania, keelan2@babel.ling.upenn.edu

Recommended Citation

Evanini, Keelan, "The permeability of dialect boundaries: a case study of the region surrounding Erie, Pennsylvania" (2009). *Publicly Accessible Penn Dissertations*. 86.
http://repository.upenn.edu/edissertations/86

The permeability of dialect boundaries: a case study of the region surrounding Erie, Pennsylvania

Abstract

This dissertation presents a dialectological study of the city of Erie, Pennsylvania, and the neighboring towns in the boundary area between the North and Midland dialect regions. The field work conducted for this dissertation consists of interviews, word lists, minimal pair tests, and grammatical acceptability judgments. In total, data from 106 speakers was analyzed to determine the course of linguistic change in the city of Erie and the current location of the dialect boundaries in the neighboring regions. In order to process the acoustic data from this large corpus, the methodology of transcription and subsequent forced alignment was applied. In order to reduce error in the formant measurements, automatic techniques for measurement point selection and formant prediction were developed. The acoustic analysis focuses on aspects of the vowel system that differentiate the North and the Midland. The results show that the merger of /o/ and /oh/ began in the city of Erie before 1900, and that it subsequently spread to Ripley, NY. On the other hand, Erie is still located on the Northern side of the boundary with respect to the fronting of the back upgliding vowels /uw/, /ow/, and /aw/. Finally, an analysis of the lexical and morphosyntactic variables shows a widespread acceptability of the Midland features in Erie. In the final section of the dissertation, the early settlement history of the region is examined, and Erie's acceptance of several Midland features is explained by the early presence of a large contingent of non-Northern, especially Scots-Irish, settlers.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Linguistics

First Advisor
William Labov

Second Advisor
Mark Liberman

Third Advisor
Jiahong Yuan

Subject Categories
Other Linguistics

THE PERMEABILITY OF DIALECT BOUNDARIES:

A CASE STUDY OF THE REGION SURROUNDING ERIE, PENNSYLVANIA

Keelan Evanini

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania

in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2009

Supervisor of Dissertation

William Labov, Professor, Linguistics

Graduate Group Chairperson

Eugene Buckley, Associate Professor, Linguistics

Dissertation Committee

Jiahong Yuan, Assistant Professor, Linguistics

Mark Liberman, Professor, Linguistics

Acknowledgements

I would like to thank my advisor, Bill Labov, for being an excellent teacher, and always being willing to provide advice and encouragement. The two other members of my committee, Mark Liberman and Jiahong Yuan, also deserve warm thanks for always being available and providing useful feedback. Finally, I would like to thank all of the remaining professors in the department for what they have taught me throughout my years at Penn. The spirit of generosity and collaboration exhibited by the professors in the linguistics department is second to none.

I gratefully acknowledge the financial and admistrative support provided by the Univerisy of Pennsylvania, the School of Arts and Sciences, the Graduate and Professional Student Assembly, the Graduate Student Associations Council, and the Department of Linguistics. In the linguistics department, the administrative support provided by Amy Forsyth and Sue Sheehan was invaluable.

The other graduate students in the department have been a constant source of support and inspiration. I would especially like to thank the other students in my dissertation year for their continuous companionship: Jonathan Gress-Wright, Laia Mayol, Jean Francois Mondon, Joshua Tauberer, and Lucas Champollion (who belongs to this group in spirit, although his official year of entry into the program is one year later than ours). Many other students also provided valuable discussion and helpful examples of excellent scholarship throughout the years, including (but not limited to): Lukasz Abramowicz, Suzanne Evans

Wagner, Michael Friesner, Josef Fruehwald, Damien Hall, Daniel Ezra Johnson, Catherine Lai, Laurel MacKenzie, James Mesbur, Maya Ravindranath, and Jonathan Wright. Aaron Dinkin deserves special mention as a constant source of inspiration and collaboration. Also, special thanks go to Tatjana Scheffler and Marjorie Pak for providing support as the other new parents in the program.

As is the case for any study involving field work, my research would not have been possible without the participation of my informants. I would like to thank all of them for graciously volunteering their time by participating in the study and sharing their stories with me. I would also like to thank the Linguistic Data Consortium for allowing me to use their call collection interface, and Chris Cieri and Kevin Walker for their invaluable support in setting it up.

Finally, the completion of this dissertation would never have been possible without the loving care and support provided by my entire family. I would especially like to thank Jamie and Gavin for keeping me focused on what is most important in life.

ABSTRACT

THE PERMEABILITY OF DIALECT BOUNDARIES: A CASE STUDY OF THE REGION SURROUNDING ERIE, PENNSYLVANIA

Keelan Evanini

Supervisor: William Labov

This dissertation presents a dialectological study of the city of Erie, Pennsylvania, and the neighboring towns in the boundary area between the North and Midland dialect regions. Erie occupies a unique place in the dialect geography of North America, in that it appears to have switched status from the North to the Midland. Since the dialect boundary between the North and the Midland has remained stable throughout the rest of North America, this switch presents an intriguing test case for theories of dialect change and phonological structure.

The field work conducted for this dissertation consists of interviews, word lists, minimal pair tests, and grammatical acceptability judgments. In addition, archival data was used to push the time depth of the analysis further back. In total, data from 106 speakers was analyzed to determine the course of linguistic change in the city of Erie and the current location of the dialect boundaries in the neighboring regions.

In order to process the acoustic data from this large corpus, the methodology of transcription and subsequent forced alignment was applied. This enabled the automatic extraction of 113,245 vowel formant measurements, an amount which would have been difficult to obtain using the standard sociophonetic procedure of manual formant extraction. In order to reduce error in the formant measurements, automatic techniques for measurement point selection and formant prediction were developed.

The acoustic analysis focuses on aspects of the vowel system that differentiate the North and the Midland. The results show that the merger of /o/ and /oh/ began in the city of Erie before 1900, and that it has subsequently spread to the town of Ripley, NY. Additionally, Erie speakers consistently have a nasal or continuous short-a system. On the other hand, Erie is still located on the Northern side of the boundary with respect to the fronting of the back upgliding vowels /uw/, /ow/, and /aw/. Finally, an analysis of the lexical and morphosyntactic variables shows a widespread acceptability of the Midland features in Erie.

In the final section of the dissertation, the early settlement history of the region is examined, and Erie's acceptance of several Midland features is explained by the early presence of a large contingent of non-Northern, especially Scots-Irish, settlers.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The Atlas of North American English (Labov et al. 2006), henceforth ANAE, represents the first comprehensive phonological study of the entire North American continent. It thus provides a detailed overview of the various dialect regions and the sound changes that are taking place in each one. ANAE, however, is a study of urban speech—the survey's methodology sampled two speakers from every Metropolitan Statistical Area with 50,000 or more inhabitants. Thus, while ANAE is able to precisely define the characteristics of the dialect regions of North America, it is not able to describe how the areas at the dialect boundaries look, since these boundaries normally lie in less populated geographic regions between urban areas.[1] The aim of this dissertation is to address this lack of coverage by studying the dialect boundary areas around a city of theoretical interest: Erie, Pennsylvania.

Erie holds a unique place in the dialectology of North America, since it is the only city to appear to have switched from the North to the Midland throughout the course of the 20th century. The earliest dialectological records of the region (Kurath 1949, Kurath and McDavid 1961) show Erie to pattern together with the North with respect to nearly all

---

[1] Except for relatively rare cases such as Windsor, ON and Detroit, MI, where the presence of a sharp dialect boundary between two neighboring cities (Labov et al. 2006:217) is due to the international border separating them.

lexical and phonological isoglosses. However, ANAE shows that the phonology of Erie is clearly no longer Northern, and shares two crucial phonological traits with the Midland: the merger of /o/ (as in *cot*) and /oh/ (as in *caught*) and a pattern of raising /æ/ before nasals (Labov et al. 2006:205).[2] This shift of allegiance from the North to the Midland is surprising, since the North is perhaps the most cohesive dialect region in North America—as evidenced by the high rate of homogeneity and consistency of its defining isoglosses (Labov et al. 2006:151)—and the boundary between the North and the Midland is one of the sharpest boundaries in North America.

The goals of this dissertation are threefold: empirical, methodological, and theoretical. First, a corpus of speech samples collected from Erie and the surrounding boundary regions provides both real time and apparent time evidence for linguistic changes that occurred in and around Erie over the past 120 years. This empirical research supplements the ANAE by describing the speech of less populated areas that have never been studied before and by determining the exact locations of the dialect boundaries around Erie. In addition to the empirical aims of this research, several methodological innovations were introduced to enable the automatic analysis of the large amount of acoustic data that was obtained through fieldwork around Erie. This application of methods commonly used in the field of automatic speech recognition to sociolinguistic research enables the analysis of a much larger amount of data in a much shorter time than would be possible with the traditional manual methods of analysis. Finally, I aim to address several theoretical questions in dialect geography through an analysis of the corpus. The following three sections provide more details about each of the aims of this dissertation.

---

[2] For a description of the vowel symbols used in this dissertation, see Appendix A.

1.1 Empirical aims

The most important empirical question that this dissertation addresses is the status of the vowels /o/ and /oh/. The fact that these two vowels are currently merged in Erie is the clearest diagnostic that Erie is no longer phonologically aligned with the North. The research conducted for this dissertation extends our knowledge about the merger in Erie in the dimensions of both time and space. The time depth of the evidence is increased by interviews with elderly Erie residents and archival data from speakers born before 1900. The geographical depth is extended by new data from several small towns between Erie and Buffalo. This data enables us to pinpoint the present location of the boundary between the merged and unmerged regions and to describe the current trajectory of the change.

On a larger scale, this research will contribute to our detailed knowledge of the progress of the merger of /o/ and /oh/ throughout North America. This change is quite likely the most widely studied phonological feature in American English, and detailed descriptions of its progress exist for many communities, including Eastern Pennsylvania (Herold 1990), the area around the border between Rhode Island and Massachusetts (Johnson 2007), Cooperstown, NY (Dinkin 2009), Charleston, SC (Baranowski 2006), West Virginia (Hazen 2005), Indianapolis, IN (Fogle 2006, 2008), Missouri (Majors 2005, Gordon 2006), Oklahoma (Bailey et al. 1993), Utah (DiPaolo 1992), and San Francisco, CA (Hinton et al. 1987, Moonwomon 1992, Hall-Lew 2009). As Labov (1994:316) states, the merger of /o/ and /oh/ is "the largest single phonological change taking place in American English." It is also the most important change from a structural standpoint, due to the large influence that the phonemic status of these two vowels has on the rest of the vowel system (Labov et al. 2006:119–151).

In addition to a documentation of the merger of /o/ and /oh/, this dissertation will also consider in detail two other aspects of the vowel system that sharply differentiate the North and the Midland: the raising of /æ/ and the fronting of back upgliding vowels. In the North, the raising and fronting of /æ/ in all environments was the triggering event for the Northern Cities Shift (Labov et al. 2006:192–195), whereas Midland speakers generally exhibit no tensing and fronting in unfavored environments (before voiceless obstruents). As is the case with /o/ and /oh/ the degree and type of raising exhibited by /æ/ has wide-ranging effects on the rest of the vowel system. Speakers from Erie pattern consistently with other Midland speakers with regard to /æ/, in that they either exhibit raising only before nasals or have a continuous system (Labov et al. 2006:180). A comprehensive study of the status of /æ/ in the boundary regions around Erie will contribute to our knowledge of how the status of /æ/ and the merger of /o/ and /oh/ are linked (Labov 1991, Labov et al. 2006).

Whereas the analyses of /o/, /oh/, and /æ/ show that Erie's phonology is aligned with Midland patterns, the analysis of the back upgliding vowels /uw/, /ow/, and /aw/ show that Erie still shares some phonological characteristics with the North. The fronting of this system of vowels is a pronounced feature of the Midland, but is either restricted in scope or completely absent from the speech of Northern speakers. In Erie, fronting of /uw/, /ow/, and /aw/ is also quite limited. An analysis of these three vowels in Erie and the surrounding boundary regions thus provides a more complete picture of the dialectological status of the area.

Finally, the methodology adopted for this dissertation enables the analysis of all vowel tokens produced by a speaker. Thus, with a combination of word list and interview data, it is possible to obtain detailed information about a speaker's entire vowel system. Individual analyses will be presented for the first two formants of each vowel in the form of natural break maps in an attempt to efficiently summarize the geographical patterns (or lack thereof) contained in the entire range of the empirical acoustic results.

1.2 Methodological aims

This dissertation employs a methodology for data analysis that has not been used before in dialect geography. After the interviews were collected, they were all transcribed in their entirety. Then, forced alignment, a procedure commonly used in the field of automatic speech recognition, was used to automatically determine the locations of the word and phoneme boundaries throughout the course of the audio file. This information was then used to facilitate the automatic extraction vowel formants for all vowels contained in the transcription. This methodology thus enables a large corpus of speech data to receive a detailed phonetic analysis in a short period of time. In contrast, sociophonetic analyses of vowel formants have traditionally relied on manual extraction of F1 and F2 data for each vowel token. This means that a large-scale analysis, such as the ANAE, takes several years to complete. For this dissertation, however, approximately the same number of vowel tokens were analyzed (around 125,000) as were analyzed for the ANAE. The total time required for the automatic analysis, even when the system development and programming time is taken into account, was substantially less than the time required for manual analysis.

Furthermore, after the process of forced alignment has been applied to a corpus, any word or phoneme of interest can be extracted easily (for either automatic or acoustic analysis), since word-level and phoneme-level time stamps are available for the entire corpus. After a manual vowel anaylsis, however, the same amount of time and labor required for the initial analysis would be necessary if the researchers desired to study another acoustic feature, such as vowel duration. Using forced alignment for sociophonetic corpus analysis is thus a good investment for future research, as well.

Of course, the techniques of forced alignment and automatic vowel analysis do introduce some errors into the data, most of which could be avoided through a manual analysis. To address this problem, a number of techniques for error reduction were employed for this dissertation, including an improved method of formant prediction. The fact that the acoustic analyses of vowels obtained through this methodology are consistent with prior manual analyses of speakers from the same locations demonstrates the validity of the automatic data extraction techniques.

A major aim of this dissertation is thus to introduce the methodology of forced alignment to the sociolinguistic community and advocate for its widespread adoption among researchers. If this were to happen, it would enable research projects to consider much larger pools of informants for the data collection stage of a project, since the analysis could be conducted quickly. Thus, findings based on social characteristics of a group of speakers would be more robust, since the number of speakers subject to analysis in each group would be much larger. Furthermore, if researchers were willing to share their data, the methodology of forced alignment would facilitate the replication of important findings and enable easy subsequent analyses of valuable corpora. As it currently is, however, if a researcher is interested in re-analyzing a corpus that has been analyzed already by a previous researcher, there is an extremely high barrier of entry. As an aid to publicizing the methodologies of forced alignment and automatic vowel analysis that were used for this dissertation, as well as facilitating their adoption by other researchers, all of the necessary software is available through the web site of the Phonetics Lab at the University of Pennsylvania.[3]

---

[3] The forced alignment software is available as part of the P2FA (Penn Phonetics Lab Forced Aligner) package, and the other software is available as part of P2TK (the Penn Phonetics ToolKit). Both sets of software are available at the following URL: *http://ling.upenn.edu/phonetics/*.

1.3 Theoretical aims

1.3.1 What causes the diffusion of linguistic change across boundaries?

One of the main theoretical questions facing dialect geographers is to explain the mechanism by which linguistic changes diffuse across dialect boundaries. The types of theories that have been put forth to explain the spread (or lack thereof) of a linguistic change from one dialect region to a neighboring one often make reference to the demographic situation that obtains in the two regions. For example, the Gravity Model (Trudgill 1974) proposed that the spread of sound change from one area to another is proportional to the population of the two areas, and inversely proportional to the square of their distances. In a similar approach, the Cascade Model (Labov 2003) proposed that changes spread from large cities to smaller ones, skipping over the sparsely populated areas in between. Both of these models are based on the idea that linguistic change is brought about through increased communication with speakers from another dialect region (Labov 1974). All of these types of explanations are ultimately based on the Principle of Density (Bloomfield 1933).

On the other hand, structural factors can also play a role in promoting or inhibiting the spread of a linguistic change across a dialect boundary. For example, the merger of /o/ and /oh/ would be expected to spread into all dialect regions in North America according to Herzog's Principle that "mergers expand at the expense of distinctions" (Labov 1994:313). However, other structural factors have inhibited its spread across certain boundaries, such as the fronting of /o/ as part of the Northern Cities Shift in the North. An example of a structural factor influencing the spread of a change is the correlated behavior of /ow/ and /ʌ/: the backing of /ʌ/ in the North appears to be influenced by the extreme back position of /ow/ in that region, whereas the fronting of /ʌ/ in the Midland and South appears to be influenced by the fronting of /ow/ (Labov et al. 2006:143).

The most important linguistic change to spread to Erie was the merger of /o/ and /oh/. Other studies that have examined areas of recent merger in depth have proposed explanations based on demographic shift to explain the merger, either through foreign immigration (Herold 1990) or through an influx of residents from a neighboring merged region (Johnson 2007). Using these earlier studies as a guide, this dissertation will attempt to determine the relative influence of social and structural factors on the spread of the merger of /o/ and /oh/ to Erie and its continuing diffusion to neighboring regions.

1.3.2 Are all dialect boundaries alike?

In conducting fieldwork around Erie, Pennsylvania, I examined three discontinuous dialect boundary regions arising from two separate dialect contact situations:

1. Erie vs. the North: This boundary exists in the two geographically distinct regions between Erie and Buffalo to the northeast of Erie and between Erie and Cleveland to the west. On the Erie side of this boundary, /o/ and /oh/ are merged, and there is no evidence of the Northern Cities Shift. On the North side of the boundary, /o/ and /oh/ remain distinct, and the NCS is represented in varying degrees. Figure 1.1 shows how Erie is located outside of most of the 8 isoglosses for the North in ANAE. Furthermore, the two that do include Erie are tenuous there, at best.[4]

2. Erie vs. the Midland / Western Pennsylvania: This boundary exists between Erie and the Midland region centered around Pittsburgh to the south of it. On the Erie

---

[4] The barred blue isogloss showing the outer limits of the North is defined by the three characteristics that are thought to be necessary conditions for the Northern Cities Shift: a short-*a* system that is not split into tense and lax class, no fronting of /ow/, and the lack of a merger between /o/ and /oh/. While the first two characteristics clearly hold for the two Erie ANAE speakers, one speaker provides very slight evidence that the merger is not complete for her: two out of the five minimal pairs were judged close in production, although the rest were judged to be exactly the same. The other speaker is clearly merged across the board in production and perception. The second Northern isogloss that is shown to include Erie Figure 1.1 is the ED line, inside of which F2 of /e/ is less than 375 Hz higher than F2 of /o/. However, the ANAE speaker from Erie who has these two vowels the closest actually has them separated by 388 Hz. So, Erie should not actually have been included inside of the ED isogloss.

side of this boundary, /ow/ and /aw/ are not being fronted as they are in much of the Midland. Furthermore, near Pittsburgh, on the southern side of the boundary, the Pittsburgh Shift is apparent in many speakers who show monophthongization of /aw/.

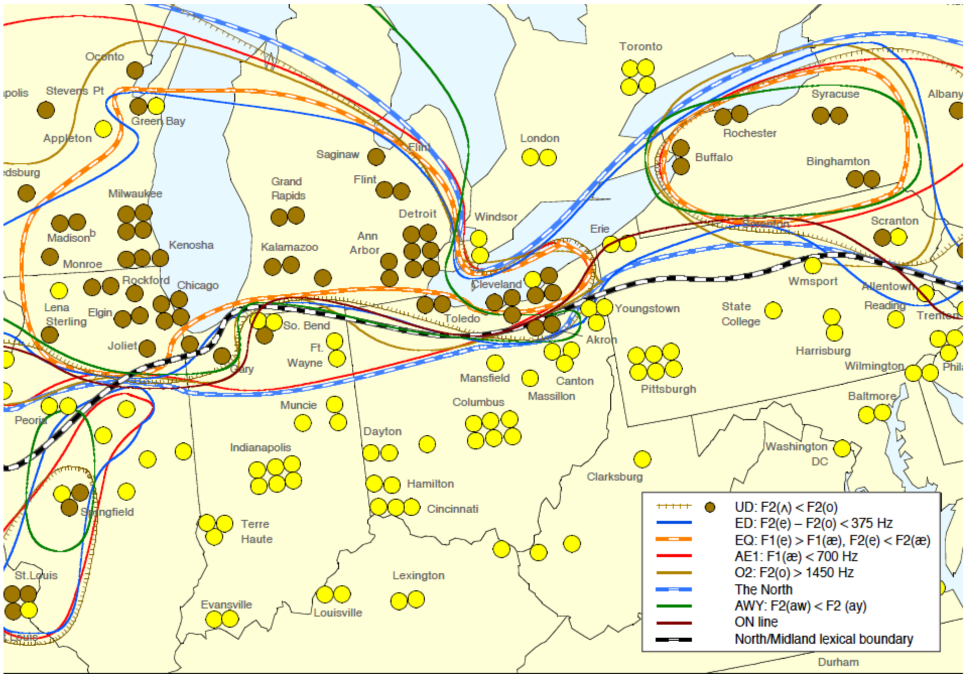

Figure 1.1: Erie and the surrounding region (Map 14.11 from ANAE)

One of the main goals of the fieldwork for this dissertation will be to determine the nature of these dialect boundaries in the areas of transition between the two regions. Several theoretical possibilities exist, based on the amount of overlap between the features of the two regions in the boundary area. For example, Chambers and Trudgill (1999:104) distinguish between *abrupt* and *gradual* transition areas. A slightly more refined taxonomy is presented by Dinkin (2009): *sharp*, *fading*, *overlapping*, and *null* boundaries. Research into the two boundary regions around Erie will determine what type of boundary exists in each area. It is hypothesized that the boundary between Erie and the North will be a sharper boundary, since the merger of /o/ and /oh/ prevents the other stages of the Northern Cities Shift from taking place. ANAE has already shown that the boundary between the North and the Midland consists of a bundle of several closely related isoglosses, and that this boundary is one of the sharpest in North America (p. 205). On the other hand, it is hypothesized that the boundary between Erie and the area to the south with strong fronting of /ow/ will be more gradual, since there are no structural barriers to the fronting of /ow/ in Erie.

1.3.3 What is the relationship between different types of isoglosses?

Chambers and Trudgill (1999) define six different types of isoglosses, based on the level of linguistic structure involved: *lexical*, *pronunciation*, *phonetic*, *phonemic*, *morphological*, and *syntactic*. Most studies that have attempted to define dialect regions to date have focused only on one or two of these types of isoglosses. For example, Carver (1987), Cassidy and Hall (1985–2002), and Kurath (1949) are all concerned with lexical isoglosses. On the other hand, Kurath and McDavid (1961) examines only pronunciation and phonetic isoglosses. ANAE is the most wide-ranging single study to date: its main focus is on phonetic and phonemic isoglosses, but it also collected some data on lexical, pronunciation, and syntactic isoglosses.

There has not been any systematic research into how the different types of isoglosses pattern differently, especially around dialect boundary areas. ANAE has shown how phonetic and phonemic isoglosses bundle together along dialect boundaries in cases of large-scale sound shifts with many related components, such as the Northern Cities Shift and the Southern Shift. However, there has been less research into the correlation of isoglosses from levels of linguistic structure that are not related. One such case that has been investigated is the close correlation between the lexical boundary dividing the North and the Midland in Kurath (1949) and the phonological boundary separating the two regions in

ANAE (Labov et al. 2006:205). ANAE showed that Erie was the only city to switch from the Northern side of the lexical boundary to the Midland side of the phonological boundary. However, it is not known whether this shift in phonological status also coincided with a shift to Midland features in other levels of linguistic structure. Since many of the lexical items surveyed in (Kurath 1949) are now obsolete, I will investigate the relationship between the phonological isoglosses around Erie and a few syntactic isoglosses that are characteristic of the Midland areas with heavy original Scots-Irish settlement, specifically positive *anymore* and *need* + Past Participle.

1.4 Outline

This dissertation is organized as follows: Chapter 2 will first review the phonological and lexical evidence from earlier linguistic atlas projects that Erie's original linguistic patterns were Northern. Next, Chapters 3 and 4 will discuss the methodology used to obtain and analyze linguistic data for this dissertation. Chapter 3 presents details about how individual speakers were selected, and described the process of data collection (interview and transcription). Chapter 4 describes the technical details of the forced alignment and vowel analysis procedures that were employed to automatically obtain vowel measurements. Chapter 5 presents natural break maps of the entire region under analysis for the F1 and F2 values for every vowel. The next two chapters contain detailed studies of a few crucial phonological features. Chapter 6 is a study of the merger of /o/ and /oh/ in the region, and Chapter 7 investigates the other vowels that differentiate the North and the Midland. Chapter 8 then presents data from lexical and morphosyntactic variables that differentiate the North and the Midland, and shows which side of the relevant boundaries Erie speakers fall on. Finally, Chapter 9 presents data relating to the settlement history of Erie in an attempt to shed light on the demographic causes of the observed linguistic patterns.

Chapter 2

Erie's Original Status as a Northern City

2.1 Introduction

All of the earliest dialectological studies indicate that Erie was aligned with the North for at least the early part of the 20th century. The evidence for this based on lexical items is quite strong, and comes primarily from *A Word Geography of the Eastern United States* (Kurath 1949), henceforth *WG*, as well as the data collected for the *Dictionary of American Regional English*, henceforth DARE, as presented in Carver (1987). The sources for the phonological evidence are *The Pronunciation of English in the Atlantic States*, henceforth *PEAS*, (Kurath and McDavid 1961) and Wetmore (1959), both based on the fieldwork done for the *Linguistic Atlas of the Middle and South Atlantic States*, henceforth *LAMSAS*.

2.2 Lexical

WG provides two types of evidence for Erie's position within the North. First of all, Erie is located within 10 of the 11 defining isoglosses of the North (Maps 5–8 in *WG*) and all 6 isoglosses that are characteristic of both the North and the Midland (Maps 39 and

40). Table 2.1 lists these words along with their non-Northern counterparts, showing the Northern version used in Erie in italics.

Northern form	non-Northern form	map in WG
whiffletree, whippletree	swingletree	5a
pail	bucket	5a
darning needle	dragonfly	5a
teeter, teeterboard	seesaw	5b
stone boat	vehicle for dragging field stones	5b
spider	frying pan	6
skaffle	scaffold (in a barn)	6
buttry	pantry	6
stoop	porch	7
dutch cheese	cottage cheese	8
stone wall	fence built of loose stone	39
hay mow	hay loft	39
grist of corn	turn of corn	39
whinny	nicker / whicker	40
corn husks	corn shucks	40
string beans	snap beans	40

Table 2.1: Northern isoglosses in *WG* that contain Erie (italicized variant used in Erie)

Table 2.2 shows that *belly-gut* is the only one of the 11 Northern lexical items that does not contain Erie.

Northern form	non-Northern form	map in WG
belly-gut	*face-down on a sled*	7

Table 2.2: Only Northern isogloss in *WG* that does not contain Erie

Secondly, evidence for Erie's original status as a Northern city is provided by the Midland isoglosses in *WG*. Erie falls outside of 8 isoglosses that define the Midland (Maps 15-18) and 5 isoglosses that are characteristic of both the South and the Midland, all of which reach northward past Pittsburgh (Maps 41 and 42) . In Table 2.3 I refer to these two types of isoglosses as the *non-Northern isoglosses*, in contrast to the isoglosses in Table 2.1 that were either distinctly Northern or shared by the North and Midland.

non-Northern form	Northern form	map in WG
I want off	*I want to get off*	15
Sook!	*call to cows*	15
snake feeder	*dragonfly*	15
blinds	*roller shades*	16
bawl	*noise a calf makes*	16
poke	*paper bag*	17
sugar tree	*maple tree*	17
worm fence	*a rail fence laid zigzag*	18
corn pone, pone	*cornbread*	41
paling fence, pale fence	*picket fence*	41
roasting ears	*sweet-corn*	41
pole cat	*skunk*	42
Christmas gift!	*Merry Christmas!*	42

Table 2.3: Non-Northern isoglosses in *WG* that do not contain Erie (italicized variant used in Erie)

Conversely, Erie shares only two of the lexical items characteristic of the Midland: *run* and *smear case* (used to define the North Midland in Map 18), and one of the items common to both the Midland and South: *spicket*. These three lexical items that Erie shares with the Midland are shown in Table 2.4.

non-Northern form	Northern form	map in *WG*
run	a small stream	18
smear case, smear cheese	cottage cheese	18
spicket	faucet	42

Table 2.4: Only non-Northern isoglosses in *WG* that contains Erie

Thus, Erie behaves like a Northern city for 88% (29 out of 33) of the relevant lexical items from *WG*. Finally, Erie is also situated outside of the three isoglosses that Kurath provides for Western Pennsylvania (Map 25 in *WG*). These three extend northward from

Pittsburgh into northwestern PA, but none of them quite reach Erie. This suggests that Pittsburgh's influence in Erie was not yet strong at that time. These three lexical items are show in Table 2.5.

Western PA form	non-Western Pa form	map in WG
hap	*quilt*	25
doodle, hay doodle	*haycock*	25
drooth	*drought*	25

Table 2.5: Three Western Pennsylvania isoglosses that do not contain Erie

The evidence from DARE is available only indirectly through Carver (1987), who trolled through the DARE fieldwork data to compile maps that capture the regional patterning of some of the lexical items used in the survey. Instead of the more traditional concept of dialect region, Carver prefers to use dialect *layers* as his descriptive apparatus. He defines a dialect layer as "the composite of a unique set of areal isoglosses, the geographical spread of its lexicon" Carver (1987:16). This concept of dialect layer is useful for the lexicon (which, in contrast to the phonology, does not have structural relationships between its elements), because it does not force the researcher to posit discrete boundaries, and thus enables the landscape to be viewed more as a continuum. Any given community can belong to several different layers, each having a different strength at that location, based on the number of items from that layer used in the community. The different layer strengths thus provide information about how strongly that community is affiliated with each dialect region.

The evidence from the maps relevant to Erie is presented in Table 2.6. The first column in the table represents the name of the dialect layer, as defined by Carver. The next two columns represent the number of DARE isogloss terms for the layer that occur in Erie, and the total number of DARE isoglosses used to define the layer, respectively. Unfortunately, even though Carver does provide lists of all of the isogloss terms he used to define the layers, there is no way to know, without consulting the original fieldwork data, which of them occur in any given geographical point—due to space limitations on the maps, Carver only depicts the number of terms, not the specific terms themselves. So, a direct comparison with the distribution of the words from WG in Tables 2.1–2.5 is not possible. The fourth column in Table 2.6 shows the strength of the boundary within which Erie is situated for each of the dialect layers. The possible types of boundaries are *primary*, *secondary*, *tertiary*, and *quaternary*, with a primary boundary containing the area where the highest percentage of terms for the layer are found. Finally, the fifth column shows the label of the corresponding map from Carver (1987).

Dialect Layer	# of terms in Erie	total # of terms for layer	boundary strength	map from Carver
North	33	82	primary	3.3
Upper North	20	62	secondary	3.7
Inland North	18	51	primary	3.9
Midland	1	40	N / A	6.5
Lower North	4	53	N / A	6.15

Table 2.6: Erie's position with regard to dialect layers in Carver (1987)

The three boundaries for the layers of the North provide good evidence for Erie's status as a Northern city at the time of the DARE fieldwork. Erie is located within primary boundaries for the North and Inland North layers; furthermore, Erie falls just outside of the primary boundary for the Upper North, which extends westward to Lake Erie, stopping just at the Pennsylvania-New York state line. On the other hand, the data provide very little evidence for associating Erie with the Midland. Map 6.15 situates Erie outside of the tertiary boundary for the Midland layer (which extends northward to around Youngstown, OH). Similarly, Erie falls outside of the two boundaries provided for the Lower North[1] layer, although the secondary boundary does stretch northward almost to Erie.

---

[1] Carver's Lower North layer corresponds closely to what is traditionally referred to as the North Midland. Carver's terminology reflects his disbelief in the existence of a separate Midland dialect region.

2.3 Phonological

Table 2.7 presents the features mapped in PEAS that are evidence for Erie's original affiliation with the North. For all of these features, the isogloss falls just south of Erie, indicating that Erie was always just on the edge of the boundary between the North and the Midland. The first column in Table 2.7 describes the feature that Erie shares with the North, and lists the lexical items that PEAS uses to illustrate this.[2] The second column describes the contrasting feature that is found just south of Erie. In some cases, this feature is widely distributed throughout the Midland region (e.g., /uw/ in *due*, Map 163), whereas in others the contrasting feature is more characteristic of Western Pennsylvania in particular (e.g., [druθ] for *draught*, Map 142). In either case, there is a clear boundary between Erie to the north and the area of Western Pennsylvania surrounding Pittsburgh to the south. The speakers from two counties immediately south of Erie appear to be transitional for many of these features, with Crawford County aligning more frequently with Western PA, and Warren County aligning more frequently with the North.

The large number of features that Erie shares with the North can be contrasted with the features in Table 2.8. These are the only three listed in PEAS that have Erie aligned with the Midland or Western Pennsylvania in opposition to the North.

Even more probative of Erie's original phonological alignment with the North is the status of the low-back vowels. This (along with the status of /ae/) is one of the two main structural features that determine the status of a dialect of present-day North American English (Labov 1991). The fact that /o/ and /oh/ are kept distinct by the two PEAS speakers from Erie proves the regions's original alignment with the North as opposed to Western Pennsylvania, where the two phonemes are merged as a low, back rounded vowel ([ɒ] in the PEAS notation).

---

[2] The isoglosses providing evidence for the low-back vowels are omitted from this table, since they are discussed separately below.

Erie feature (shared with the North)	contrasting Midland / Western PA feature	PEAS Map #
/i/ monophthongal in *crib*	ingliding diphthong [ɪᵊ]	4
/e/ monophthongal in *bed*	ingliding diphthong [ɛᵊ]	4
/ey/ more close (i.e. [eɪ]) in *day*, *bracelet*	[ɛɪ]	18, 19
non-fronted /ow/ in *ago*, *coat*	fronted to [ɜʊ]	20, 21
raised nucleus for /ay/ in *nine*	not raised	26
fronted [ɛʊ] in *mountain*, *(worn) out*	[aʊ ∼ ɑʊ]	28, 29
unrounded [a] in *father*	rounded [ɒ]	32
/iw/ in *dues* and *tube*	/uw/	33
vowel in *four* and *forty* distinct	merged	44
[e] before /r/ in *married*	[æ]	51
[ʊ] in *root*	[u]	113
[ʌ ∼ ɵ] in won't	[o]	125
[drauθ] for *draught*	[druθ]	142
/iw/ in *blue*, *chew*, and *suit*	/uw/	147
[ɪ] in final unstressed syllable of *careless*, *houses*, *haunted*, and *bucket*	[ə]	148
/iw/ in *due*, *new*, and *Tuesday*	/uw/	163, 164, 165
yeast pronounced as [jist]	[ist]	166
/ð/ in *without*	/θ/	170
/s/ in *greasy*	/z/	171

Table 2.7: Northern phonological isoglosses that contain Erie in *PEAS*

Table 2.9 presents the /o/ and /oh/ words that are mapped in PEAS. First of all, the merger in vowel quality is shown clearly for the Western PA speakers: they have a rounded low-back vowel for all tokens. Interestingly, however, the atlas does mark a distinction in length: the three /oh/ words (*law*, *salt*, and *dog*) have a half-length mark, and are also characterized by the possible presence of a central offglide. The Northern speakers, on the other hand maintain a clear distinction in quality between the two types. The /o/ words have an unrounded low-central vowel, possibly slightly fronted in *oxen*, whereas the /oh/ words all have a consistently low-back rounded vowel.

Erie feature (shared with the Midland / Western PA)	contrasting Northern feature	PEAS Map #
/i/ is ingliding [ɪᵊ] in *whip*	[ɪ]	5
/e/ is ingliding [ɛᵊ] in *fence*	[ɛ]	9
trisyllabic pronunciation of *mushroom* ending in /n/	disyllabic ending in /n/	177

Table 2.8: Midland / Western PA phonological isoglosses that contain Erie in *PEAS*

Lexical Item	Erie	North	Western PA	Map #
oxen	[ɑ ∼ ɑ̙ ∼ a̘]	[ɑ ∼ ɑ̙ ∼ a̘]	[ɒ ∼ ɔ]	15
wash	[ɑ]	[ɑ]	[ɒ ∼ ɔ]	135
fog	[ɑ]	[ɑ]	[ɒ ∼ ɔ]	136
on	[ɑ]	[ɑ]	[ɒ ∼ ɔ]	138
law	[ɒˑ ∼ ɒˑᵊ]	[ɔˑ ∼ ɔˑᵊ]	[ɒˑ ∼ ɒˑᵊ]	22
salt	[ɒˑ ∼ ɒˑᵊ]	[ɔˑ ∼ ɔˑᵊ]	[ɒˑ ∼ ɒˑᵊ]	23
dog	[ɒˑ ∼ ɒˑᵊ]	[ɔˑ ∼ ɔˑᵊ]	[ɒˑ ∼ ɒˑᵊ]	24

Table 2.9: /o/ and /oh/ words in PEAS in Erie, the North, and Western PA

Wetmore (1959) reaches the same conclusion based on a larger body of evidence from the LAMSAS fieldnotes. In addition to the data presented in PEAS, he examined the lexical items *pot*, *fought*, *shock*, *god*, *off*, *cloth*, *sauce*, *costs*, *frost*, *all*, *John*, *gone*, *launch*, *strong*, *saw*, *swamp*, and *wasp* (Wetmore 1959:109) for speakers from Western Pennsylvania (although it is unclear exactly which of these had data from the two Erie speakers). Based on this evidence he lists both Erie informants as having a distinction between /ɑ/ and /ɔ/ (Wetmore 1959:113).

Thus, it is clear from the earliest survey data available that Erie's linguistic original linguistic affiliation was with the North. Both the lexical data from WG and DARE as well as the phonological data from PEAS and further LAMSAS field records show that Erie was located inside of most of the Northern isoglosses and outside of most of the Midland /

Western PA isoglosses.

Chapter 3

Data Collection

3.1 Introduction

In this chapter I will describe the procedures I employed to collect the data for this dissertation. First, Section 3.2 will discuss the methodology in recruiting speakers to participate in interviews. Due to this project's non-traditional method of vowel analysis (involving transcription of the interviews and forced alignment) as well as the demographic characteristics of the desired speakers, the methods I employed for contacting speakers were different than the methods used for many other sociolinguistic studies. Next, Sections 3.3 and 3.4 will describe the types of speakers I targeted for inclusion in the corpus, and provide tables with demographic information for every speaker whose data was analyzed. Then, Section 3.5 will briefly summarize the demographic characteristics of the corpus as a whole. Section 3.6 will describe the interview procedure and the formal methods used for targeting specific variables. Finally, Sections 3.7 and 3.8 will provide a high-level description of the methodologies for data analysis involving transcription, forced alignment, and automatic vowel analysis (for a more technical description of this methodology, see Chapter 4).

3.2 Selection of individuals

While great care is often taken in the selection of individual speakers for inclusion in a sociolinguistic corpus to ensure a random, stratified sample including the social categories of interest (e.g., (Sankoff and Sankoff 1973)), this is not usually possible to do in a dialect geography study. Due to the increased time and effort involved in sampling speakers from a larger geographical area, only a handful of speakers are usually selected in each community. Thus, it is not possible to adequately control for speaker characteristics such as age, sex, and socioeconomic status. Rather, speakers in such studies are usually selected by a somewhat ad hoc process, as described in DARE (Cassidy and Hall 1985:xiv): “the intention was to maximize the collection of materials by going to the places and people most likely to furnish the largest amount of appropriate data.” This general approach was also followed in the sampling of speakers for this dissertation. In Erie, the focus was on older speakers, in an attempt to document the earliest stages of the city’s shift away from the North. In other areas, the closest attention was paid to boundary regions, especially the town of Ripley, NY, in order to obtain a detailed apparent time view of how the boundaries have been changing recently.

One relatively easy way of conducting targeted sampling based on geographic location is to select individuals from phone book listings. This is the is the method of participant recruitment that was employed by the ANAE (Labov et al. 2006:24–27). The main benefit of this procedure is that it enables the researcher to obtain an extremely wide geographical coverage without needing to travel to conduct the research. However, there is one main drawback to this procedure: since the calls are unsolicited and unexpected, the researcher is not able to prepare the speaker in advance with any specific materials. Thus, it is impossible to use formal methods that require the speaker to have access to a physical object, such as word lists, reading passages, or picture naming materials. The ANAE overcame this difficulty with a combination of verbal tasks, such as targeted elicitation of lexical items, naming of lexical items from a restricted set (e.g., numbers, days of the week, articles of clothing), semantic differentials, etc.—see Labov et al. (2006:32–34) for a complete list of all such formal elicitation tasks employed by the ANAE researchers. However, the best way to ensure that there is a large set of words that are uttered by all speakers in a corpus is to use a word list. The ANAE researchers realized this, and attempted to arrange follow-up interviews with speakers after sending them a word list in the mail (Labov et al. 2006:29); however, this second interview was only carried out with a minority of the ANAE speakers. Thus, in order to ensure that all speakers would participate in the word list and reading passage tasks, and to greatly speed up the minimal pairs task, I decided to not conduct any unsolicited telephone interviews for this dissertation. I did conduct several interviews over the telephone, as described in Section 3.6.2; however, in all cases the speakers had been contacted prior to the interview, and a specific appointment had been arranged (this was necessary to ensure that the speakers would have internet access at the time of the interview).

A second method of efficiently obtaining data for a dialect geography study is the Short Sociolinguistic Encounter (SSE), described in Ash (2002). This method involves anonymous face-to-face interviews of people in public places in the town of interest, thus enabling the use of word lists, reading passages, or any other printed formal method, during the initial contact. An SSE is usually shorter in duration than a full interview, thus enabling the researcher to collect data from several individuals during a single day of field work in a given town. To find subjects, the researcher approaches individuals in public places such as parks, cafes, stores, etc., and conducts the SSE in the same location. The technique was originally envisioned "to provide the maximum amount of data on a specified, small set of variables" (Ash 2002:2). For example, it was used successfully by Ash (2002) to obtain large amounts of data about the distribution of /æ/ in the Mid-Atlantic region, and by Johnson (2007) to investigate the distribution of /ah/, /o/, and /oh/ along the border between Massachusetts and Rhode Island. Recent work by Dinkin (2009) has maintained the basic approach of the SSE, but extended it somewhat in order to obtain enough data to describe each speaker's entire vowel system.

While the SSE is an efficient method that enables the dialect geographer to obtain a decent sample of the speech in a given location in as little as a single day, it does have some drawbacks. First of all, since the interviews are usually conducted anonymously, the speakers can not be contacted again in the future if follow-up research is necessary. More importantly, the anonymity of the encounter makes it much more difficult to ask the informant for referrals to other potential participants. Such personal introductions are often crucial in enabling the researcher to come in contact with speakers of a targeted demographic group that might not be easy to encounter in an SSE. This point is important for this dissertation, since the employment of the SSE as the sole method of data collection would have made finding a large number of elderly speakers very difficult.

However, the most severe problem of the SSE is that the recordings are often of lower quality. Due to the nature of how speakers are met in an SSE, recordings often take place outside (e.g., in public parks), or in public establishments such as cafes or retail stores. In both types of environments, there is almost always a substantial level of background noise that is also picked up during the recording. For the purposes of a standard soiciophonetic analysis of vowel quality, this would not be a problem, since the manual extraction of vowel formants is not hindered greatly by the presence of such background noise. However, for the purposes of this dissertation, the technique of forced alignment was employed to automatically segment the audio signal into phonemes (see Section 4.2 for a description of this methodology) which, in turn, enabled the automatic extraction of vowel formants. For the forced alignment procedure to perform optimally, the speech signal should be as clean as possible; background noise can be mis-recognized as speech from the speaker, thus causing errors in phoneme segmentation. For this reason, I attempted to conduct interviews in locations with a minimal amount of background noise, ideally in the speaker's residence, or a quiet public place, such as a meeting room in a library. Such interviews are not normally possible when using the SSE methodology.

In order to overcome the drawbacks of random telephone interviews and the SSE, a somewhat more laborious method was chosen as the primary way of selecting speakers. Advertisements were put up in libraries and community centers in all of the towns of interest, and potential informants were asked to contact me by telephone or email if they were interested in participating. Upon initial contact, a bried demogrphic interview was conducted to make sure that the speaker was born in the town of interest and lived there continuously until the age of 18. An effort was made to screen for speakers who also lived their entire adult lives in the town of interest, but speakers were not excluded if they lived outside of the town for short periods in their adult lives (for example, during college or for a job transfer). If the speaker met the residency criteria, a time would be arranged for the interview, either at the speaker's residence or at a maximally quiet public place. Response to the advertisements was surprisingly heavy, especially in the smaller towns, and many participants were initially contacted in this manner.

After conducting each interview, the participant was then asked to recommend any friends or family members who met the residency criteria and would be willing to participate. Almost all informants were willing to provide contact information for at least one other participant; when these references were pursued, they also usually led to more successful interviews. Thus, the bulk of the speakers were initially found through recommendations from previously interviewed speakers; usually only a single response to the initial advertisement in a given town was required to obtain an endless string of potential speakers. The only interviews that were not conducted in this manner were the ones at the Sun Valley retirement community (described below in Section 3.3.1), a few initial interviews in the city of Erie that were obtained through personal contacts, and a few early SSE's in neighboring towns that were conducted before the negative effects of the poorer quality recordings were fully realized. Despite the fact that the chosen methodology required a larger investment of time in speaker recruitment, and meant that more time had to be spent in the field, I believe the resulting increased quality of the recordings and deeper personal connection with the participants made the effort worthwhile.

Finally, in the later stages of the field work, several interviews were also conducted over the telephone using a call collection interface provided by the Linguistic Data Consortium. Speakers were initially contacted in the same manner as described above, but were interviewed over the telephone while I was in Philadelphia. This was done simply to reduce the number of trips that I would have to take to the area. The process of conducting telephone interviews through the LDC's interface led to very high quality recordings for the purpose of vowel analysis, despite the restricted frequency range transmitted over the telephone. Background noise in these interviews is almost non-existent, and the LDC's interface separates the input from the two phone lines into two separate audio files; both of these factors led to easier transcription and improved performance of the forced alignment system.

3.3 Selection of speakers: time depth

The first goal of selecting speakers for analysis was to push the time depth for our knowledge of Erie as far back as possible. As shown in Chapter 6, the data from the two LAMSAS speakers suggests that the merger of /o/ and /oh/ occurred in Erie sometime after 1910. In order to test this hypothesis, an effort was made to record elderly Erieites with the hope of finding some who were born before the merger took place. If present-day speakers with the /o/ $\sim$ /oh/ distinction could be found, it would enable us to pinpoint the date of the merger with a high degree of certainty, and, thus, convincingly demonstrate when Erie ceased to be a Northern city. In addition, data was obtained from several archival sources in an attempt to collect data from an earlier time period than is possible with living speakers. This real time evidence complements the apparent time evidence obtained from my own fieldwork and provides a more complete picture of the course of linguistic change in Erie and the neighboring region.

3.3.1 Sun Valley residents

As discussed above, my first goal in data collection was to find several elderly Erieites who were born before or around the date suggested by the LAMSAS evidence for the merger of /o/ and /oh/ in Erie. The set of data providing this apparent time evidence comes from one-on-one interviews I conducted in person at an upscale retirement community in Erie, which I will call Sun Valley. I contacted the facility's commuity director, and she put an announcement in the weekly newsletter advertising my survey. Volunteers who were interested in participating in the interviews contacted her, and she set up appointments throughout the course of October, 2007. In total, I conducted 12 interviews at Sun Valley with life-long residents of Erie and three other cities of interest; these speakers ranged in age from 66 to 95. The oldest Sun Valley resident was born in 1912, and would thus be a good candidate for maintaining the distinction, assuming the chronology in Section 2.3 is correct. Table 3.1 displays the demographic characteristics of these 12 speakers.

3.3.2 Archival material

In addition to interviewing the older Erie residents from Sun Valley, I made an attempt to obtain real time data from speakers that would extend the time depth even further into the past. For this purpose, archival material was obtained from three distinct sources described below.

Name	Born	City	State	Occupation
Dan R.	1912	Erie	PA	engineer
Robert E.	1916	Erie	PA	doctor
Mary D.	1919	Erie	PA	
Flora R.	1919	Erie	PA	teacher
Eloise B.	1925	Erie	PA	
Charles B.	1925	Erie	PA	manager
Dottie A.	1926	Erie	PA	teacher
Sally W.	1928	Erie	PA	deputy sheriff
Dana W.	1941	Erie	PA	teacher
Jane S.	1915	Oil City	PA	
Marge K.	1919	Pittsburgh	PA	
Walter K.	1927	Buffalo	NY	engineer

Table 3.1: Demographic characteristics of the 12 Sun Valley residents

The first source was the Seasonal Workers in Viticulture (SWV) corpus. This corpus was compiled as part of an oral history project conducted in 1988 to document the local grape growing industry around North East, PA. The town of North East is located in the northeastern corner of Erie County, about 20 miles from downtown Erie, and directly across the state line from New York. The town has always been a center for grape production, with a focus on producing juice from the Concord variety. A few other grape varieties are also grown, and several wineries exist along both sides of Route 20. Viticulture has always been one main source of jobs for North East residents, in addition to other types of agriculture.

The SWV project attempted to interview older native residents of North East who had owned vineyards or who had worked as grape pickers, although a few younger people and a few in-migrants were also interviewed. In all, 50 recordings were made, each about one hour in length. The interviews are available to the public as cassette tapes at the Erie County Historical Society. I selected the two oldest, native North East residents from the corpus for analysis, since they were most likely to have maintained a distinction between /o/ and

ID	Born	Interviewed	City	State	Source
SWV 039	1906	1988	North East	PA	SWV
SWV 046	1907	1988	North East	PA	SWV

Table 3.2: Demographic characteristis of the two speakers from the SWV corpus whose speech was analyzed manually

/oh/. These two speakers were born in 1906 and 1907. Their demographic information is summarized in Table 3.2.

The recordings for these two speakers are quite poor: the microphone was positioned far away from the speaker's mouth, and each recording contains a large amount of background noise. They are thus poor candidates for the procedure involving forced alignment and automatic vowel analysis that will be used for the rest of my data. Therefore, the vowels for these two speakers were analyzed manually, and they will not be included in the Natural Break maps in Chapter 5. However, the manual measurements for these two speakers will be included in the analysis of the merger of /o/ and /oh/ in Erie in Chapter 6, since they provide a useful early source of data.

The second source of archival material is a set of recordings from the Dictionary of American Regional English (Cassidy and Hall 1985–2002). I analyzed the interview speech from three DARE speakers from Erie County (two from North East and one from Union City). In addition, I analyzed the "Arthur the Rat" reading passage recordings from these three speakers and an additional 11 DARE speakers from the boundary regions around Erie.[1] The demographic information for these 14 speakers is provided in Table 3.3. Personal names of the speakers in the DARE corpus are not public information; the names provided in Table 3.3 are pseudonyms that I created. The DARE ID numbers are also

---

[1] The digitized DARE audio files were provided by Joan Hall, chief editor of DARE, at the University of Wisconsin.

Name	ID	Born	City	State	Occupation
Nancy S.	PA129	1908	North East	PA	teacher
Sarah N.	PA130	1897	North East	PA	homemaker
Gladys T.	PA131	1899	Meadville	PA	policeman
Bill C.	PA133	1950	Meadville	PA	librarian
Agatha S.	PA181	1907	Warren	PA	teacher
Steven G.	PA182	1915	Warren	PA	medicine
Maggie S.	PA234	1900	Union City	PA	teacher
Anne B.	NY099	1898	Fredonia	NY	homemaker
Leslie B.	NY100	1897	Fredonia	NY	teacher
Wallace L.	NY101	1892	Fredonia	NY	factory worker
Jonas H.	NY102	1898	Ripley	NY	vintner
Clarence T.	NY103	1886	Ripley	NY	engineer
Jill C.	NY104	1889	Ripley	NY	seamstress
Ted L.	NY215	1904	Jamestown	NY	craftsman

Table 3.3: Demographic characteristics of 14 DARE speakers from archival sources whose acoustic data were analyzed, interviewed 1968 - 1969

Name	Born	City	State	occupation
H.O. Hirt	1887	Erie	PA	CEO

Table 3.4: Demographic characteristics of H.O. Hirt, interviewed in 1977 for the Erie Insurance Company archives

provided, to enable reference to the specific speakers in published DARE material.

Additionally, I obtained a VHS tape of an interview with H.O. Hirt, the founder of Erie Insurance Exchange.[2] Hirt was born in Erie in 1887, founded the company in 1925, and served as its CEO until 1976. The interview was conducted with him in 1977 and a 20-minute segment of it was released as a publicity tape by the Erie Insurance Group. This interview thus represents the oldest recorded Erieite that I have so far been able to discover. H.O. Hirt's demographic information is summarized in Table 3.4.

Finally, seven speakers from the ANAE corpus were reanalyzed for this dissertation

---

[2] The tape was given to me by the staff archivist at the Erie Insurance Group.

Name	ID	Born	City	State	Occupation
Irvin H.	TS168	1932	Cleveland	OH	special education
Samuela S.	TS364	1964	Erie	PA	car wash attendant
Ken K.	TS545	1961	Pittsburgh	PA	student
Gwen S.	TS355	1929	Pittsburgh	PA	unknown
Cecilia S.	TS356	1933	Pittsburgh	PA	student
Henry K.	TS544	1935	Pittsburgh	PA	teacher
Charlotte S.	TS739	1961	Pittsburgh	PA	secretary

Table 3.5: Demographic characteristics of the 7 ANAE speakers whose acoustic data were re-analyzed, interviewed 1994 - 1996

using the methodology of transcription and forced alignment. Most of the ANAE speakers chosen for reanalysis (five out of the seven) were from Pittsburgh. They were selected in order to provide a more complete description of this city for comparison with Erie, since my own fieldwork recordings only contained three speakers from Pittsburgh. Because the method of analysis for these seven ANAE speakers involved transcription and forced alignment, their vowel analyses are based on their entire interviews, not only the words that were measured manually by the ANAE annotators. Their demographic information is provided in Table 3.5. Again, personal names of the speakers in the ANAE corpus are private; the pseudonyms in Table 3.5 are the ones provided by the public version of the ANAE database released with the corpus.

3.4 Selection of speakers: geographical depth

The second aim of selecting speakers for analysis was to collect data from the small towns around Erie, in an attempt to determine the nature of the dialect boundaries between Erie and the North, on the one hand, and Erie and Pittsburgh, on the other.

First of all, a more complete description of the city of Erie itself was needed, since the

Name	Born	Occupation
Barry G.	1938	car salesman
Laurie G.	1946	nurse
Pam R.	1945	homemaker
Sophie D.	1950	computer programmer
Jane L.	1953	gardener
Tom L.	1953	welder
Greg A.	1980	unemployed
Sally L.	1982	office worker

Table 3.6: Demographic characteristics of 8 speakers from the city of Erie

only study to date of Erie speech involving acoustic analysis is the ANAE, in which two speakers were analyzed. I recorded and analyzed the speech of nine native Erieites. Their demographic information is presented in Table 3.6.[3]

In order to investigate the boundary region between Erie and the North, I visited several towns in Chautauqua Co., NY, located in the western part of the stated along the border with Erie Co., PA. The speakers from these towns are listed in Table 3.7. As can be seen from Table 3.7, a strong focus was placed on the town of Ripley, NY. This town, the first one in NY after crossing the state line from PA, was discovered to have an apparent time distribution of /o/ and /oh/ indicating that these two phonemes began to merge about two generations ago. Thus, I sought a higher number of speakers there than in other towns in Chautauqua Co. in order to provide a more complete apparent time distribution of this change.

Due to the time spent on the field work in the boundary area between Erie and the North in Chautauqua Co., NY, I was only able to interview a few speakers from the other boundary area between Erie and the North, namely to the west of Erie in Ohio. These

---

[3] All names given for speakers that I interviewed are pseudonyms. They are provided to facilitate reference to the vowel plots for specific speakers that will be presented in Chapter 6 and 7.

Name	Born	Town	Occupation
Winifred S.	1925	Ashville	waitress
Mae S.	1925	Bemus Point	secretary
Bill R.	1930	Dunkirk	insurance salesman
Daisy T.	1921	Fredonia	secretary
Joan P.	1938	Jamestown	social worker
Barbara C.	1952	Jamestown	librarian
Amy G.	1963	Jamestown	teacher
Amy C.	1937	Westfield	librarian
Ralph O.	1934	Ripley	grape farmer
Margaret B.	1940	Westfield	historian
Stan R.	1948	Ripley	grape farmer
Rachel A.	1951	Ripley	daycare provider
Larry K.	1952	Ripley	town supervisor
John M.	1953	Ripley	town supervisor
Pam O.	1958	Ripley	winery owner
Daphne R.	1958	Ripley	grape farmer
Shelly I.	1960	Westfield	
Rachel C.	1963	Ripley	town clerk
Troy R.	1989	Ripley	student
Ryan N.	1994	Ripley	student
Grace N.	1997	Ripley	student
Jeff H.	1952	Buffalo	teacher

Table 3.7: Demographic characteristics of 23 speakers from New York (all speakers are from Chautauqua County except Jeff H.)

Name	Born	Town	County	Occupation
Patti N.	1957	Conneaut	Ashtabula	librarian
Brenda W.	1937	Ashtabula	Ashtabula	
Lisa C.	1940	Cleveland	Cuyahoga	librarian

Table 3.8: Demographic characteristics of 3 speakers from northeastern Ohio

Name	Born	Town	County	Occupation
Bob O.	1947	Ford City	Armstrong	teacher
Mary N.	1948	Butler	Butler	librarian
Abe M.	1944	Franklin	Venango	teacher
Ed W.	1932	Franklin	Venango	engineer
Bart P.	1946	Franklin	Venango	goverment administrator
Gary S.	1930	Franklin	Venango	goverment administrator
Carol H.	1942	Warren	Warren	health care administrator
Charlene O.	1934	Pymatuning	Crawford	teacher
Allison N.	1932	Pittsburgh	Allegheny	
Sara B.	1958	Pittsburgh	Allegheny	health care administrator
Kevin W.	1974	Greensburg	Westmoreland	banking

Table 3.9: Demographic characteristics of 11 speakers from western PA

speakers are listed in Table 3.8.

Additionally, research was conducted in the region of western PA between Erie and Pittsburgh in an attempt to ascertain the extent of the influence of the Pittsburgh system in the region. These speakers are listed in Table 3.9.

Finally, speakers were sought from most of the small towns in Erie Co., PA, especially those in the eastern and southern portions of the county, in an attempt to provide a more complete picture of the boundary areas between Erie and the North and Pittsburgh, respectively. These speakers are listed in Table 3.10.

Furthermore, I conducted abbreviated interviews with an additional 18 speakers who could not be recorded due to time constraints. Most of these speakers completed the por-

Name	Born	Town	Occupation
Jane W.	1948	Edinboro	principal
James N.	2000	Edinboro	student
Irene C.	1927	Wattsburg	telephone operator
Cathy A.	1955	Lawrence Park	librarian
Catherine F.	1942	North East	banker
Betty W.	1936	North East	teacher
Sharon N.	1931	Union City	
Cindy M.	1943	Girard	teacher
Charlotte S.	1955	Waterford	teacher
Marjorie S.	1986	Waterford	student

Table 3.10: Demographic characteristics of 10 speakers from Erie Co.

tions of the survey involving minimal pairs and acceptability judgments (see Section 3.6.1 for more information about the components of the interview), although a few only completed one or the other. Since there is no acoustic data available for these speakers, they will only be included in the maps for minimal pairs and lexical and morphosyntactic variables, when appropriate. Table 3.11 provides demographic information for these 18 speakers.

3.5 Characteristics of the corpus

One of the main research goals of this dissertation is to explore the oldest stages of the Erie system, in order to discover when and how Erie ceased to be part of the North. Many of the speakers thus fit the profile of the NORM (Non-mobile, Older, Rural, Male) speaker that is traditionally the target speaker for a dialect geography study (Chambers and Trudgill 1999:29) (although many elderly female speakers were also interviewed). Younger speakers were only targeted specifically when a change in progress was detected in the community, and a more complete apparent time distribution was desired. Figure 3.1 displays a histogram of the birth year for all 106 speakers in the corpus, both from interviews con-

Name	Born	Town	County	State	Occupation
Susan B.	1959	Westfield	Chautauqua	NY	unknown
Sheila T.	1950	Ripley	Chautauqua	NY	waitress
Jane L.	1960	Ripley	Chautauqua	NY	waitress
Tracy N.	1972	Ripley	Chautauqua	NY	waitress
Heather E.	1990	Ripley	Chautauqua	NY	high school student
Trevor J.	1990	Ripley	Chautauqua	NY	high school student
Teri F.	1990	Ripley	Chautauqua	NY	high school student
Carrie B.	1990	Ripley	Chautauqua	NY	high school student
Vanessa T.	1990	Ripley	Chautauqua	NY	high school student
Chloe S.	1990	Ripley	Chautauqua	NY	high school student
Adam R.	1990	Ripley	Chautauqua	NY	high school student
Rebecca R.	1980	Ripley	Chautauqua	NY	baker
Charles S.	1956	Buffalo	Erie	NY	public services
Edith N.	1933	Ashtabula	Ashtabula	OH	clerk
Dan A.	1923	Erie	Erie	PA	retail manager
Tess E.	1945	Erie	Erie	PA	telephone operator
Laura S.	1953	Erie	Erie	PA	unknown
Sadie N.	1960	Girard	Erie	PA	clerk

Table 3.11: Demographic characteristics of 18 unrecorded speakers (they will only be displayed in the maps for minimal pairs and lexical / morphosyntactic variables)

ducted specifically for this dissertation (including speakers with no audio data) and from archival material.

The corpus contains 68 females and 38 males; thus the ratio of female to male speakers is 1.8:1. This ratio is nearly identical to the overall female to male ration of 1.7:1 in the ANAE corpus (Labov et al. 2006:28).

The map in Figure 3.2 shows the geographic location of all 88 speakers whose acoustic data will be analyzed in this dissertation, including both speakers from archival sources and speakers from interviews conducted for the dissertation.

3.6 Interview procedure

3.6.1 Materials

For speakers who had enough time (ca. 30 minutes or more), the interview consisted of approximately 20 minutes of conversation and the formal methods. The conversation was targeted first to extract the necessary demographic information from the speaker, and then to ask them to describe their town (what the downtown is like, how it has changed) and talk about any other nearby towns or cities they frequently go to. Other topics that were often discussed at length include jobs, family, and hobbies. The conversations were thus similar in style to the interviews conducted for the ANAE, and are more properly characterized as dialectological interviews than sociolinguistic interviews.

The formal methods section of the interview consisted of a word list, a set of minimal pairs, and an acceptability judgment task. The word list consisted of 159 words, and was designed to provide a complete view of a speaker's vowel system. Extra tokens of words with vowels crucial to the present study (namely /o/, /oh/, /æ/, and /ow/) were included to ensure the reliability of mean values for these vowels (see Appendix C for the complete

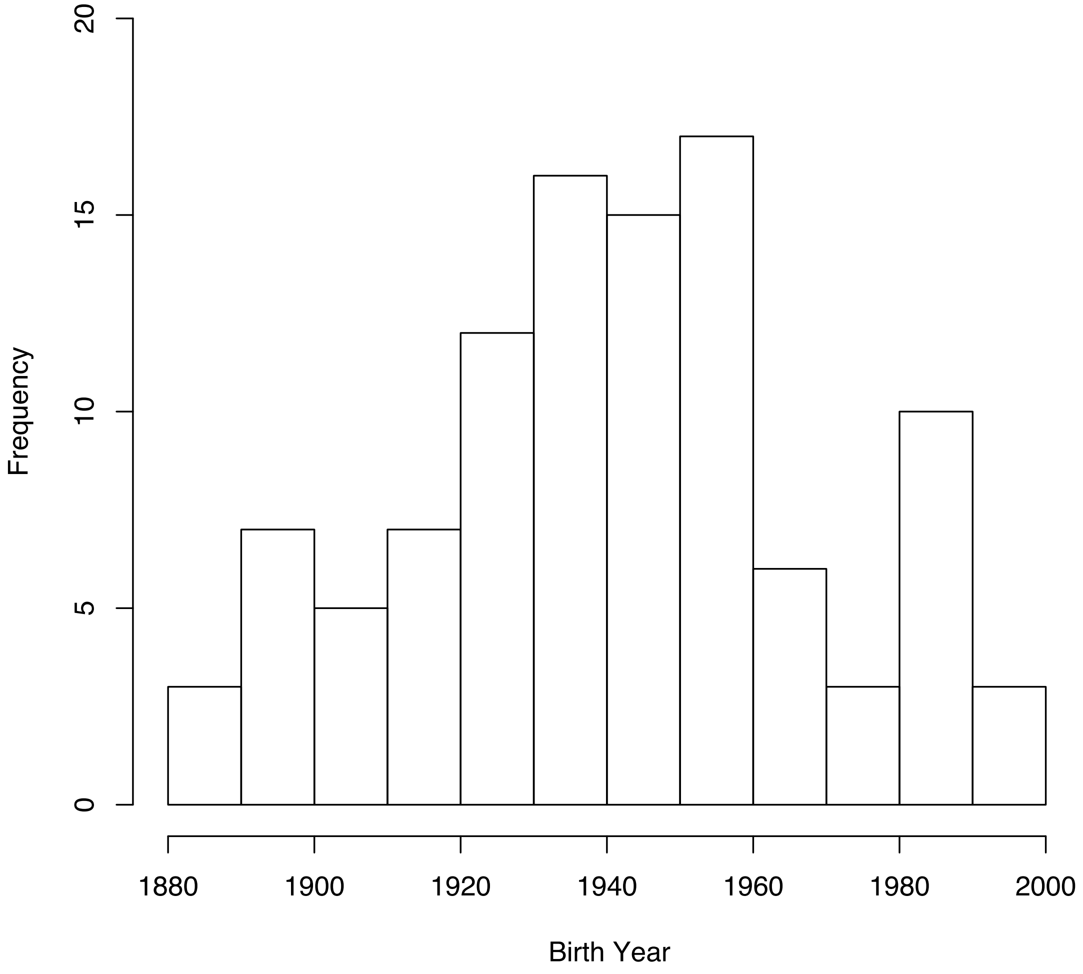

Figure 3.1: Histogram of birth year for all 106 speakers in the corpus (from both new interviews and archival sources)

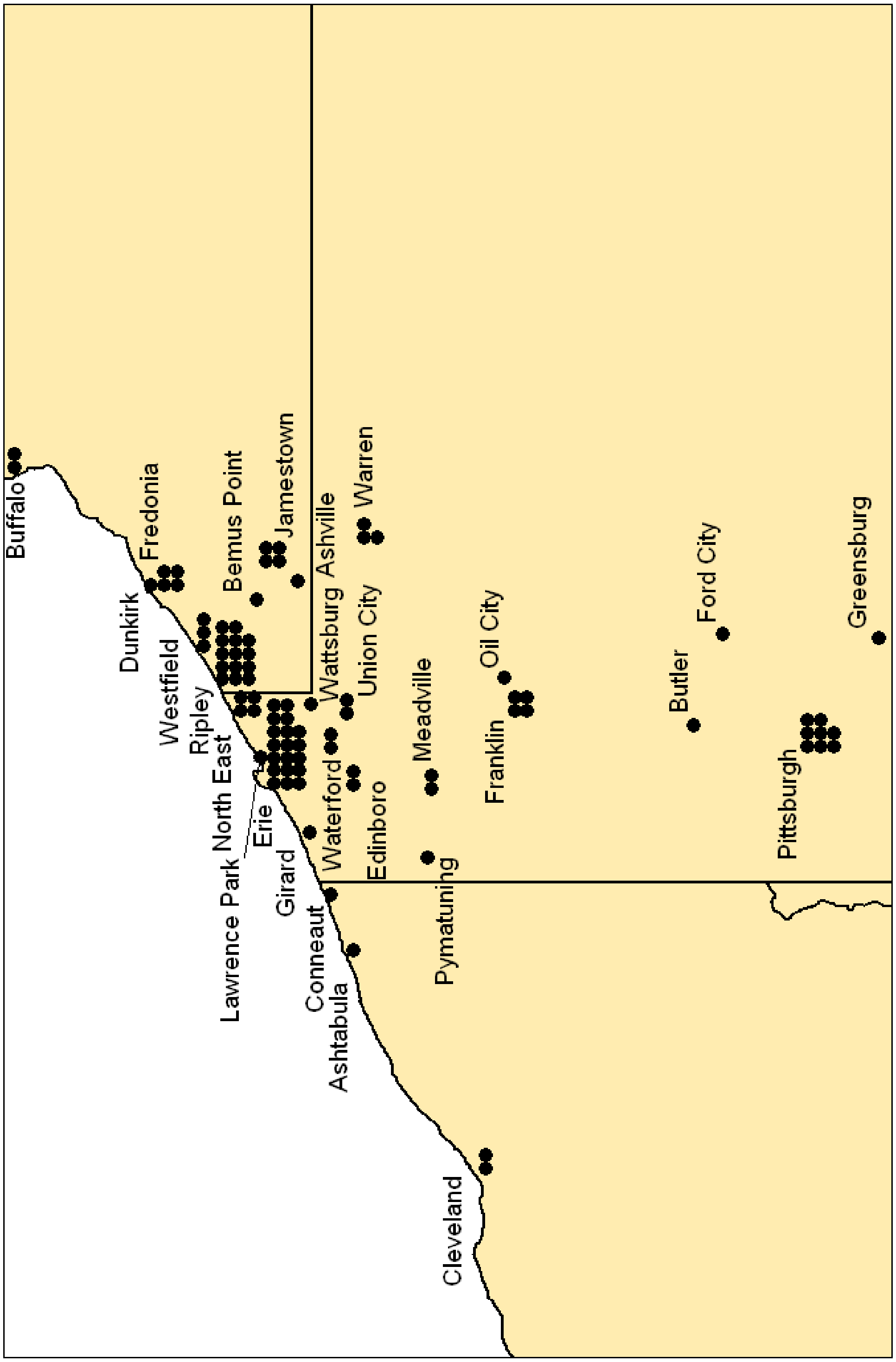

Figure 3.2: Locations of all 88 speakers whose acoustic data will be analyzed

word list). The minimal pair section contained 17 minimal pairs and 4 near-minimal pairs (see Appendix B for a complete list). Seven of the items involved the main distinction of interest, /o/ vs. /oh/, and two tested the contrast of /uw/ and /u/ before /l/. Two further near-minimal pairs examining the status of /ah/ were included, as well as one designed to determine the phonemic status of the vowel in *on*. The remaining pairs were included as filler material, and were not expected to vary within the geographic region under consideration. Finally, the acceptability judgment task consisted of 10 items designed to test the geographic extent of certain Midland lexical and morphosyntactic usages (see Appendix D for a complete list). For a few speakers who only had a limited amount of time, the interview consisted of at least the minimal pairs, and sometimes also the word list and sentence judgment task.

The audio data for the speakers from the DARE corpus consists of the "Arthur the Rat" reading passage and guided conversation. The methodology used for collecting this data is described in detail in the Introduction to Cassidy and Hall (1985). For this dissertation, interview data from three DARE speakers from Erie County will be analyzed, as well as "Arthur the Rat" data from 14 speakers (including the three from Erie County as well as 11 from the neighboring boundary regions.) The full text of the version of the "Arthur the Rat" reading passage used by DARE is included in Appendix E (Cassidy and Hall 1985:xliii).

The interviews conducted for the ANAE consist of a variety of elicitation techniques designed to encourage the speakers to say certain words, minimal pair tests, grammaticality acceptability ratings, and guided conversation (see Labov et al. (2006:29–35) for a complete description of the ANAE interview procedure). For this dissertation, the entire interviews with the speakers in Table 3.5 were transcribed and analyzed.

3.6.2 Equipment

The face-to-face interviews were conducted using an M-Audio Microtrack 24/96 solid state recorder with 16-bit quantization and a sampling rate of 44,100 Hz. A Sony ECM-717 lavalier microphone was attached to the speaker's shirt near the chin. The telephone interviews were conducted using a call collection interface provided by the Linguistic Data Consortium, and were separated into separate μ-law encoded audio files for each of the two channels. During the telephone interview, participants accessed a web site that contained the word list, minimal pairs, and sentence judgment task.

3.7 Transcription

The speech samples to be analyzed come from a wide variety of sources: interviews conducted specifically for this dissertation, interviews from archival sources (DARE, ANAE), word lists, and reading passages (DARE's "Arthur the Rat"). All of these types of speech materials were analyzed using the methodology of forced alignment and automatic vowel analysis described in Chapter 4. In this section, I will describe the procedure I followed to prepare the data for this type of analysis.

Before a sound recording can be processed by forced alignment, the speech must be transcribed orthographically. In order to complete the transcriptions quickly and efficiently, I developed a program that enabled me to transcribe the speech and non-speech sounds in each recording with a minimum of stopping and rewinding. The program, called `quickTrans`, is a collection of Python, Praat, and shell scripts, and can be downloaded freely as part of P2TK, the Penn Phonetics Toolkit.[4]

The general approach that `quickTrans` uses is to automatically segment the sound file into small chunks based on intensity levels of the signal. This is done by setting a mini-

---

[4] PT2K is available at *http://www.ling.upenn.edu/phonetics/p2tk/*.

mum intensity threshold level for a segment of speech to be considered part of an utterance, as well as a duration threshold for pauses. If a segment of speech has an intensity level continuously below the intensity threshold for a period of time longer than the minimum pause duration, then it is marked as a pause. The portions on either side of the pause with greater intensity are then marked as utterance chunks. These small chunks, which are intended to be short enough to fit into short-term memory, are then played sequentially by an audio player (Praat). Transcription can then proceed with a minimum amount of rewinding to repeat utterances.[5]

After some experimentation, the intensity threshold for pauses was set to 1/3 of a standard deviation less than the mean, and the minimum pause duration was set to 200 msec. In other words, if there is a segment of the audio file which has an intensity level consistently less than 1/3 of a standard deviation below the mean intensity level of the sound file for greater than 200 msec, then it is marked as a pause between utterances. These configuration values worked well for most of the recordings in my corpus. However, they needed to be modified for some recordings with a large amount of overlapping speech, constant background noise, or heavily unbalanced intensity levels between two interview partners (e.g., because the microphone was much closer to one interlocutor than the other). In such cases, the `quickTrans` configuration variables were modified until they produced utterance chunks that were mostly short enough in duration to fit into short-term memory and that did not omit segments of speech data.

This approach used by `quickTrans` is similar to the `AutoSegmenter` tool from the LDC (Glenn and Strassel 2008). It enables quick transcription, where the goal is simply to "get the words right" as quickly as possible. The LDC reports transcription rates of approximately seven to ten times real-time using `AutoSegmenter` while following

---

[5] If the transcribed is not able to transcribe an entire utterance on the first pass, the `quickTrans` interface does enable the entire utterance to be repeated. However, more complex manipulations, such a replaying part of an utterance or listening to previous utterances must be done in an external sound editor.

the quick transcription guidelines described in LDC (2004). For my own transcriptions of my dissertation corpus, I was able to transcribe most files at a rate of around five times real-time. This fast rate of transcription was partially due to the fact that I omitted all information from the transcription that was not necessary for the purposes of forced alignment and automatic vowel analysis. Specifically, the transcriptions include no punctuation and are all in lower case. On the other hand, certain transcription conventions were followed in order to make the forced alignment as accurate as possible (see Evanini et al. (2009b) for a more complete description of these transcription practices and how they affect the output of forced alignment). These included:

- **Disfluencies**: Partial word disfluencies, such as the false start of *sch-* for *school* are transcribed with a hyphen representing the part of the word that was not spoken. Pause fillers, such as *um* and *uh*, are always transcribed.

- **Non-speech sounds**: The forced alignment system recognizes five symbols for non-speech sounds: {BR} for breath, {CG} for cough, {LG} for laughter, {LS} for lip smack, and {NS} for background noise. These were transcribed as such when they were loud enough to be included in the utterance chunks produced by `quickTrans`.

- **Unknown words**: Words that are uttered in the audio file but that do not have a corresponding transcription in the forced alignment system's pronouncing dictionary are problematic. If no pronounciation is available, then the system can not include them in the alignment; this also leads to sub-optimal alignments for the words in the transcription surrounding the missing word. Therefore, after a post-transcription check with the pronouncing dictionary, all words in the transcription that were missing from the dictionary were manually provided with pronunciations.

- **Multiple speakers**: For audio files with multiple speakers, it is necessary to annotate which speaker uttered each word. Without this, automatic vowel analysis would

> be impossible, since the identity of the speaker for each token would be unknown. The convention I used in my transcriptions was to prepend a single capital letter corresponding to the speaker that produced each utterance for every audio file with speech data from more than one speaker.[6]

For example, a portion of the transcription for the interview with Dottie A. is reproduced below. It illustrates the transcription of a partial word disfluency (*sch-* for *school*), filled pauses and backchannels (*uh* and *uh huh*), non-speech sounds ({LG} for laughter, {BR} for breath, and {NS} for background noise) and input from multiple speakers (Speaker `A` is Dottie A. and Speaker `B` is the interviewer (me)).

```
A i taught in panama new york
B uh huh A and uh
A then i came back to erie
A and when i got a teaching job again
A i taught just middle sch- middle school and junior high
B and how did you enjoy that job
A well i liked it i liked the subject of home economics
A but i
A wasn't too crazy about the discipline i had to {LG}
A {BR}
A to uh manifest in order to
A keep the children
{NS}
A so that they would
A be getting some kind of an education
```

3.8 Vowel measurements

Vowel formants were extracted automatically for all speakers in the corpus (except for the two speakers from the SWV corpus, as described above) according to the procedures

---

[6] Because all interviews were transcribed in lower case, this convention never produced any ambiguities with words from the transcription.

described in Chapter 4. A collection of Python scripts were written to implement the automatic vowel analysis techniques developed for this dissertation. They are included in the program `extractFormants`, also available as part of the P2TK package.

The `extractFormants` configuration variables were set so that only vowels longer than 50 msec in duration (as determined by the output of the forced alignment procedure) were measured. In total, 113,245 F1 and F2 measurements were extracted (vowel tokens from the interviewer and other speakers in the recording that will not be analyzed are excluded from this figure). This number includes vowels with all three levels of lexical stress: primary, secondary, and none, as well as vowels in all segmental environments.

After the vowel formants were extracted, they were normalized on a per-speaker basis to reduce the effects of individual variation in vocal tract length, according to the log-mean algorithm in Nearey (1977) (see also Labov et al. (2006:39–40) for a concise summary of the procedure). The group log mean value from the ANAE, 6.896874, was used for this corpus as well, since it was calculated from a larger number of speakers.

In an attempt to remove outliers due to errors introduced by the automatic formant measurement process, the automatic measurements were compared to the database of manual measurements taken for the ANAE. Any normalized automatic measurement that was outside the range of all of the normalized hand measurements for each vowel class was excluded from the set of automatic measurements. This approach is justified, since it is quite unlikely that such measurements represent accurate formant values for the vowel. Since the ANAE database is both larger than the database for this dissertation and drawn from a wider variety of dialects, it would be expected to exhibit a wider range of formant values for each vowel class. Thus, any measurement outside of these ranges can be treated as erroneous. A total of 1,282 automatic measurements were excluded based on this metric, approximately 1% of the total database. Afterwards, 111,963 vowel measurements remained for analysis.

As an example of the usefulness of excluding automatic measurements outside the bounds of the manual ANAE measurements, consider the vowel plot of /o/ and /oh/ for Dan R. from Erie shown in Figure 3.3. A single measurement from the /o/ class is a clear statistical outlier. This token of *got*, shown in the lower-left corner of the vowel plot, was provided with an F1 measurement of 1247 Hz and an F2 measurement of 2115 Hz by the automatic vowel analysis procedure described in Chapter 4.[7] However, this measurement is clearly an error, since the articulation of a vowel with such a high F1 value would be physically impossible for this speaker. Additionally, manual inspection of this token shows that the correct (unnormalized) F1 and F2 measurements should be around 501 Hz and 1206 Hz, respectively. Clearly, the automatic formant prediction procedure erred in this case by substituting F2 for F1 and F3 for F2.

The exclusion procedure described above applies to this token of /o/ in *got* for both the F1 and the F2 values. The normalized F1 value for this measurement is 1520 Hz, and the maximum normalized F1 value for all manual measurements of /o/ in the ANAE database is 1264. Additionally, the normalized F2 value for this automatic measurement is 2578, and the maximimum normalized F2 value for /o/ in the ANAE is 1915. Thus, both the F1 and F2 measurements for this token of *got* are outside the range of manual measurements for /o/, and this token is correctly labeled as an error by this procedure and excluded from consideration.

The exclusion of such outliers caused by measurement errors is useful, because such gross errors can have a disproportionately large effect on the mean values. For example, Dan R.'s F1 and F2 mean values for 56 tokens of /o/ before the exclusion of outliers are 717 Hz and 1361 Hz, respectively. After this single token of *got* with the gross measurement error is excluded, the mean values change to 704 Hz and 1338 Hz. This change in the mean values is not so drastic for Dan R., because his number of /o/ tokens is quite high. However,

---

[7] These values represent the original unnormalized measurements produced by the system, whereas the plot in Figure 3.3 shows the measurements after normalization. That is why the stated values do not correspond exactly to the coordinates in the figure.

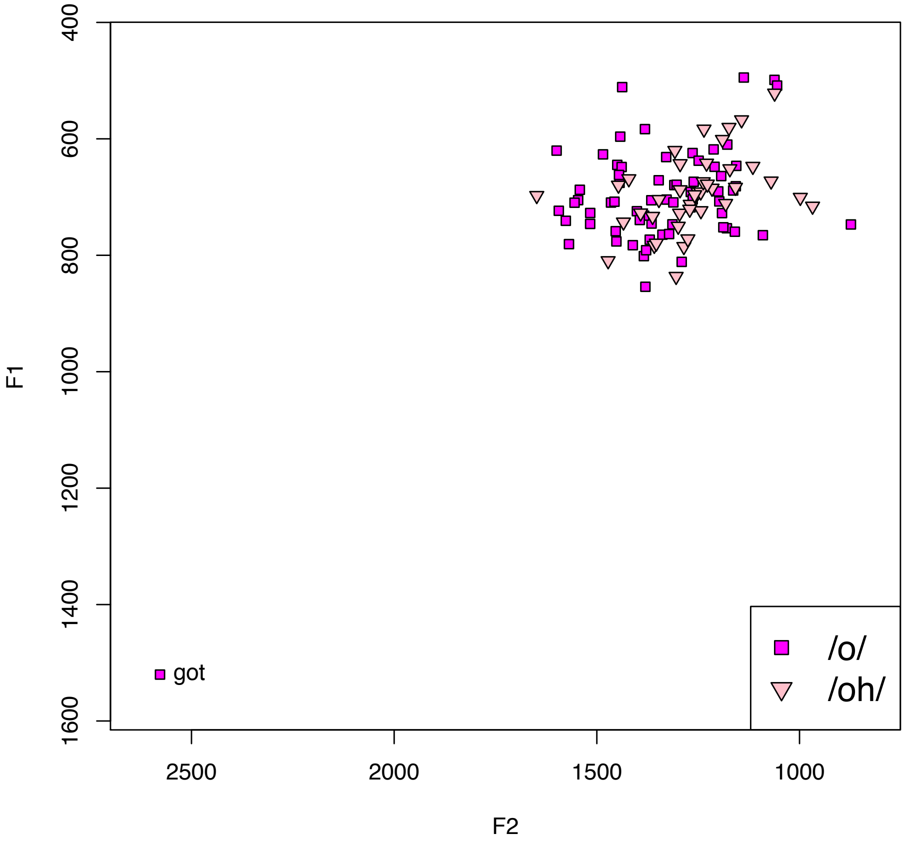

Figure 3.3: Automatic formant measurements for the vowels /o/ and /oh/ for Dan R. from Erie. The token of *got* in the lower-left corner is a gross measurement error that should be excluded from a vowel analysis.

a single gross measurement error such as this one can cause the mean values to shift by a few hundred Hz when the number of tokens is small. This point becomes important for the DARE speakers whose /o/ and /oh/ values are analyzed from the "Arthur the Rat" reading passage. As Section 6.7.3 will describe, each speaker produces about 14 tokens of /o/. For these speakers, the exclusion of gross errors is necessary in order to obtain reliable mean values.

After gross errors were excluded by this procedure, a subset of the remaining 112,087 vowel measurements were then selected when calculating vowel means for individual speakers in Chapters 5 through 7. The following list provides details about which tokens were excluded:

- vowels before /l/ and /r/
- vowels after /w/ and /y/
- vowels after obstruent+liquid onset clusters
- tokens of /i/, /e/, /æ/, and /aw/ before nasals
- vowels with secondary stress and unstressed vowels (as indicated by the phonemic transcription in the CMU pronouncing dictionary)
- vowels in a set of high-frequency function words that often undergo reduction: *and, but, for, he, he's, huh, I, I'll, I'm, is, it, it's, its, my, of, oh, she, she's, that, the, them, then there, they, this, uh, um, up, was, we, were, what, you*

The first four sets of exclusions based on the neighboring segmental environments correspond to the exclusions that the ANAE authors made before means were calculated (Labov et al. 2006:77). The exclusion of vowels not marked with primary lexical stress and vowels from words in the list of stop words represents an attempt to reproduce the ANAE's selection criteria. Under their approach to vowel analysis, the annotators only measured vowels bearing primary lexical stress. Additionally, most words chosen for analysis also bore primary phrasal stress (Labov et al. 2006:37). Focusing only on these tokens for vowel mean measurements reduces the centralizing effect of vowel reduction and presents a more accurate view of the phonetic targets for each vowel.

After these exclusions were applied, the total number of remaining vowel measurements was 44,599. Unless stated otherwise, all maps and figures displayed below were generated from this smaller subset of 44,599 vowels.

Chapter 4

Automatic Vowel Analysis

4.1 Introduction

In this section I will describe the method that will be used for automatically analyzing vowel tokens in my corpus. Three separate procedures are combined in the process of automatically extracting F1 and F2 measurements for a given vowel:

- **Forced alignment**: This process takes as input the audio file and a word level transcription and outputs time stamps for all words contained in the transcription as well as all phonemes contained in each word. Forced alignment thus enables the researcher to automatically locate all vowels of interest (e.g., all stressed vowels or all tokens of vowels involved in the Northern Cities Shift) in the speech signal.
- **Selection of measurement point**: In order to easily compare individual vowel tokens as well as mean values from a given vowel class (i.e., across speakers, dialect regions, etc.), it is necessary to abstract away from the dynamic trajectory of each vowel token and drastically reduce the dimensionality of its representation. For this research, the standard sociolinguistic practice of selecting a single point within the duration of a given vowel to represent its "central tendency" (Labov et al. 1972), (Labov et al.

2006:38) will be adopted, and methods for automatically selecting this point will be compared.

- **Formant prediction**: After an LPC analysis is performed on the speech signal, it is necessary to determine the location of the formants based on the poles and associated bandwidths produced by the LPC analysis. In most sociolinguistic research, the default automatic formant tracking procedures contained in standard software packages, such as Praat, are used initially. However, manual inspection of the formant tracks for each vowel token is considered necessary in order to correct potential errors in formant prediction. If an error is detected, the researcher changes the parameters of the formant tracking software until usable values are produced. In an attempt to eliminate this step of manual intervention, I develop a procedure for predicting vowel formants from an LPC analysis that simulates this human error correction process and reduces much of the error that is normally associated with automatic formant prediction.

Each of these three procedures will be described in turn in the sections below.

4.2 Forced alignment

Before any procedure for vowel analysis, either manual or automatic, it is first necessary to determine the locations of the vowels to be analyzed. In the standard sociolinguistic methodology using manual formant analysis, the vowel locations are determined by listening to the audio file and pausing it each time a relevant token is heard. For automatic vowel analysis, on the other hand, it is necessary to first provide a phoneme-level segmentation of the audio file. This can be done automatically by forced alignment when an orthographic transcription of the audio file is available.

Forced alignment of an audio file with a word-level orthographic transcription proceeds as follows: first, a pronouncing dictionary is used to obtain a phoneme-level transcription. Then, each of the phonemes in this sequence is replaced by a statistical model that was trained on multiple realizations of the phoneme. Next, the speech signal is converted into a low-dimensional representation on a frame-by-frame basis. Finally, dynamic programming is used to align each frame from the speech signal with the acoustic model in the phonemic transcription that was most likely to have produced it. The most widely used type of statistical model for phoneme alignment is the Hidden Markov Model (Rabiner 1989), the two most commonly used representations of the speech signal are MFCCs (Davis and Mermelstein 1980) and PLPs (Hermansky 1990), and the optimal phoneme alignment path is normally computed using the Viterbi algorithm (Viterbi 1967).

For this research, the data will be processed using the Penn Phonetics Lab Forced Aligner, henceforth P2FA, developed by Jiahong Yuan at the University of Pennsylvania (Yuan and Liberman 2008a)[1]. P2FA uses GMM-based monophone HMM acoustic models with 32 mixture components on 39 PLP coefficients trained on 25.5 hours of speech from the SCOTUS corpus (Supreme Court oral arguments). The phonemic transcriptions are taken from the CMU pronouncing dictionary. For a detailed, step-by-step tutorial on using P2FA for forced alignment, see Evanini et al. (2009b).

Several other recent studies have successfully used forced alignment as a tool in phonetics research. Examples of phonetic phenomena that have been investigated include vowel formants (Konopka and Pierrehumbert (2008), Yuan and Liberman (2008b), Chen et al. (2009), intonation (Anufryk (2008)), vowel duration (Tauberer and Evanini (2009)), and the quality of /l/ in English (Yuan and Liberman (2009)).

---

[1] P2FA can be freely downloaded from *http://www.ling.upenn.edu/phonetics/p2fa/*

4.2.1 Forced alignment accuracy

The procedure of orthographic transcription combined with forced alignment can quickly produce phone boundary labels for relatively large speech corpora. However, these automatically produced labels will necessarily be less accurate than manual labels due to the fact that it is impossible for acoustic models to account for all the variation that could be present in the speech signal. Nevertheless, for the purposes of automatic vowel analysis, a small degree of error in the phone alignments is acceptable, as long as the formant extraction procedure is still able to obtain accurate results from the output of the forced alignment. In one study using the P2FA forced alignment system, Yuan and Liberman (2008a) report that the vast majority of automatically generated word onset boundaries differed from the manual boundaries by less than 50 msec.

In order to test P2FA's performance on the current corpus, the phone boundaries for all stressed vowels from two word list recordings were manually segmented. One recording was taken from a face-to-face interview, and one was taken from a telephone interview. This was done in order to determine whether P2FA performs worse on telephone speech (since the acoustic models were not trained using telephone speech). The results for the two recordings were similar, though, and both are pooled together for the analysis below.

For this experiment, 324 vowels with primary stress were manually provided with onset and offset labels. Figure 4.1 shows a histogram comparing the absolute difference between the FA onset boundaries and the manual ones, and Figure 4.2 provides the same comparison for the vowel offset boundaries. These results are quite good: a difference of 10 msec or less is by far the most common result.

Table 4.1 summarizes the results presented in the two histograms in Figures 4.1 and 4.2, and shows that about two thirds of the automatically assigned boundaries fall within 20 msec of the manual ones in both cases, and all but one fall within 50 msec for the vowel onset. These numbers are promising, especially since accurate alignment performance

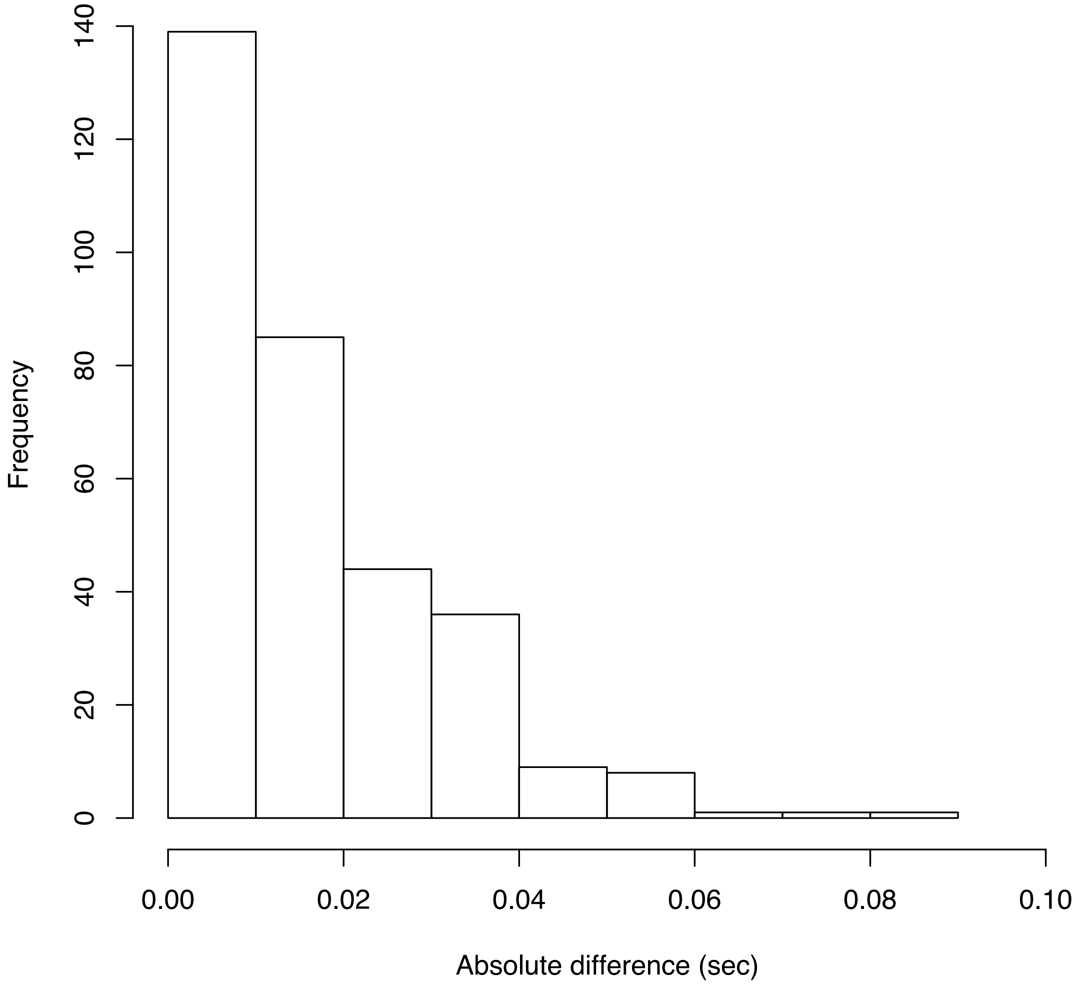

Figure 4.1: Difference between manual and FA vowel onset boundary labels

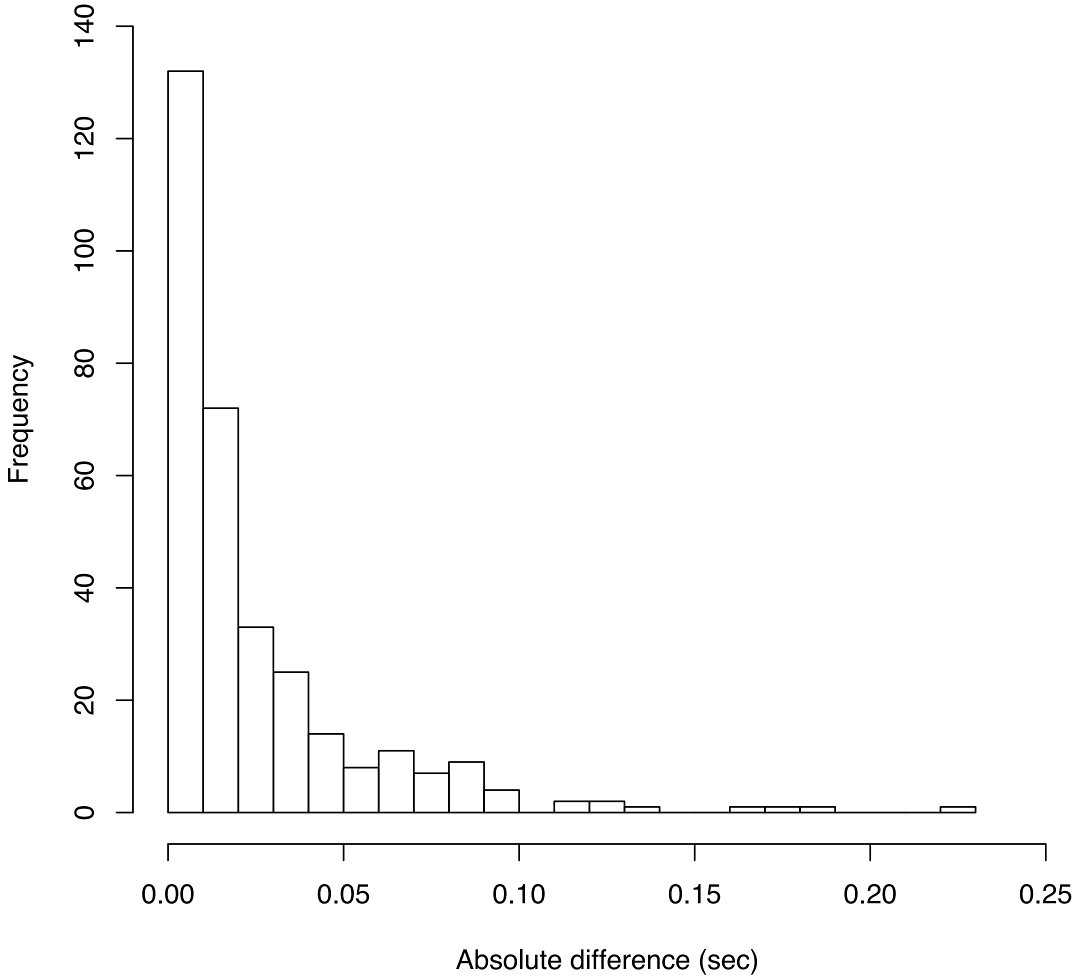

Figure 4.2: Difference between manual and FA vowel offset boundary labels

	within 20 msec	within 50 msec
V Onset	224 (69.1%)	323 (99.7%)
V Offset	204 (63.0%)	276 (85.2%)

Table 4.1: Comparison between FA and manual vowel boundaries (N = 324)

is most important for the vowel onset. As Section 4.3.4 will show, most manual vowel formant measurements are taken closer to the onset. The approach that will be adopted for automatic measurement point selection in Section 4.3 takes this into account, so alignment errors that occur in the vowel offset will generally not have an effect on automatic formant extraction.

As an additional metric for evaluating the performance of P2FA for the purpose of automatic formant extraction, we can consider the number of cases in which the manually selected measurement point falls inside of the vowel boundaries produced by P2FA. Cases in which this does not occur are serious errors, since the point marked as the correct measurement point by the human annotator is not available to the automatic vowel analysis system. Alternatively, for all cases in which the measurement point does fall inside the automatically produced vowel boundaries, the automatic vowel analysis system has the potential to choose the same point for formant measurement as the human annotator did.

In order to test this, manual vowel formant measurements were extracted for the 324 tokens from the preceding analysis, and the manual measurement points were compared with the vowel boundary labels produced by P2FA. In only 6 out of the 324 cases (less than 2%) did the manual vowel measurement point fall outside of the FA vowel boundary labels.[2] In four out of these six cases, the FA boundary errors were caused by neighboring

---

[2] There was one further case, in the word *route*, that initially appeared to be a mis-alignment. However, upon visual inspection it was clear that the FA boundaries were correct, and that the manual formant measurement had in fact been taken during the duration of the /r/ phone.

liquid consonants in the words *rider*, *fool*, and two tokens of the word *full*. The other two errors were in the words *hammer* and *began*, and were caused by a mis-alignment of the preceding segment.

Based on these results, it seems safe to conclude that the forced alignment results obtained by using P2FA will be accurate enough for conducting automatic vowel analysis. Though the mis-alignment rate for interview data is likely to be higher than the rate determined in this study for word list speech, this study suggests that any further errors that arise will be due to problems inherent to interview data, such as disfluencies, laughter, etc., and not due to inadequacies of the forced alignment system. In the case of interview data, there are generally thousands of tokens from each speaker. Thus, even with some measurement errors due to mis-alignments, the Law of Large Numbers will ensure that the means calculated for each vowel will be stable.[3]

4.3 Automatic measurement point selection

After the onset and offset of each vowel have been determined through forced alignment (as described in the preceding section), the next task is to decide on the point(s) within the duration of the vowel where the formant measurements are to be extracted. The standard sociolinguistic procedure is to manually select a single point based on visual analysis of the spectrogram combined with auditory analysis of the vowel. However, this method is quite labor-intensive, and it is impractical to analyze large corpora in this manner. Furthermore, the use of human intervention in every measurement introduces a degree of subjectivity to the results that makes comparisons with future studies by other researchers more difficult. In order to overcome these difficulties, it was decided that the measurement point for all

---

[3] However, for studies of change in progress where the position of individual tokens is crucial, such as the study of the regularity of sound change, this result may be unsatisfactory. In these cases, an additional examination of tokens with extreme values must be carried out in order to determine whether these tokens represent measurement errors or true outliers.

vowels should be determined automatically.

Despite the fact that a wide range of approaches to vowel analysis have been used, no studies have sytematically compared the results of the different methods to determine which is most accurate and reliable. Specifically, for the present study, it is necessary to determine which method can produce the best results when formant extraction is done completely automatically. In order to determine this, several different methods of automatic vowel analysis were compared using a set of manual formant measurements from the ANAE as a baseline.

In the following section I will first discuss my rationale for selecting only a single point to measure for each vowel. Then, in Section 4.3.2, I will describe the techniques for measurement point selection compared in the study. Next, in Section 4.3.3 I will describe the study and present its results. Finally, in Section 4.3.4 I will show how these results correspond well with two databases of manual formant measurements.

4.3.1 How many measurements to take?

The most common way to extract vowel formants in the course of either manual or automatic vowel analysis is to choose the formants from a single point in time to represent the entire vowel. This allows the analyst to easily disregard a vowel's onset and offset where the formants can be heavily influenced by the neighboring consonants. However, this approach is not without problems. The main drawback to abstracting a vowel's characteristics to a single point in time is that it discards all information about a vowel's trajectory over time. This clearly represents a large loss of information for diphthongal vowels, such as /ay/. It is also a problem for most phonologically monophthongal vowels, which are actually acoustically complex. This is espcially problematic in cases of dialect variation where a sound change in one dialect is causing a monophthong to become more diphthongal in nature, such as /i/ and /e/ in the South.

Recent sociolinguistic research has pursued methodology to compare entire formant trajectories across vowels in an attempt to preserve the inherent dynamic nature of most vowels. For instance, Watson and Harrington (1999) modelled the formant contours for many vowels using discrete cosine transform coefficients and used these values for vowel classification. Additionally, Nycz and Decker (2005) and Baker (2005) both used the technique of Smoothing Spline ANOVA to investigate /æ/-tensing and the merger of /o/ and /oh/. However, neither of these approaches have yet been successfully used to compare individual vowel tokens within a single speaker and mean values across speakers for a wide range of vowels. For these purposes, the abstraction of the two-dimensional vowel plot based on a single pair of F1 and F2 measurments for each vowel token remains a useful simplification. For this reason, and to facilitate comparison with other studies of vowel variation in North America (especially ANAE), I will continue the standard practice of representing vowels by a single formant value.

4.3.2 Methods for determining a vowel's measurement point

Once the decision has been made to use a single set of formant values to represent a vowel, there are two possible approaches to selecting this measurement point. First, the formant extraction procedure can be based on an examination of the time dimension. Under this approach, the formant values at a single point in time, such as the midpoint of the vowel, are taken to represent the vowel as a whole. A slightly more complex version of the same approach is to extract multiple measurements at fixed points and combine them in some manner to produce a single value (e.g., by taking the median value). Such techniques of vowel formant extraction based on the temporal dimension are attractive to researchers conducting automatic vowel analysis, since they are easy to implement.

The second general approach to formant extraction is based on an examination of the formant values themselves. This is done in an attempt to determine the formant values that are most representative of a given vowel, based on a model of that vowel's characteristic formant trajectories. For example, such approaches could look for a period of minimal change in the formant values or for an F1 minimum. Vowel measurements taken in this manner are often believed to correspond more closely to the "central tendency" of the vocalic nucleus (Labov et al. 2006:36) than measurements taken at a fixed point in time. However, automatic extraction of measurements based on formant trajectories can be difficult to automate; most researchers who employ these methods conduct manual vowel analysis.

By far the simplest approach, and the one most commonly used, is to select a point for measurement based on a percentage of the duration of the vowel. The majority of studies simply choose the midpoint for all vowels. A sampling of few studies illustrating the use of this methodology includes Steinlen (2002), Pierrehumbert et al. (2004), and Chen et al. (2009).

Selecting the midpoint is the easiest way to be sure that the formant values are not influenced by the transitions to the onset or offset consonants. However, measuring at the midpoint of a vowel is often not the best approach for non-monophthongal vowels. In diphthongal vowels, the transition from the target for the nucleus (which is assumed to be the desired measurement point) to the target for the glide often begins before the midpoint. This can be especially problematic in vowels where the nucleus and offglide are quite far apart in F1 and F2 space, e.g. /aw/ and /ay/.

As an example of a vowel for which the midpoint is not the ideal measurement point, Figure 4.3 shows a token of the word *house* spoken by ANAE speaker # 12, a 35-year-old man from Danville, IL. A measurement taken at the midpoint of the vowel /aw/, indicated by the vertical line in Figure 4.3, produces an F1 value of 651 Hz and an F2 value of 1076 Hz. An examination of the vowel's trajectory, however, shows that the midpoint is already well into the offglide /w/, and suggests that these formant values may not be the

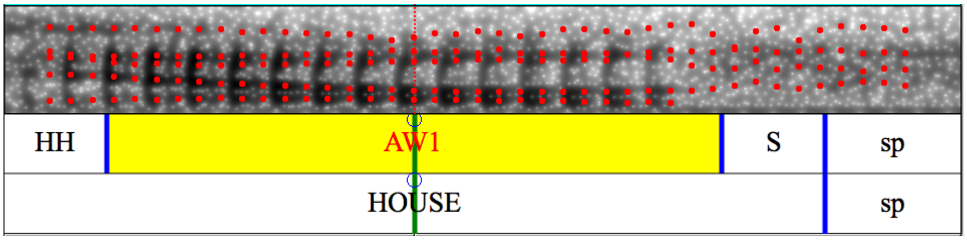

Figure 4.3: The midpoint of the vowel /aw/ in an ANAE token of *house* (formant values at this point do not represent the nucleus of the diphthong well, since they are influenced heavily by the offglide)

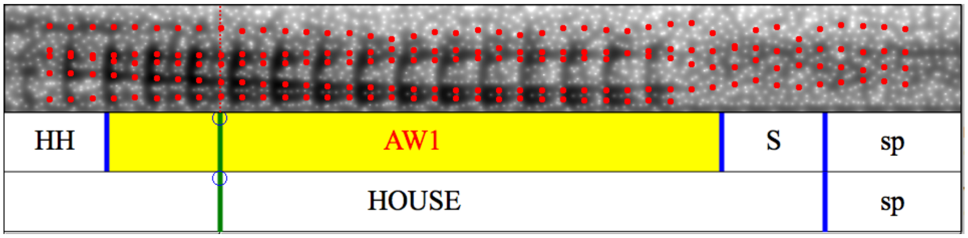

Figure 4.4: The measurement point selected manually by the ANAE annotators for a token of *house* (note that this point is much earlier than the midpoint)

best representation for this vowel. Indeed, the measurement taken manually by the ANAE annotator after an examination of the spectrogram was much earlier in the vowel. This manual measurement point is indicated in Figure 4.4. This measurement produced a value of 725 Hz for F1 and 1451 Hz for F2; thus, there is a difference of 375 Hz in the F2 domain for this token between the manual measurement and the one taken at the midpoint.

Due to such discrepencies, it might make more sense to measure a vowel at a point earlier than the midpoint. This point would then be more likely to represent the nature of a diphthong's nucleus, and not the transition from the nucleus to an offglide. For monophthongal vowels, there should not be much difference between a measurement taken at the midpoint and one taken earlier, since the formant values are relatively constant. In order to test the assumption that the midpoint is not the optimal point for extracting formant measurements, two other points earlier than the vowel's midpoint were tested, namely, the points at 1/3 and 1/4 of the way into the vowel's duration.

As discussed above, the other main method of selecting a vowel's measurement point is based on examination of the formant trajectory and a comparison of that trajectory to a model of the most representative portion of each vowel. By far the most common model that is used in vowel analysis looks for a steady state in the formants. The main purposes of using this model are to avoid the periods of formant movement that correspond to the transitions to and from neighboring consonants and to find the period of relative stability that is assumed to be the vowel's target. For diphthongal vowels, the assumption of a steady state target is clearly incorrect; however, it could be argued that the target of a diphthong's nucleus is a steady state.

An early and influential study of formant values in *hVd* words in American English used the steady state approach for manual vowel analysis, and described their methodology as targeting: "a part of the vowel following the influence of the [h] and preceding the influence of the [d], during which a practically steady state is reached" (Peterson and Barney 1952:177). Another influential study of vowel formants in North American English attempted to replicate this procedure (Hillenbrand et al. 1995); however, they also took measurements automatically at three fixed points in the vowel's duration.

There have been several attempts to automatically detect a vowel's steady state for the purposes of automatic formant analysis. An early approach developed in Lennig (1978) calculates a Coefficient of Change, c, at each formant measurement, i, based on the absolute value of the difference between that measurement and the preceding and following ones. The equation for calculating the coefficients is shown in 4.1.

$$c_i = \frac{|\mathrm{F1}_i - \mathrm{F1}_{i-1}| + |\mathrm{F1}_i - \mathrm{F1}_{i+1}|}{\mathrm{F1}_i} + \frac{|\mathrm{F2}_i - \mathrm{F2}_{i-1}| + |\mathrm{F2}_i - \mathrm{F2}_{i+1}|}{\mathrm{F2}_i} \tag{4.1}$$

The measurement for the vowel is then taken at the point where c_i is the smallest. This approach is rather simplistic, in that it only examines a single preceding and following formant measurement, when, in theory, a larger window could be considered. Furthermore, it treats F1 and F2 on the same scale—a similar approach could normalize F1 and F2 to account for the fact that the range of variation in F2 values is larger than in F1 values. However, Lennig (1978) reports that this procedure produced useful results for Parisian French, which has mostly monophthongal vowels. Other procedures that have been used to automatically detect steady states in vowels are reported in Miller (1989), ... For this study, the method from Lennig (1978) will be used, despite its potential flaws, due to its ease of implementation.

The steady state approach, however, is not optimal for English, since many vowels, even ones that are phonologically monophthongal, are characterized by dynamic formant contours. The authors of the ANAE recognized this fact, and instead based their model of a vowel's most representative point on the nature of how the formants change, specifically the peaks and valleys. Their approach is as follows:

> The central tendency of most short vowels and many long upgliding vowels is a downward movement of the tongue into the nucleus, followed by a rise out of the nucleus into the glide or following segment. The acoustic reflection of this fall and rise is a rise and fall in F1, with a maximal value of F1 representing the lowest point reached by the tongue. Vowels displaying this tendency were therefore measured at the point where F1 reached its maximal value. F2 was then measured at the same point (Labov et al. 2006:38).

There are, however, a few vowels that this procedure does not apply to, namely ingliding vowels. The ANAE authors explain as follows:

> The major exception to the principle of using the F1 maximum as a point of measurement occurs with those vowels whose central tendency is not so much a lowering and raising of the tongue as a movement of the tongue towards and then away from the front or rear periphery of the vowel space; these are ingliding vowels. In these cases, a point of inflection in F2, indicating maximum displacement toward the front or back periphery, was used as the point of

Method name	Algorithm for determining measurement point
Mid	measure at $t/2$
Third	measure at $t/3$
Fourth	measure at $t/4$
Lennig	measure at point where Coefficient of Change (see Equation 4.1) is smallest
ANAE	if vowel == /æ/, measure at F2 max elif vowel == /oh/, measure at F2 min else measure at F1 max

Table 4.2: Summary of automatic vowel analysis methods under comparison

> measurement, with F1 measured at the corresponding point. Vowels whose tendency was movement toward and away from their front periphery were measured at their F2 maxima; those moving toward and away from the rear periphery were measured at their F2 minima (Labov et al. 2006:38).

This procedure is problematic for a purely automatic approach, however, since it is difficult to determine a vowel's central tendency without examining its formant trajectories. For the purposes of this experiment, the following simplification was made: the vowels /æ/ and /oh/ were treated as ingliding, since they are the only two vowels that are often ingliding. Thus, /æ/ was measured at its F2 maximum, and /oh/ was measured at its F2 minimum.

In summary, Table 4.2 shows all of the methods discussed in this section that will be used in the study reported in the following section. The first three (Fourth, Third, and Mid) are simplistic, but easily automated, methods based on the vowel's temporal domain, whereas the last two (Lennig and ANAE) are somewhat more complex, and are based on the vowel's formant measurements.

Method	|Manual − Automatic|		|Manual − Automatic| / Manual	
	F1	F2	F1	F2
Third	64.7	216.1	10.4%	12.5%
Fourth	67.1	216.4	10.7%	12.5%
Mid	70.0	232.6	11.3%	13.5%
ANAE	103.8	247.8	18.5%	15.5%
Lennig	110.8	304.4	17.4%	18.8%

Table 4.3: Mean differences between manual and automatic formant measurements for five different measurement points (N = 110,399)

4.3.3 A comparison of methods for measurement point selection

In order to compare the methods described in the previous section, a study using the manual vowel formant measurements from the ANAE was conducted. This comparison used a subset of 110,399 pairs of un-normalized manual F1 and F2 measurements from 406 speakers that were obtained from the original Telsur log files. For each of these vowels that was measured manually by ANAE annotators, F1 and F2 values were obtained automatically according to the five procedures in Table 4.2. The overall results compare the mean absolute differences between the manual measurements and the automatic measurements for each method. In addition, since the range of raw Hz measurements is different for F1 and F2, the ratio of the measurement differences to the manual measurements was also calculated. These results are presented in Table 4.3.

The method that performs the best overall takes the formant measurements one third of the way through the vowel's duration, and the results at one fourth of the duration are quite similar. The two methods that select the measurement point based on the vowel's formant trajectory perform substantially worse. This, however, could potentially be due to formant tracking errors, since these two methods are much more sensitive to this type of error than methods based on the temporal domain. Whatever the cause, it is clear that the methods based on the vowel's duration produce better results in a purely automated formant extraction procedure.

The overall mean differences between manual and automatic formant measurements reported in Table 4.3 give a somewhat misleading picture of the performance results for each method, since the distributions are skewed. For example, a majority of the differences for the measurements taken at $t/3$ are less than the mean: 67.6% of them for F1 and 73.0% for F2. Thus, there is a long tail of large differences in the distribution, most likely caused by formant tracking errors. Figures 4.5 and 4.6 show histograms of the differences for the F1 and F2 values produced by the Third method. For both F1 and F2, the largest number of automatic measurements are within 5% of the manual measurements.

Thus, Table 4.3 suggests that a method for automatic vowel analysis based on selecting a fixed point in the time domain will produce better results than one based on an examination of the formant contours. Additionally, selecting either $t/3$ or $t/4$ for the measurement point achieves slightly better performance than the commonly used midpoint method.

4.3.4 Where are manual measurements taken?

The preceding section showed that the best results for automatic vowel analysis were attained by the method that measured F1 and F2 at the point represented by one third of the duration of the vowel. Further support for this approach can be obtained by examining large databases of manual measurements. The two corpora that will be examined here are the ANAE and Hillenbrand et al. (1995).

Hillenbrand et al. (1995) is a useful corpus for determining where within a vowel's duration humans most often take formant measurements. The utterances in it consist of 12 different isolated *hVd* words produced by 139 different speakers of North American English, for a total of 1669 tokens. All vowel onsets and offsets were annotated manually, and vowel formant measurements were taken manually for each token by two separate

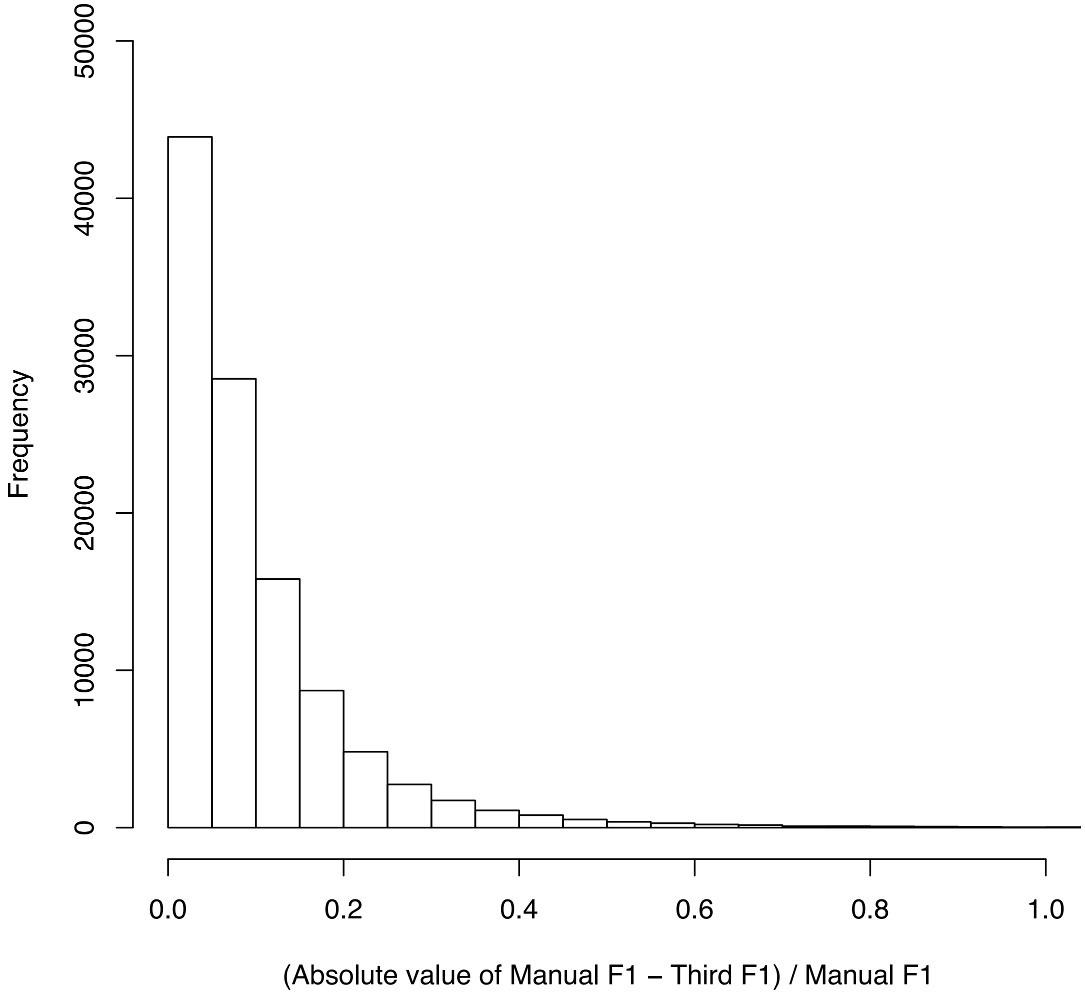

Figure 4.5: Histogram of differences bewteen automatic and manual F1 measurements using the Third method (N=110,399)

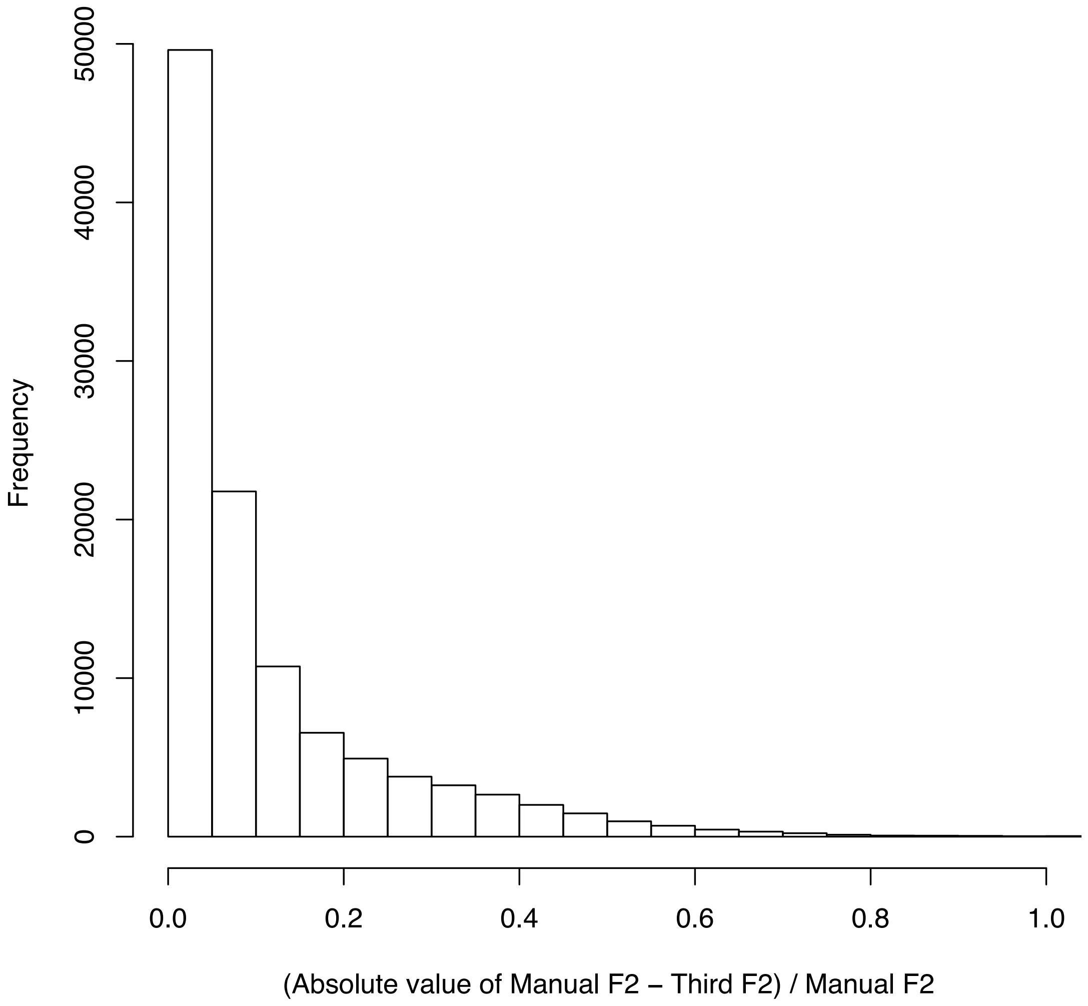

Figure 4.6: Histogram of differences bewteen automatic and manual F2 measurements using the Third method (N=110,399)

Corpus	Measurement point
Hillenbrand et al. (1995), Judge #1 (N=1,669)	0.27
Hillenbrand et al. (1995), Judge #2 (N=1,669)	0.29
ANAE (N=120,119)	0.32

Table 4.4: Average measurement points for all vowels in two corpora

judges. The procedure they used to select the measurement point was by looking for steady states in the vowel formants (Hillenbrand et al. 1995:3000).

The ANAE corpus is an even better dataset for determining where manual measurements are most often taken. It is a much larger corpus, and contains vowels in all segmental environments, not just *hVd*. However, the ANAE vowels were not segmented, so it is more difficult to determine the location of the measurement point relative to the vowel's onset and offset. In order to provide a solution to this problem, all ANAE tokens were aligned automatically on the phoneme-level through forced alignment (see Section 4.2). The resulting values for the vowel's onset and offset were then used together with the time stamp for where the formant measurements were taken to calculate the ratio of the measurement point to the vowel duration.

Table 4.4 presents the overall results for all vowels in the two corpora. The measurement point value represents the ratio of the distance from the onset (i.e., the difference between the measurement point and the vowel's onset) to the vowel's duration (i.e., the time difference between the vowel's offset and onset)—smaller values represent measurements taken closer to the onset, larger values represent measurements taken closer to the offset. The results for the ANAE corpus show that the mean measurement point is just around one third of the duration of the vowel, while the mean measurement point for the Hillenbrand et al. (1995) corpus is between one fourth and one third. The slightly earlier times in Hillenbrand et al. (1995) are probably due to the fact that all words in the corpus

Vowel	ANAE	Hillenbrand et al., Judge #1	Hillenbrand et al., Judge #2
/oy/	0.21	—	—
/ey/	0.22	0.22	0.23
/aw/	0.24	—	—
/ᴉ̩/	0.25	0.33	0.34
/iy/	0.26	0.38	0.40
/ay/	0.27	—	—
/ow/	0.28	0.20	0.23
/uw/	0.30	0.35	0.33
/i/	0.32	0.18	0.24
/æ/	0.34	0.19	0.22
/e/	0.36	0.31	0.33
/o/	0.38	0.27	0.28
/oh/	0.39	0.26	0.27
/ʌ/	0.39	0.34	0.33

Table 4.5: Average measurement points for each vowel in two corpora

are *hVd* words, and thus do not have formant transition effects due to the initial consonant.

Table 4.5 presents the mean measurement point on a per-vowel basis, ordered from earliest to latest point in the ANAE corpus (the diphthongs /oy/, /aw/, and /ay/ were not included in (Hillenbrand et al. 1995)). In general, the measurement points for the diphthongs and diphthongal long vowels are earlier than the measurement points for the monophthongal short vowels. This makes sense, because the measurer normally tries to select a measurement point that characterizes the vowel's nucleus and to avoid the portion of the vowel leading into its offglide. Overall, most of the vowels have mean measurement points in the range between $t/4$ and $t/3$. None of the vowels have mean measurement points later that 0.40, again indicating that the midpoint is not the optimal measurement point.

Figures 4.7 - 4.9 present a more complete view of the distributions used to calculate the mean overall measurement points reported in Table 4.4. Values less than 0 and greater than

1 for the ANAE data in Figure 4.7 are due to alignment errors–they represent cases where the manual measurement point was outside of the boundaries for the vowel, as determined by forced alignment. These errors, though, do not seem to be biased in either direction, so the location of the peak and the shape of the distribution should not be affected.

Finally, Figures 4.10 - 4.12 show the histogram distributions for all vowels in the two corpora (alignment errors in the ANAE corpus are excluded from Figure 4.10 by limiting the values on the x-axis to be between 0 and 1).[4] Nearly all of the distributions are approximately Gaussian, with a longer tail towards the right side. These distributions suggest that an automatic approach which selects the measurement point based on the most likely measurement point for each vowel from these distributions will be close to the manually-selected measurement point in most cases.

4.4 Formant prediction

4.4.1 Manual formant analysis

When sociolinguists conduct manual formant analysis, the procedure involves the following steps:

- Use the default formant prediction algorithm from a speech analysis software package (most commonly, LPC autocorrelation in Praat) to track the formants for a vowel token.
- Examine the predicted formants in conjunction with the spectrogram while listening to the sound file to determine if the predicted formants are correct.

---

[4] These figures use the ARPABET vowel symbols, since they are easier to display as labels in the graphs (they do not require the use of non-ASCII characters for /æ/ and /ʌ/. See Appendix A for a key for the ARPABET symbols.

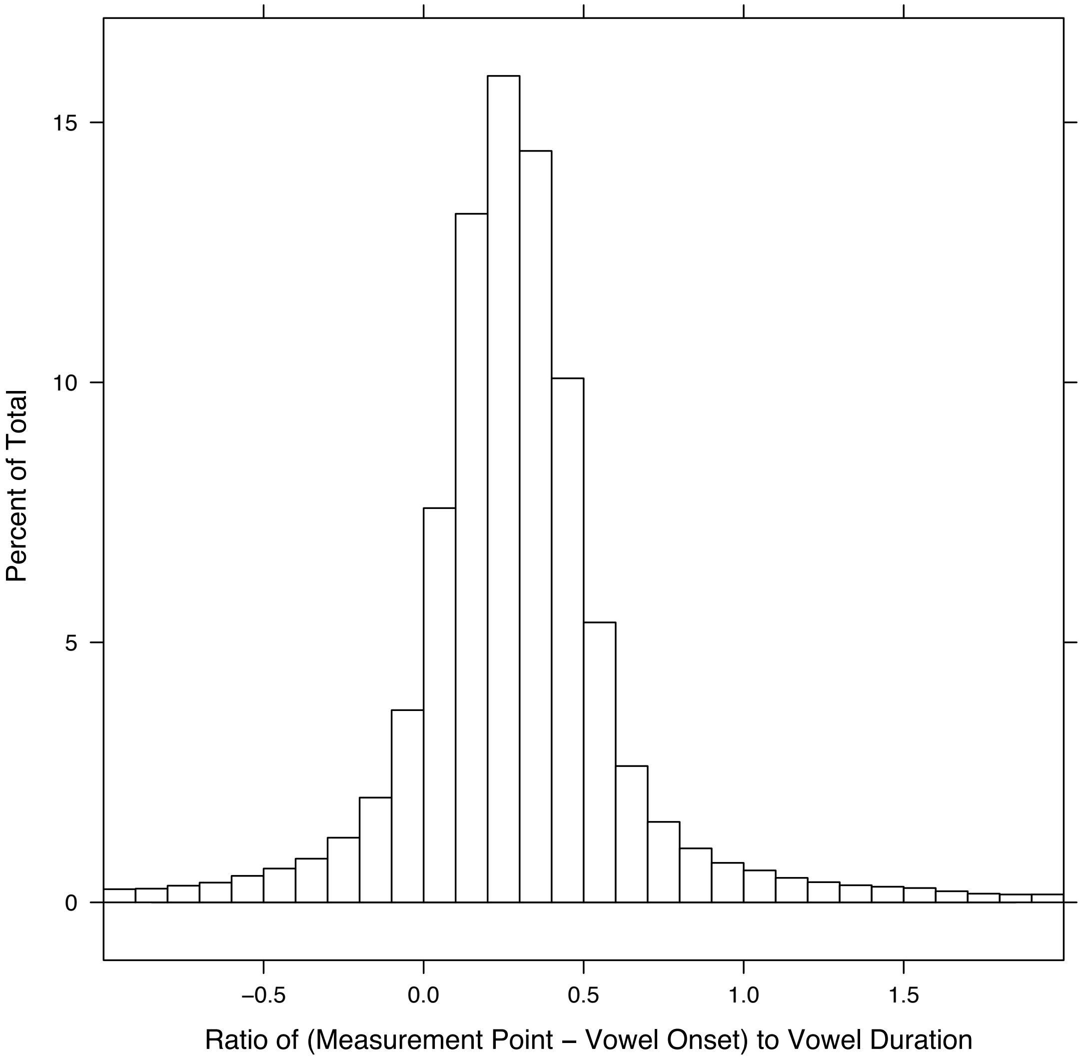

Figure 4.7: Location of measurement points for all ANAE vowels (values greater than 1 or less than 0 are caused by errors in the forced alignment procedure).

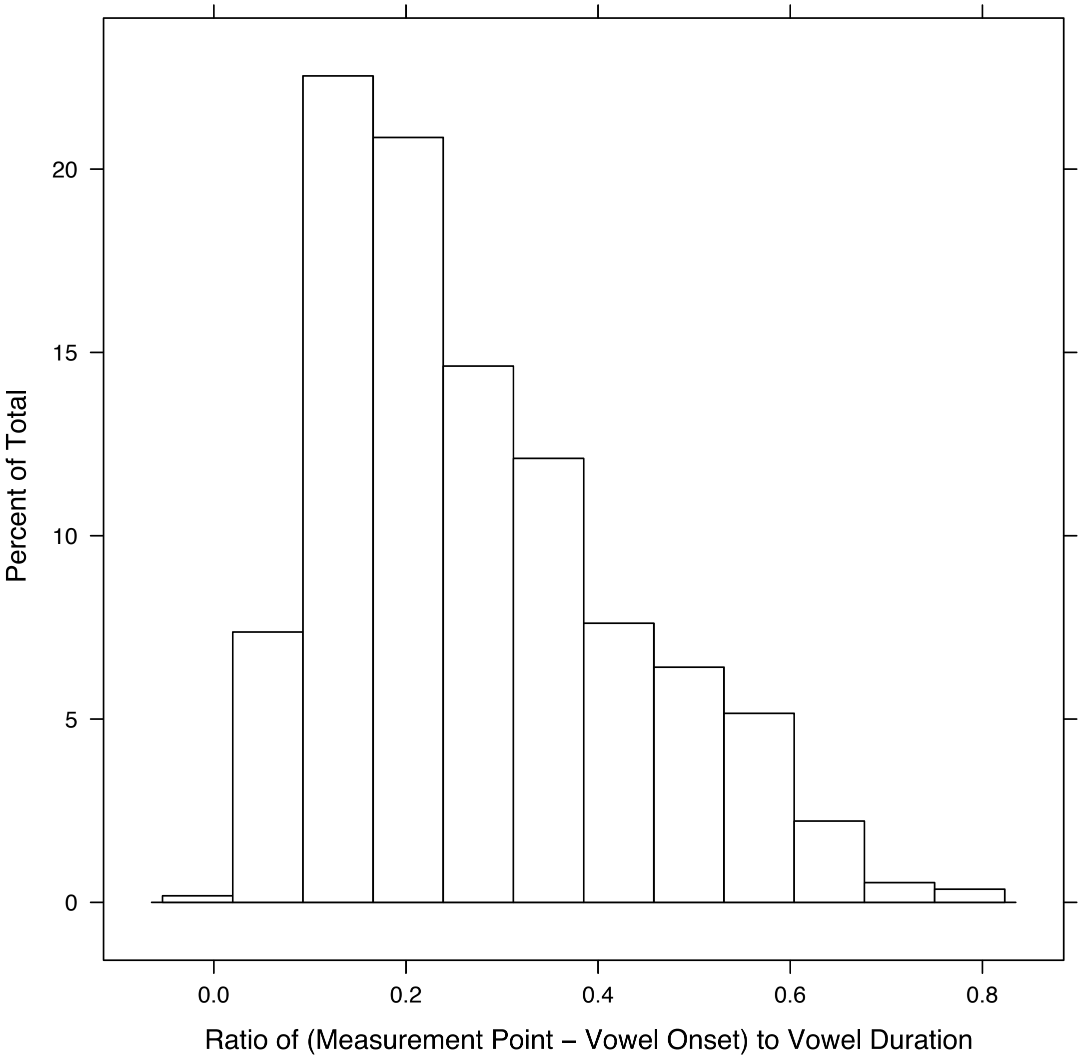

Figure 4.8: Location of measurement points for vowels in Hillenbrand et al. (1995), Judge #1

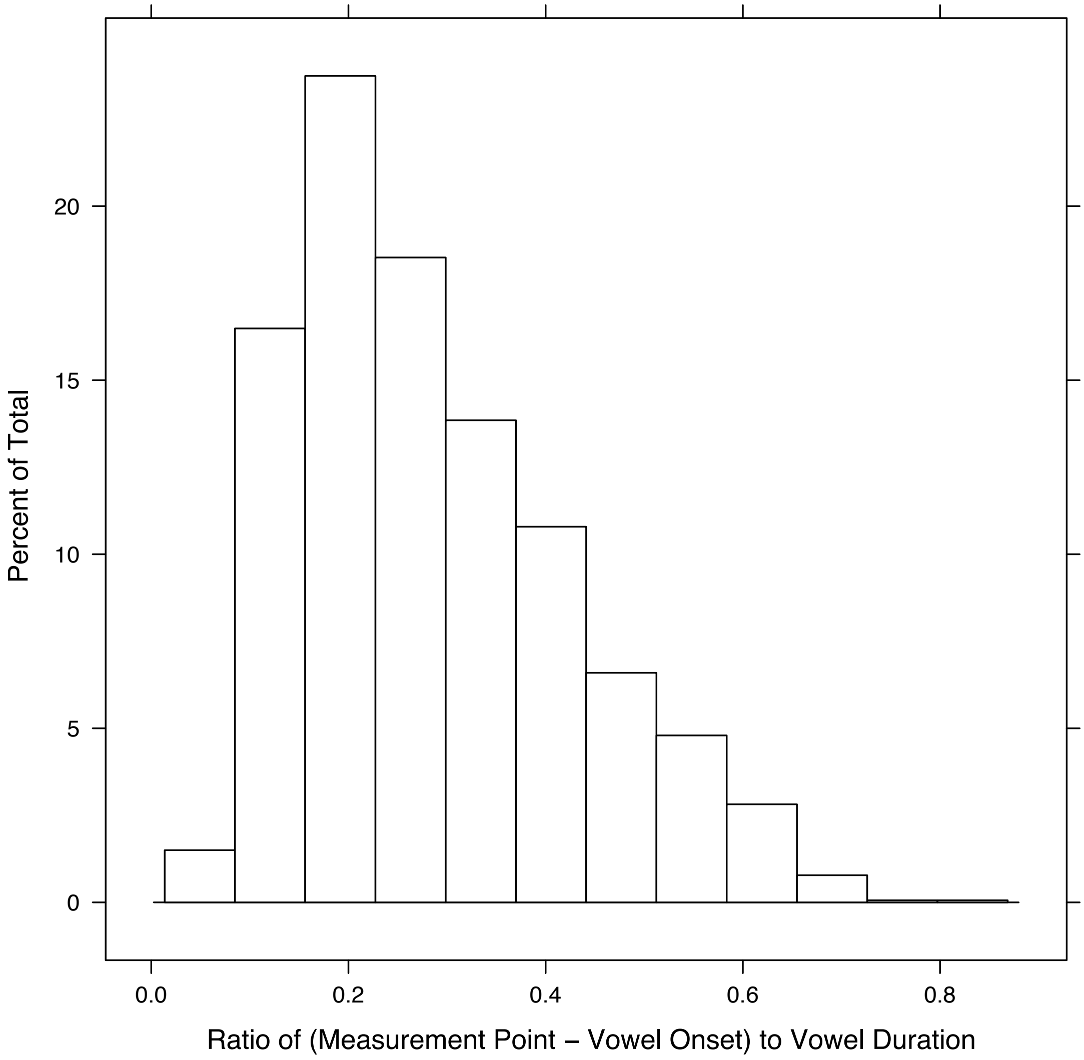

Figure 4.9: Location of measurement points for vowels in Hillenbrand et al. (1995), Judge #2

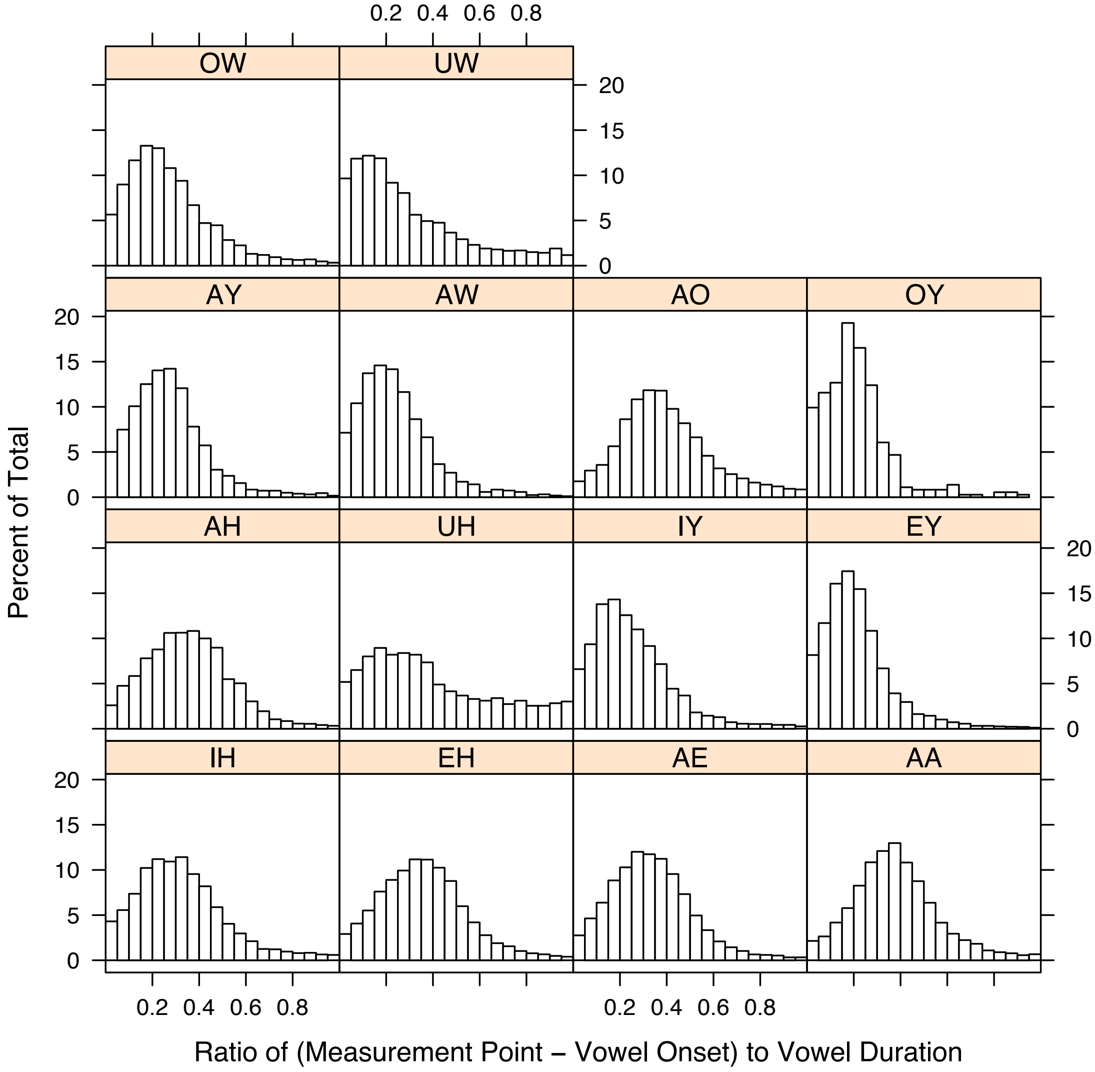

Figure 4.10: Location of measurement points by vowel class in ANAE

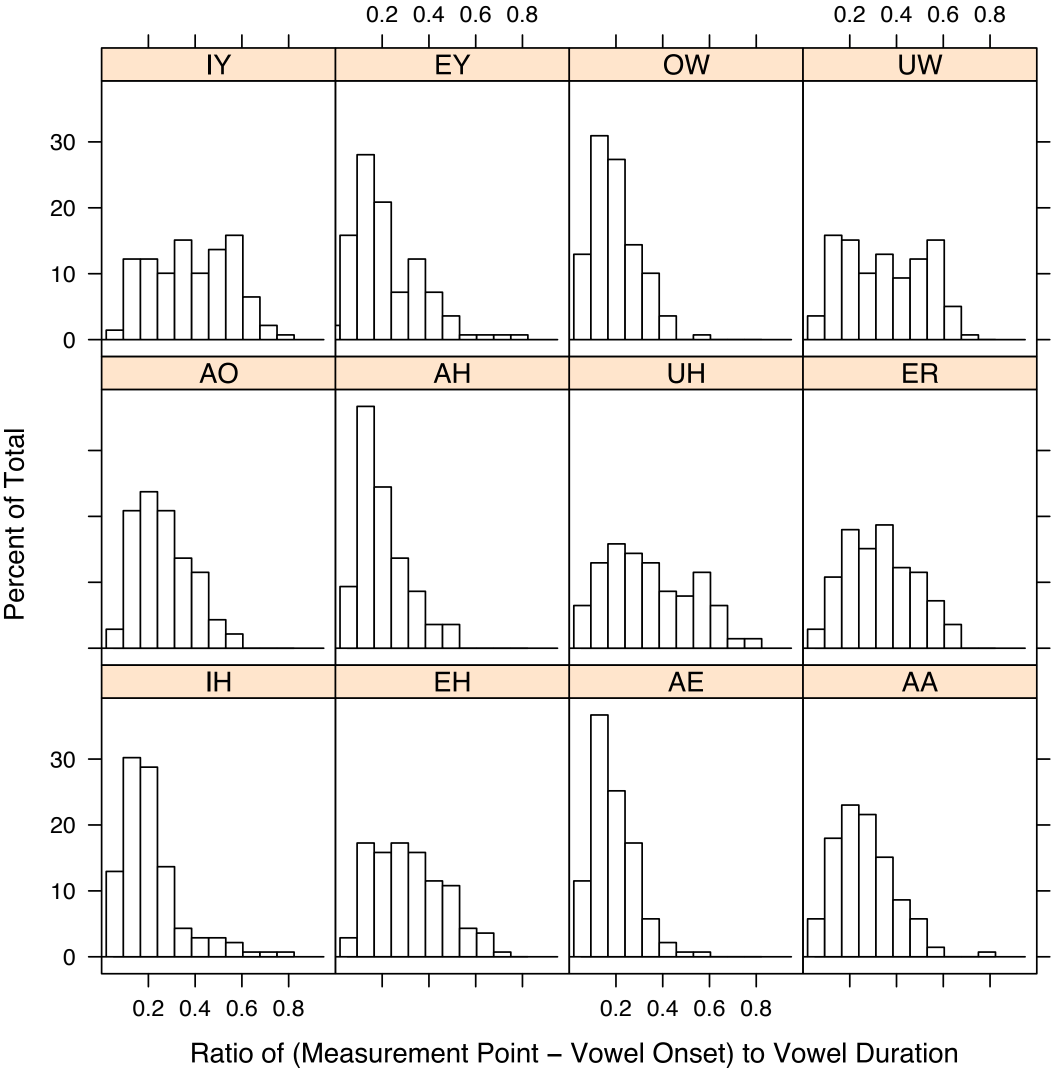

Figure 4.11: Location of measurement points by vowel class in Hillenbrand et al. (1995), Judge #1

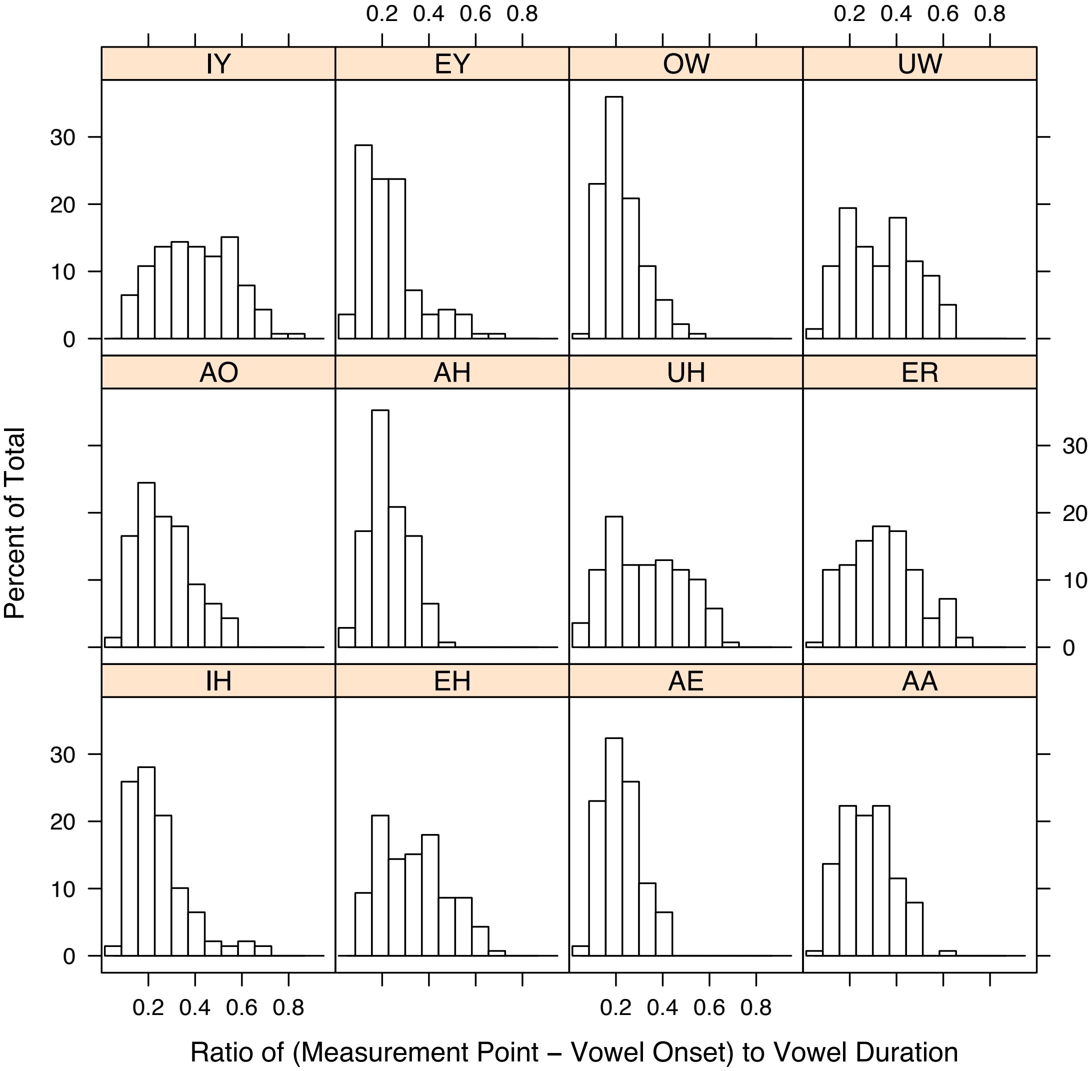

Figure 4.12: Location of measurement points by vowel class in Hillenbrand et al. (1995), Judge #2

- If the predicted formants are judged to be erroneous, modify the parameters of the formant tracker until the predicted formants are judged to be correct.[5]

The crucial addition in manual formant analysis, then, is that the researcher is able to listen to the vowel being analyzed and know its identity before deciding on appropriate values for F1 and F2. The researcher bases this decision on a model of expected formant values for each vowel; when formants predicted by the automatic formant tracker do not match the model of expected values, they are considered to be potential errors.

For example, Figure 4.13 shows a clear case of an erroneous formant track produced by a Praat LPC analysis for the vowel /æ/ in the word *sack*. This token was produced by a 49-year-old female speaker from Pittsburgh reading the word list (see Appendix C). The formant track in Figure 4.13 was produced by setting the `Maximum formant` parameter to a value of 5500, and the `Number of formants` parameter to 5 (these are the suggested default values for an adult female speaker). The first predicted formant shown in the Figure 4.13 has a value of around 775 Hz, which is within the expected range for /æ/. However, the second predicted formant, around 1350 Hz, would correspond to an extremely back token of /æ/, and does not correspond with either the auditory impression of the token or the darker areas of the spectrogram. This spurious second formant would lead to an incorrect value being produced by a standard automatic vowel analysis procedure, but is easily corrected in a manual procedure.

To produce a more accurate formant track for this token, the `Number of formants` parameter was changed from 5 to 4. The formants predicted by this analysis are shown in Figure 4.14. In this figure, the second predicted formant is around 2000 Hz, a much more likely value for /æ/, and there is no trace of the spurious second formant from Figure 4.13.

---

[5] In the case of the ANAE, a final step involving an inspection of the entire vowel system to detect outliers is also conducted.

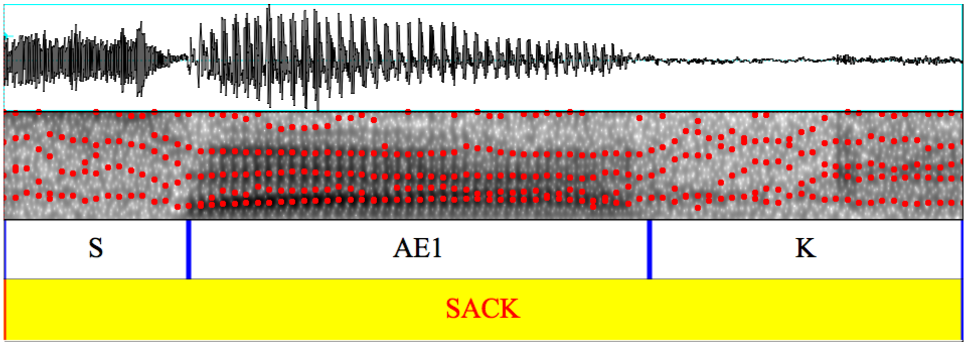

Figure 4.13: Praat LPC analysis of *sack* with 5 predicted formants, including a spurious second formant

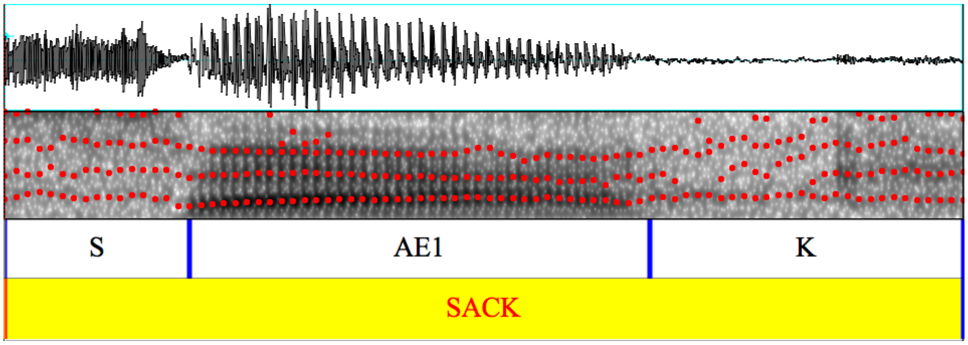

Figure 4.14: Praat LPC analysis of *sack* with 4 predicted formants

Such manual adjustments to the LPC parameters used in formant prediction are frequently necessary. For example, consider the ANAE corpus, which contains manually extracted formant measurements for over 125,000 vowels. The ANAE analysts were instructed to write a comment in their analysis log files whenever they had to modify the number of poles in the LPC analysis to produce a reasonable formant track for a given vowel. An examination of the log files (by using a regular expression pattern matching search to count all entries that had an annotation for the number of poles) shows that 12,847 such annotations were made; thus, at least 10% of the vowel tokens in the corpus required human intervention to produce accurate F1 and F2 values.

The prior knowledge that is used by the analyst in manual formant extraction, i.e., the expected distributions of F1 and F2 values for different vowels, can be approximated by an automatic formant prediction procedure through supervised learning. If the vowel's identity is known to the procedure beforehand (e.g., in the case of data that has undergone forced alignment), the human procedure described above can be simulated by the automatic procedure. If a suitable training corpus is available for constructing a model of expected formant values, the automated formant prediction procedure should be able to produce values that are similar to the ones produced by humans in manual extraction. The ANAE is an ideal corpus for this purpose, since it contains a large quantity of manually corrected F1 and F2 values for all vowels in North American English. The following section will describe a method for using the ANAE measurements to train a model of expected formant values that can be used in the formant prediction procedure.

4.4.2 Formant prediction using the Mahalanobis distance metric

The general approach taken by the proposed formant prediction method is to simulate the procedure used by a human annotator by incorporating prior knowledge of the distribution of formant and bandwidth combinations for specific vowels. For each vowel, a model of formant and bandwidth combinations was trained by computing the means and full covariance matrices for the manual F1 and F2 measurements from the ANAE corpus along with their respective bandwidths. Since bandwidth information is not provided in the ANAE corpus, the bandwidth values associated with automatically predicted formant values produced by the ESPS formant tracker were used when the formant values were close to the hand formants. The criterion for determining whether to use a token's ESPS bandwidth data in the training set was if both the predicted F1 and F2 values were within 7% of the respective hand measurements. This criterion led to a total of 61,048 training tokens (55% of the total corpus) with manual F1 and F2 values plus bandwidth data from the ESPS measurements (tests were also conducted with models trained using only F1 and F2 data from all 111,810 tokens, i.e., without bandwidth data, but this led to decreased performance). Additionally, the bandwidth measurements were converted to the log domain for both training and testing in order to make the bandwidth distributions closer to Gaussian (converting the formant measurements to the log domain did not improve performance).

To predict F1 and F2 using the proposed method for a given test vowel, all possible pairs of poles and their associated bandwidths returned by the ESPS LPC analysis for the vowel were considered. This results in $\binom{n}{2}$ test instances, where n is the number of poles provided by ESPS. Each test instance, x_i, is thus a vector consisting of four values: the two potential formant values and their associated bandwidths. To determine the most likely F1 and F2 values, the Mahalanobis distance, D_i, between the model for the vowel and each test instance x_i is computed. The equation for D_i is given in Equation 4.2, where μ and Σ are the means and covariance matrix for the formant and bandwidth values for the vowel.

$$D(x_i) = \sqrt{(x_i - \mu)^T \Sigma^{-1} (x_i - \mu)} \tag{4.2}$$

The test instance most likely to contain F1 and F2, $\hat{x}$, is predicted by minimizing D_i,

Pole (Hz)	Bandwidth (Hz)
514.2	709.1
921.1	61.6
1937.7	413.0
2362.1	489.9
3106.7	130.8
4080.7	581.4

Table 4.6: 6 pairs of poles and bandwidths returned by a 12th-order autocorrelation LPC analysis in ESPS for a token of the word *sack*

according to Equation 4.3. The pair of poles and bandwidths contained in x_i which has the smallest Mahalanobis distance to the training values, are then taken to be the values for F1 and F2 for the test vowel.

$$\hat{x} = \operatorname*{argmin}_{x_i} D(x_i) \tag{4.3}$$

For example, consider the vowel /æ/ in the token of *sack* shown in Figures 4.13 and 4.14. The six poles and their associated bandwidths produced by a 12-th order LPC analysis in ESPS for the point at $t/3$ are shown in Table 4.6.

There are thus $\binom{6}{2}$, i.e., 15, test instances consisting of pairs of pole and bandwidth information to be compared to the training values in order to predict F1 and F2. These test vectors are shown in Table 4.7.

Table 4.8 displays the mean values for the vowel /æ/ that were calculated by using the ANAE as the training database, and Table 4.9 displays the covariance matrix (in these two tables, F1 and F2 represent the first and second formants; B1 and B2 represent the bandwidths of the first and second formants, respectively).[6]

---

[6] As discussed above, the bandwidth values were converted to the log domain before other calculations were done. The bandwidth values are shown in their unconverted forms in Table 4.8 for ease of interpretation.

test instance	Poles		Bandwidths	
x_1	514.2	921.1	709.1	61.6
x_2	514.2	1937.7	709.1	413.0
x_3	514.2	2362.1	709.1	489.9
...	...	...	...	...
x_{15}	3106.7	4080.7	130.8	581.4

Table 4.7: 15 potential F1 and F2 pairs from the LPC data in Table 4.6 that will be considered by the Mahalanobis distance metric in Equation 4.2

	F1	F2	B1	B2
$\mu_{/\text{æ}/}$	728.8	1893.2	101.0	146.9

Table 4.8: $\mu_{/\text{æ}/}$, the mean formant and bandwidth values for /æ/ used by the Mahalanobis distance metric in Equation 4.2

To predict F1 and F2 for this token of /æ/ using the proposed method, the 15 Mahalanobis distances between the test vectors in Table 4.6 and the training data in Tables 4.8 and 4.9 are computed according to Equation 4.2. These distances are shown in Table 4.10, along with all 15 test vectors.

Finally, the test instance which minimizes the Mahalanobis distance to $\mu_{/\text{æ}/}$ and $\Sigma_{/\text{æ}/}$ is chosen to represent F1 and F2. For the 15 test instances and Mahalanobis distances

	F1	F2	B1	B2
F1	20,287.0	-945.8	23.7	19.1
F2		85,500.7	32.4	42.6
B1			0.41	0.06
B2				0.24

Table 4.9: $\Sigma_{/\text{æ}/}$, the covariance matrix for the formant and bandwidth values for /æ/ used by the Mahalanobis distance metric in Equation 4.2

test instance	Poles		Bandwidths		$D(x)$
x_1	514.2	921.1	709.1	61.6	5.6
x_2	514.2	1937.7	709.1	413.0	4.6
x_3	514.2	2362.1	709.1	489.9	4.6
x_4	514.2	3106.7	709.1	130.8	5.4
x_5	514.2	4080.7	709.1	581.4	8.0
$\boldsymbol{x_6}$	**921.1**	**1937.7**	**61.6**	**413.0**	**2.7**
x_7	921.1	2362.1	61.6	489.9	3.2
x_8	921.1	3106.7	61.6	130.8	5.2
x_9	921.1	4080.7	61.6	581.4	8.1
x_{10}	1937.7	2362.1	413.0	489.9	8.7
x_{11}	1937.7	3106.7	413.0	130.8	10.5
x_{12}	1937.7	4080.7	413.0	581.4	11.7
x_{13}	2362.1	3106.7	489.9	130.8	13.4
x_{14}	2362.1	4080.7	489.9	581.4	14.3
x_{15}	3106.7	4080.7	130.8	581.4	19.7

Table 4.10: Mahalanobis distances for the 15 potential F1 and F2 pairs from the LPC data in Table 4.6; the pair with the closest distance, x_6, is highlighted

shown in Table 4.10, the one which minimizes the distance is x_6, shown in bold. The F1 and F2 values predicted by the proposed method for this token of *sack* are thus 921.1 and 1937.7, with 61.6 and 413.0 as their respective bandwidths. The first pole returned by the LPC analysis at 514.2 Hz is rejected by the proposed method. Even though the distance between this pole and the mean F1 value for /æ/, 728.8 Hz, is about the same as the distance between the second pole and the mean F1 value, the first pole is rejected because its bandwidth is so much higher than the mean bandwidth value for F1 in the training data for /æ/. This behavior is desired, since lower bandwidth values for predicted poles generally correspond to actual spectral peaks, i.e., formants; higher bandwidth values for predicted poles correspond to more diffuse diffuse spectral peaks and suggest that they are less likely to be actual formants (Johnson 1997:88).

4.4.3 Evaluation of formants predicted by the Mahalnobis distance metric

In order to evaluate the performance of the proposed formant prediction method, a comparison was done between its predicted values and a set of baseline values produced by the default formant tracking algorithm in ESPS. Again, the ANAE corpus was used for this experiment, and the measurement points used for both sets of measurements were the manual measurement points for each vowel.[7]

The baseline set of automatic formant measurements was extracted by using the `formant` command from ESPS (Talkin 1987). Most default settings for the `formant` command were used, resulting in the following formant analysis parameters: 12 order autocorrelation LPC analysis using a 49 msec raised cosine window at 100 Hz with a preemphasis factor of 0.7. The only setting given a different value was the number of formants to predict: this was set to 3, since the corpus consists of telephone speech and the signal thus has a maximum frequency component of 3500 Hz (tests conducted with the default setting of 4 formants resulted in similar performance). After the `formant` command was run on each token, the predicted F1 and F2 values at the point in time closest to the hand measurement were extracted.

Table 4.11 shows the overall improvement over the ESPS formants using our predicted formants for all 111,810 tokens in the corpus, obtained by applying 10-fold cross validition to the entire data set (see Deng et al. (2006) for similar results comparing ESPS and manual formants on a different database). The proposed method reduces the global mean absolute difference from the hand measurements by 10% in F1 and 20% in F2. The performance improved for all 30 individual vowel classes in F1; in F2, a large performance gain was obtained for non-front vowels, whereas performance declined slightly for front non-low

---

[7] See (Evanini et al. 2009a) for a more complete description of this experiment.

vowels.

	Mean abs. difference		RMS difference	
	F1	F2	F1	F2
ESPS default	54.8	112.8	97.9	297.3
Proposed method	49.4	90.6	79.8	199.9

Table 4.11: Differences between two formant prediction methods and manual measurements (N = 111,810)

Figures 4.15 and 4.16 graphically illustrate the comparison between the proposed method and the baseline for 17,954 tokens from three vowel classes: /iy/ as in *heat*, /o/ as in *hot*, and /uw/ as in *hoot*.[8] The most striking difference between the two sets of predicted formant values is in the lower-left quadrant of the two plots: in the ESPS plot in Figure 4.15 there are many tokens of /o/ and /uw/ erroneously predicted to be in this quadrant, whereas the plot from the proposed method looks much more similar to the distributions obtained by the hand measurements shown in Figure 4.17.

---

[8] The tokens of /uw/ used in this experiment all occur after non-coronal onsets, and thus correspond to the /Kuw/ allophonic class in the ANAE notation. The tokens occuring after coronal onsets were excluded, since most North American dialects have substantial fronting of /uw/ in this environment. Limiting the analysis to /Kuw/ tokens means that the F2 values between the /iy/ and /uw/ groups overlap much less.

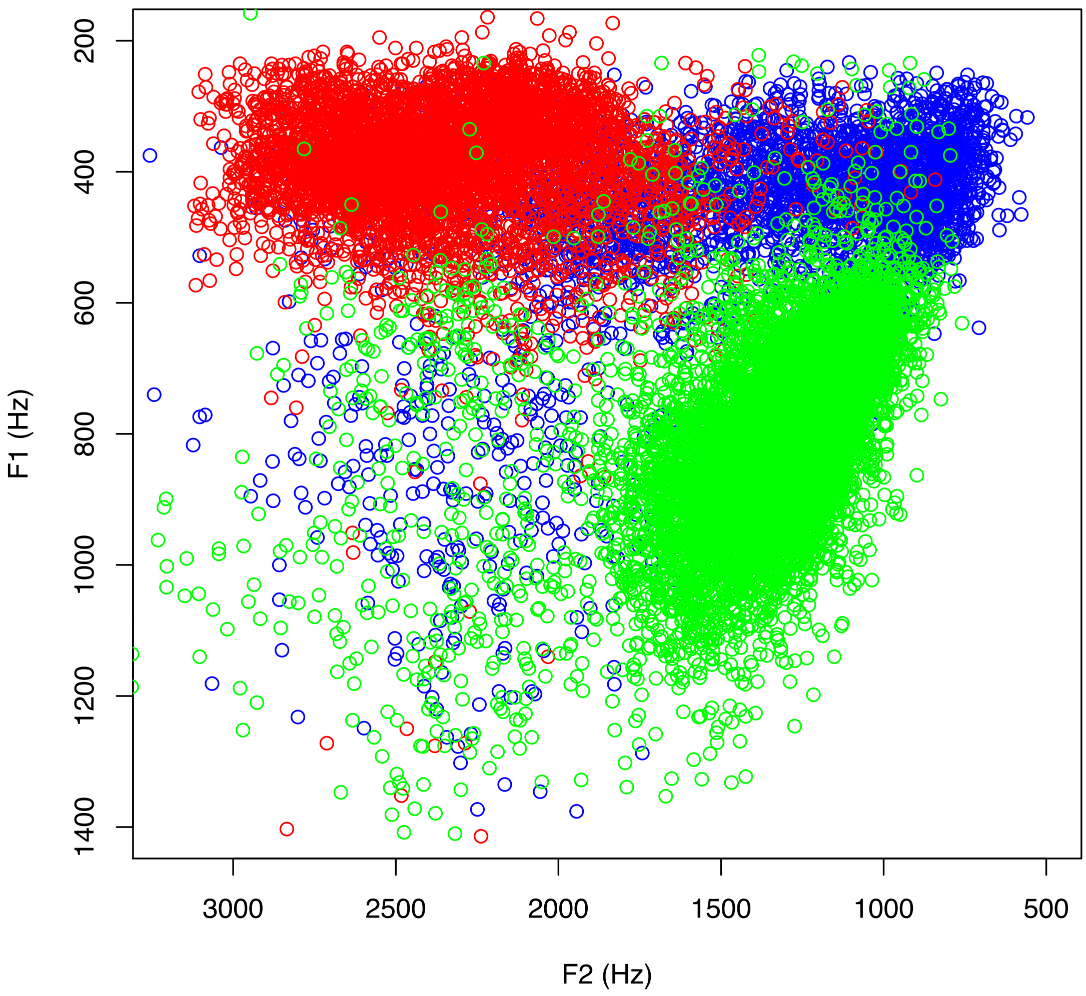

Figure 4.15: ESPS F1 and F2 measurements for /iy/, /uw/ and /o/

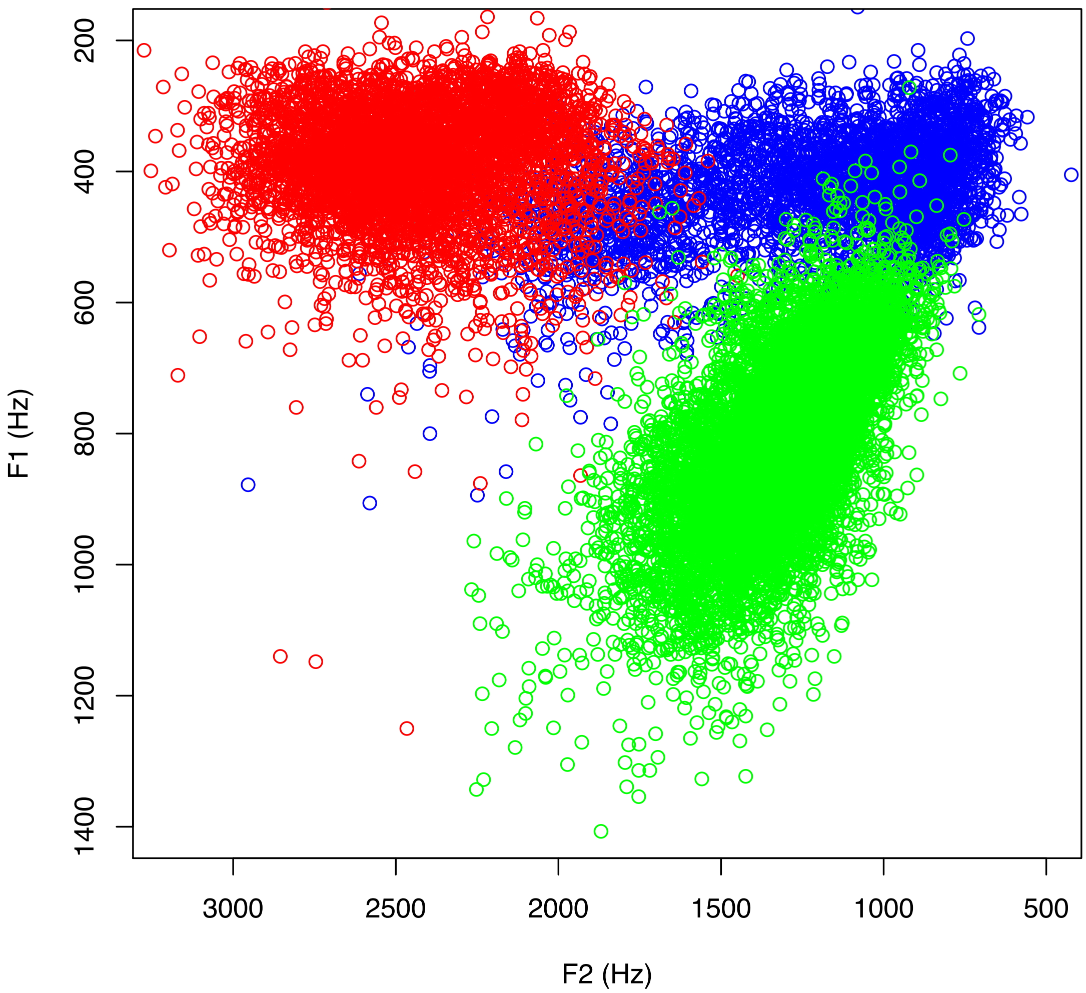

Figure 4.16: F1 and F2 measurements for /iy/, /uw/ and /o/ predicted by the Mahalanobis distance metric

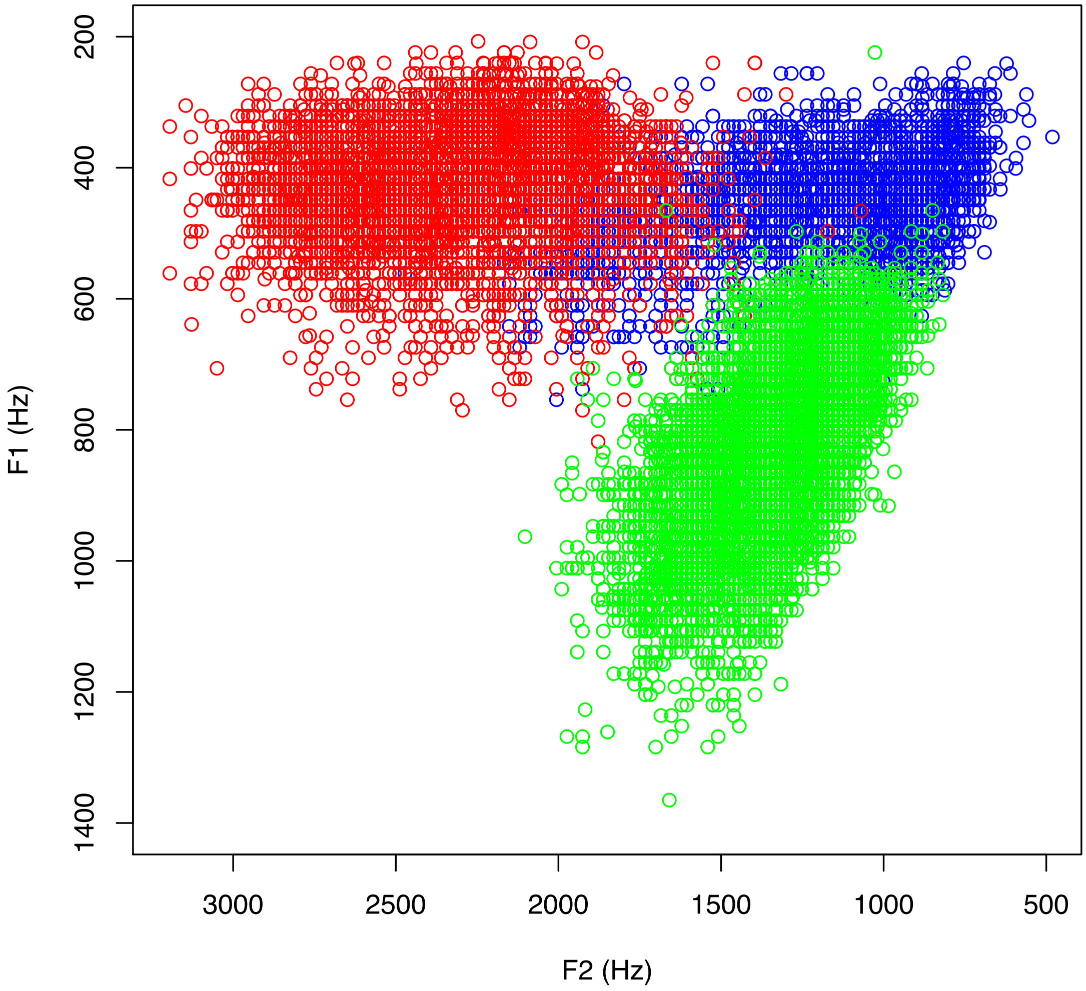

Figure 4.17: Manual F1 and F2 measurements for /iy/, /uw/ and /o/

In addition to evaluating the proposed formant prediction method by examining the distance for each individual measurement from the corresponding manual measurement, the most important test for the applicability of any automatic formant prediction method from a sociolinguist's perspective is whether the predicted values demonstrate the same group trends as the values measured by hand. That is, for any sociolinguistically important group of speakers based on sex, age, geographic region, etc., the group's vowel means from the predicted formants must demonstrate the same type of variation as the means from the hand measurements. Even when a small number of automatic formant measurements are gross errors, the sociolinguistic analysis can still be conducted successfully if they are not systematically biased in any direction.

As a demonstration of this approach to validating the automatic formants predicted for the ANAE, Figure 4.18 displays the vowels participating in the Northern Cities Shift (NCS). The normalized vowel means for NCS vowels (from non-prenasal tokens) for the 52 Inland North speakers are shown for the manual formant measurements (in red) and the formant measurements predicted by the proposed algorithm (in blue). As a reference for the non-shifted forms, Figure 4.18 also shows the mean values from the 332 non-Inland North speakers in black.

The automatically predicted formants capture the salient characteristics of the NCS quite well: tense AE is higher and fronter than EH, which has lowered and moved back; the NCS strong fronting of /o/ is also clearly visible. All of the NCS vowel means predicted by the automatic method are close enough to the ones produced by the hand measurements that the relative positions of the vowels are preserved. Thus, a sociolinguistic analysis using the automatically predicted formants would reach the same conclusions as one using the manual formants.

4.5 Combining proposed methods for measurement point selection and formant prediction

The results in Table 4.11 for the performance of the proposed formant prediction method are somewhat artificial, since the automatic measurements for each vowel were taken at the same point in time as the manual measurements. A more realistic evaluation of the

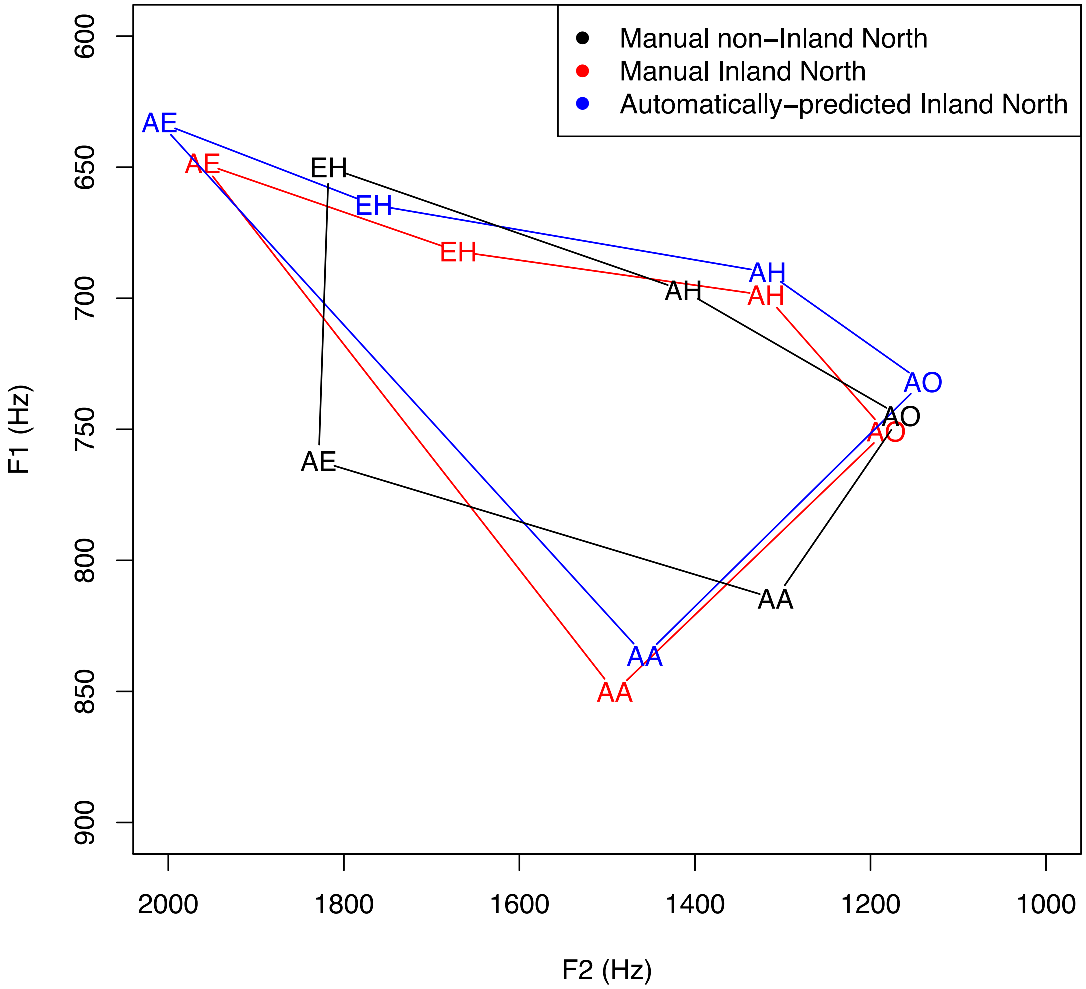

Figure 4.18: Comparison of NCS vowel means produced by manual and automatic measurements for 52 Inland North speakers from ANAE

Method	|Manual - Automatic|		|(Manual - Automatic)| / Manual	
	F1	F2	F1	F2
Praat formant tracker	64.7	216.1	10.4%	12.5%
Distance metric proposed in Section 4.4.2	57.8	126.4	9.0%	8.6%

Table 4.12: Mean differences between manual and automatic formant measurements taken at one third of the vowel's duration using two different formant prediction techniques

automatic vowel analysis approach would combine the proposed method for measurement point selection from Section 4.3.2 with the proposed method for formant prediction from Section 4.4. Table 4.12 presents the results of an experiment that did just this. Formant measurements at one third of the vowel's duration were predicted from the LPC poles using the proposed distance metric, and the results are compared to the formants predicted by Praat's default formant tracker, taken at the same point in time (these results are also presented in Table 4.3.

4.5.1 Comparison to manual formant tracking

Table 4.12 shows the performance of the proposed method for automatic formant tracking which involves taking formant measurements at one third of the vowel's duration and using a training set of manual measurements to predict the most likely formant values based on the poles and the bandwidths produced by an LPC analysis. The best performance attained an overall average difference from manual measurements of 57.8 Hz for F1 and 126.4 Hz for F2. This section will discuss the validity of this automatic approach by comparing these results to inter-annotator agreement in manual formant tracking studies.

A number of studies have compared inter-annotator agreement for manual formant tracking, and the results vary widely depending on the study and the technique used. In an early study, Lisker (1948) reported a random error of 12 Hz in his manual formant estimation procedure. Labov et al. (1972:32) report on four reliability tests that each compared two sets of manual formant extraction results using a total of three different annotators. The number of vowel tokens in the samples were relatively small: 25, 42, 42, and 20. The absolute F1 differences they report for the four studies are 40.5 Hz, 31.5 Hz, 35 Hz, and 33.5 Hz; the F2 differences are 84 Hz, 60 Hz, 38 Hz, and 64.8 Hz. Deng et al. (2006) report agreement results that are similar, although slightly larger in the F1 domain: they report an average absolute difference per frame of 55 Hz for F1 and 69 Hz for F2. However, the experiment in Deng et al. (2006) is much larger, and potentially more difficult than the one reported in Labov et al. (1972). For that task, manual formant measurements were taken at every frame during the duration of the vowel, not just at a single point representing its central tendency, and results were averaged over five pairs of labelers who each examined 16 distinct sentences from TIMIT. Finally, Hillenbrand et al. (1995:3101) report exceptionally good agreement between their two annotators: the absolute differences between formants were 9.2 (1.5%) for F1 and 17.6 (1.0%) for F2.

Another type of comparison can be made with studies that examined difference limens for vowel formant frequencies. These measurements correspond to the just noticable differences in formant frequencies, and thus represent an upper bound the accuracy that is necessary to be achieved in formant prediction. Flanagan (1955) measured difference limens for both F1 and F2 of a synthetic steady state vowel and three different frequencies, and reported limens that ranged between 3% and 5%. In an attempt to produce somewhat more natural stimuli, Mermelstein (1978) investigated difference limens in both steady state vowels and vowels with consonantal transitions. That study reports difference limens in the steady state context of 50 Hz for F1 and 142 Hz for F2. The difference limens for the consonantal context were 49 Hz for F1, 174 Hz for F2 in the /b/ context, and 199 Hz for F2 in the /g/ context. Thus, Mermelstein (1978) showed that the just noticable differences for F2 targets are higher when they are produced with consonantal transitions. This result is important, since it corresponds more closely to the situation that obtains in natural perception of continuous speech.

Chapter 5

Natural Break Maps

Before an in-depth analysis of the acoustic data is conducted, the F1 and F2 mean values for each speaker will be presented geographically in this section in a pre-theoretic manner, without the construction of isoglosses or manually-defined ranges of values. This will enable the actual distributions for each vowel to be observed in an unobstructed way. Furthermore, it facilitates direct and objective comparison with similar maps provided in Chapter 10 in the ANAE.

Each map in this section presents each speaker's mean value for the automatic F1 and F2 measurements for 17 vowel classes. The points for each speaker are color coded to correspond to four ranges of values that span the entire range of possible values for each vowel. For ease of comparison, the color scheme of the ANAE (Labov et al. 2006:77) was adopted. Under this representation, F1 values are arranged from lowest to highest in the order red, yellow, green, blue; i.e., a red symbol represents the highest vowel and a blue symbol represents the lowest vowel. For F2, values are arranged from lowest to highest in the order blue, green, yellow, red (the opposite order from F1); i.e., a red symbol represents the most front vowel, and a blue symbol represents the most back vowel.

The values for the four ranges in each map were determined by using the Natural Break algorithm in MapInfo (MapInfo Professional, version 7.8). This procedure attempts to find natural clusters of individual data points by iteratively minimizing the distance between the individual data values and the average value for each range until the clusters are as homogenous as possible. The MapInfo documentation does not provide any details about the implementation of their algorithm (such as how the clusters are initialized or whether multiple rounds of clustering are performed to avoid local minima), but they cite Jenks and Caspall (1971) as the source for their algorithm. In essence, the algorithm developed in that paper is a one-dimensional *k*-means clustering procedure[1] where the stopping criteria take into account a set of four accuracy indices defined by Jenks and Caspall (1971). For an initial examination of cartographic data, the use of such clustering techniques for determining the ranges on the map is preferable to defining the ranges by hand. This is because this procedure minimizes any subjective bias that could arise due to the researcher's preconceived notions of which groupings are important.

The natural break maps in this chapter were created from a subset of 45,756 tokens from the total of 112,848 automatic vowel measurements that were extracted from the corpus (see Section 3.8 for information about how these tokens were selected). The mean F1 and F2 values were computed from these tokens for the following 17 vowel classes: /i/, /e/, /ae/, /æN/, /o/, /ʌ/, /u/, /iy/, /ey/, /ayV/, /ay0/, /oy/, /Tuw/, /Kuw/, /ow/, /aw/, and /oh/. Three pairs from this set of vowel classes represent allophones of the same phonemic class: 1) /æN/ signifies tokens of /æ/ occuring before nasals, 2) /ayV/ signifies tokens of /ay/ that occur before voiced consonants or word-finally, whereas /ay0/ signifies /ay/ before voiceless consonants, and 3) /Tuw/ signifies /uw/ after coronal onsets, whereas /Kuw/ signifies /uw/ after non-coronal onsets.

The maps below contain F1 and F2 means for these 17 vowel classes for the 88 speakers listed in Tables 3.1 through 3.11. They contain tokens from all stylistic contexts that were

---

[1] For a description of the *k*-means clustering algorithm see, for example, Bishop (2006:424–428).

availble for each speaker. Thus, tokens from word lists, reading passages, and interviews are combined. While not ideal, this combination of data from different stylistic contexts was deemed necessary in order to provide a more complete geographic representation, since interview data is missing for a subset of the speakers. For example, for most of the DARE speakers, only the tokens from the "Arthur the Rat" reading passage were analyzed. Additionally, a few speakers in my corpus were only recorded reading the word list—time constraints prevented the completion of a full interview. Thus, data from all stylistic contexts are combined in the maps, so that these speakers who lack interview datacan also be included. Furthermore, studies have shown that the NCS and the merger of /o/ and /oh/, two of the most important changes in the dialect boundary region around Erie, are occurring below the level of consciousness. Thus, the amount of style-shifting between different contexts is minimal, and the combination of data from these three sources is justified.

A few of the speakers who were only able to provide word list data and who were recorded in the early stages of my fieldwork did not provide any examples of the vowels /iy/, /ey/, /u/, /oy/, /ay/, and /uw/. (An early version of the word list did not contain any tokens of these vowels.) In these cases, the speaker's symbol is not shown on the map, but the symbols for the other speakers from the same town are left in place. For example, the maps for /u/ in Figures 5.13 and 5.14 show that the speaker from Ford City and four speakers from Erie are missing.

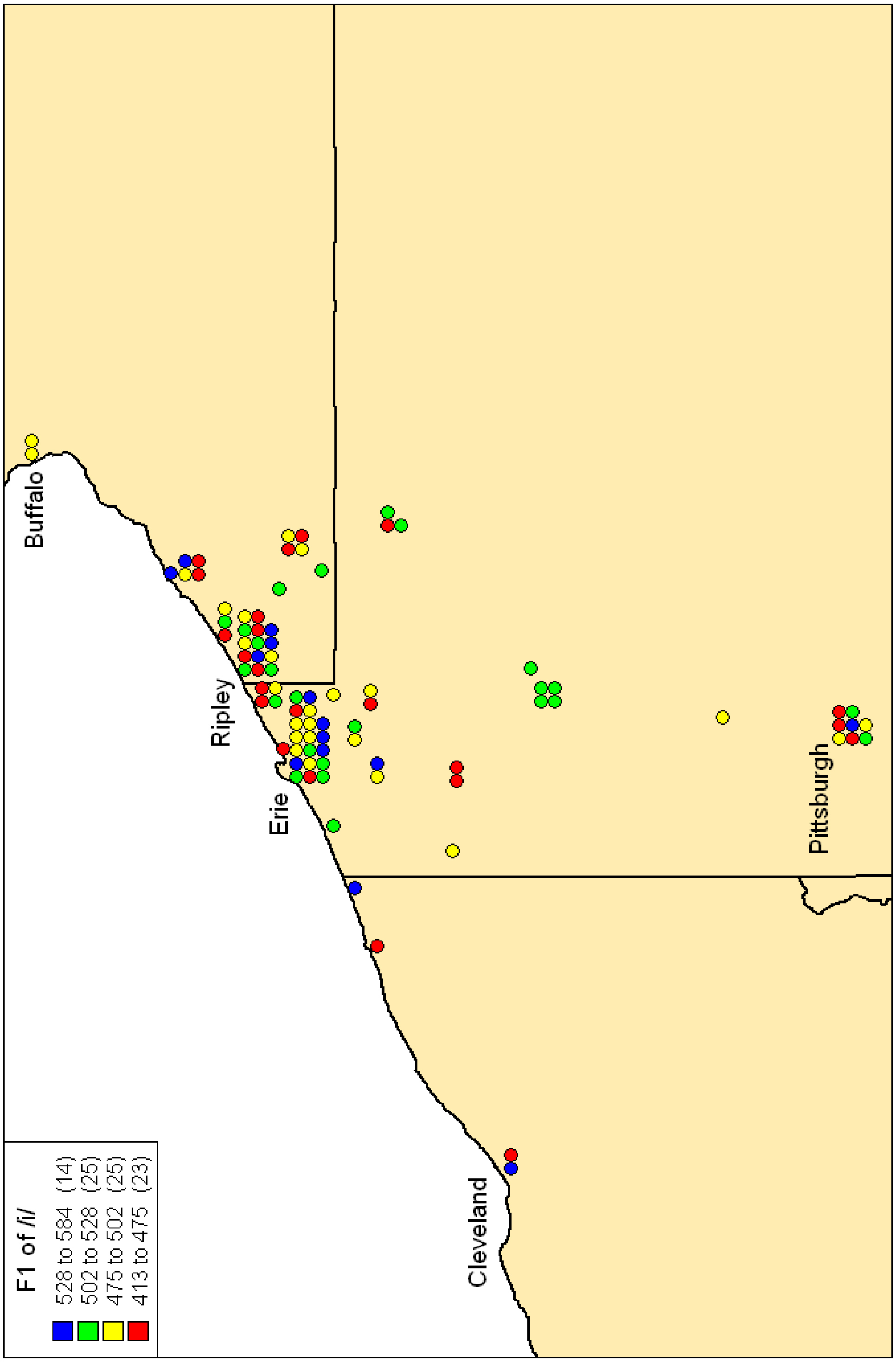

Figure 5.1: Natural break map for F1 of /i/

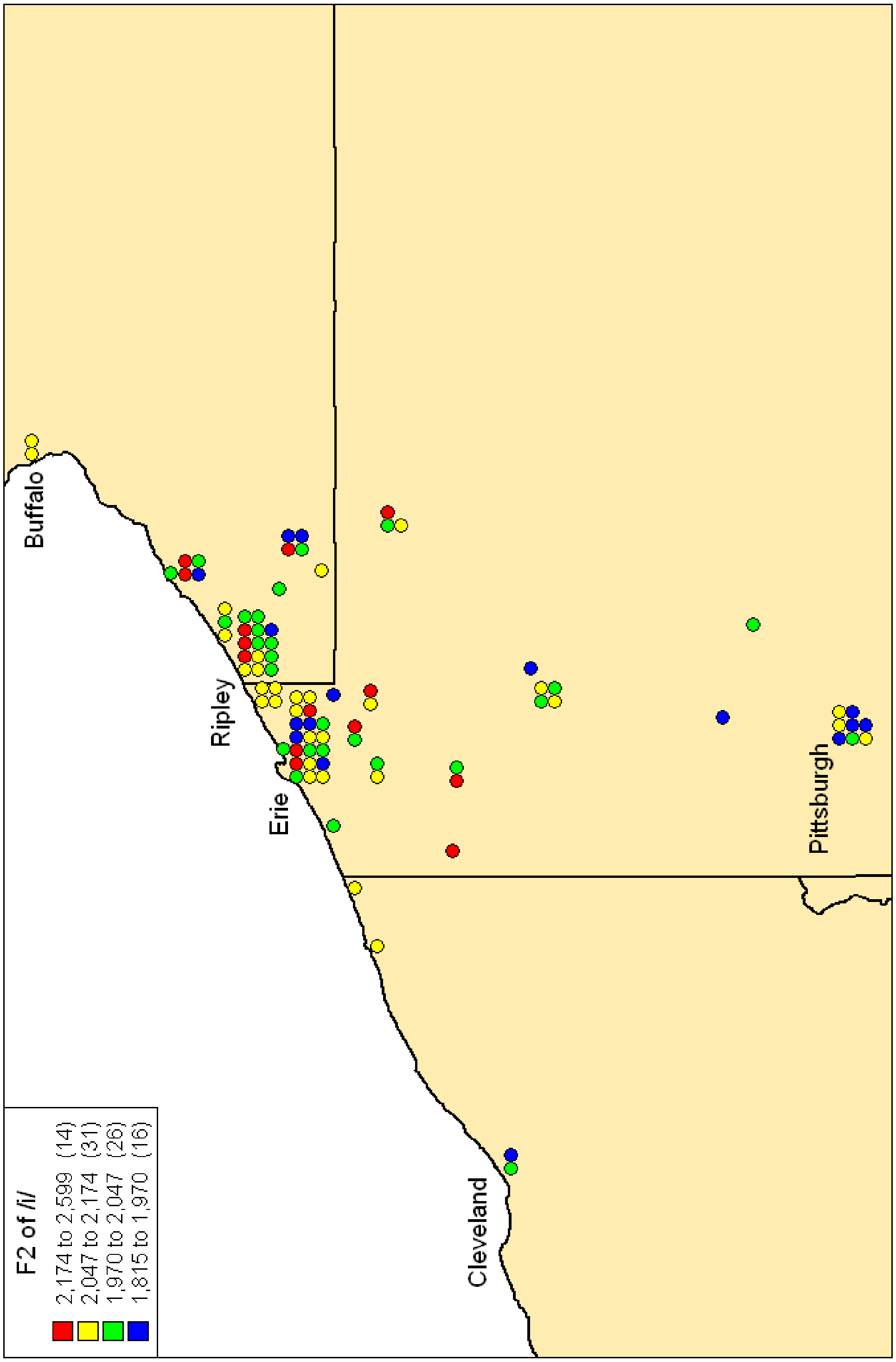

Figure 5.2: Natural break map for F2 of /i/

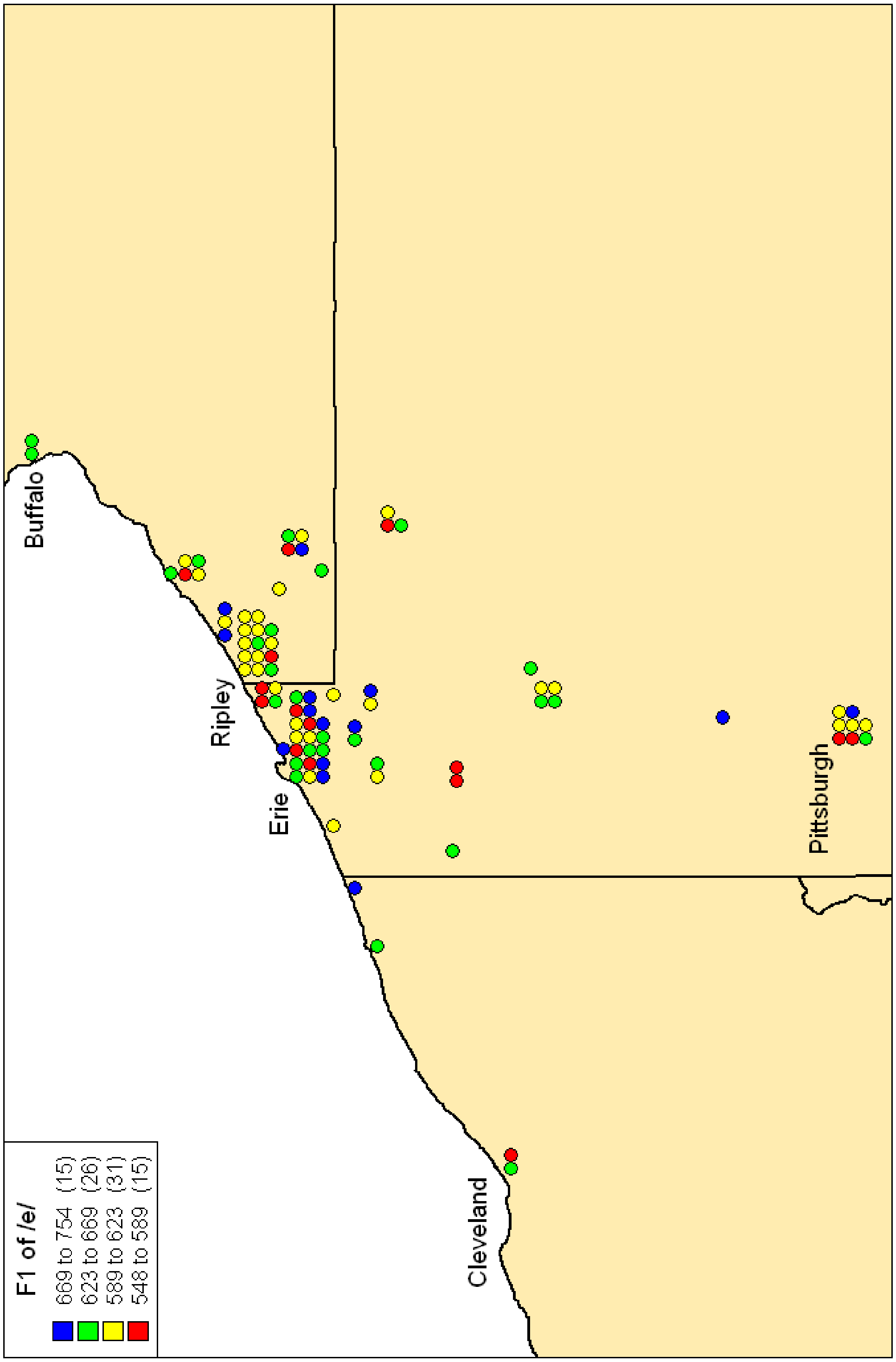

Figure 5.3: Natural break map for F1 of /e/

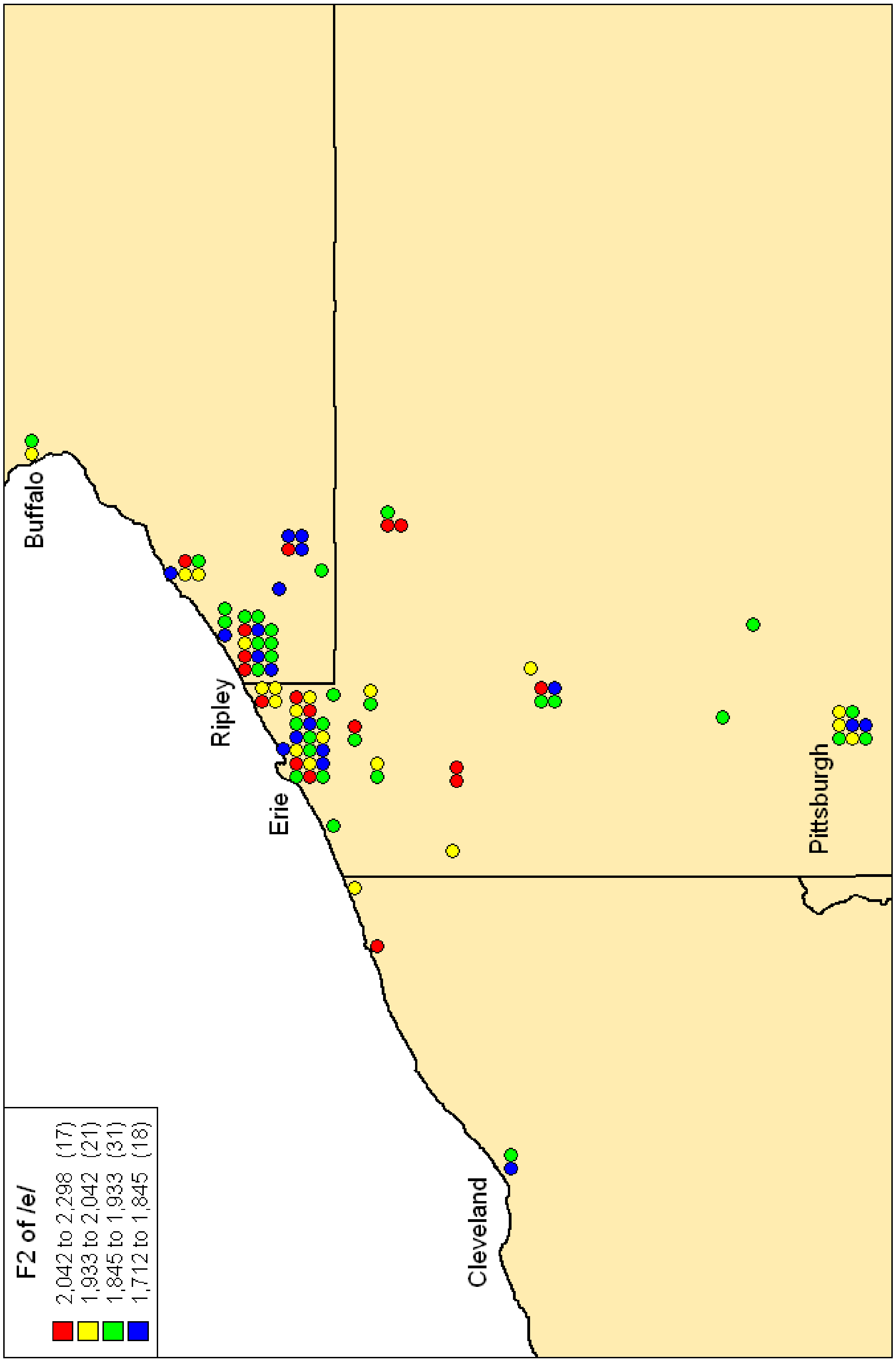

Figure 5.4: Natural break map for F2 of /e/

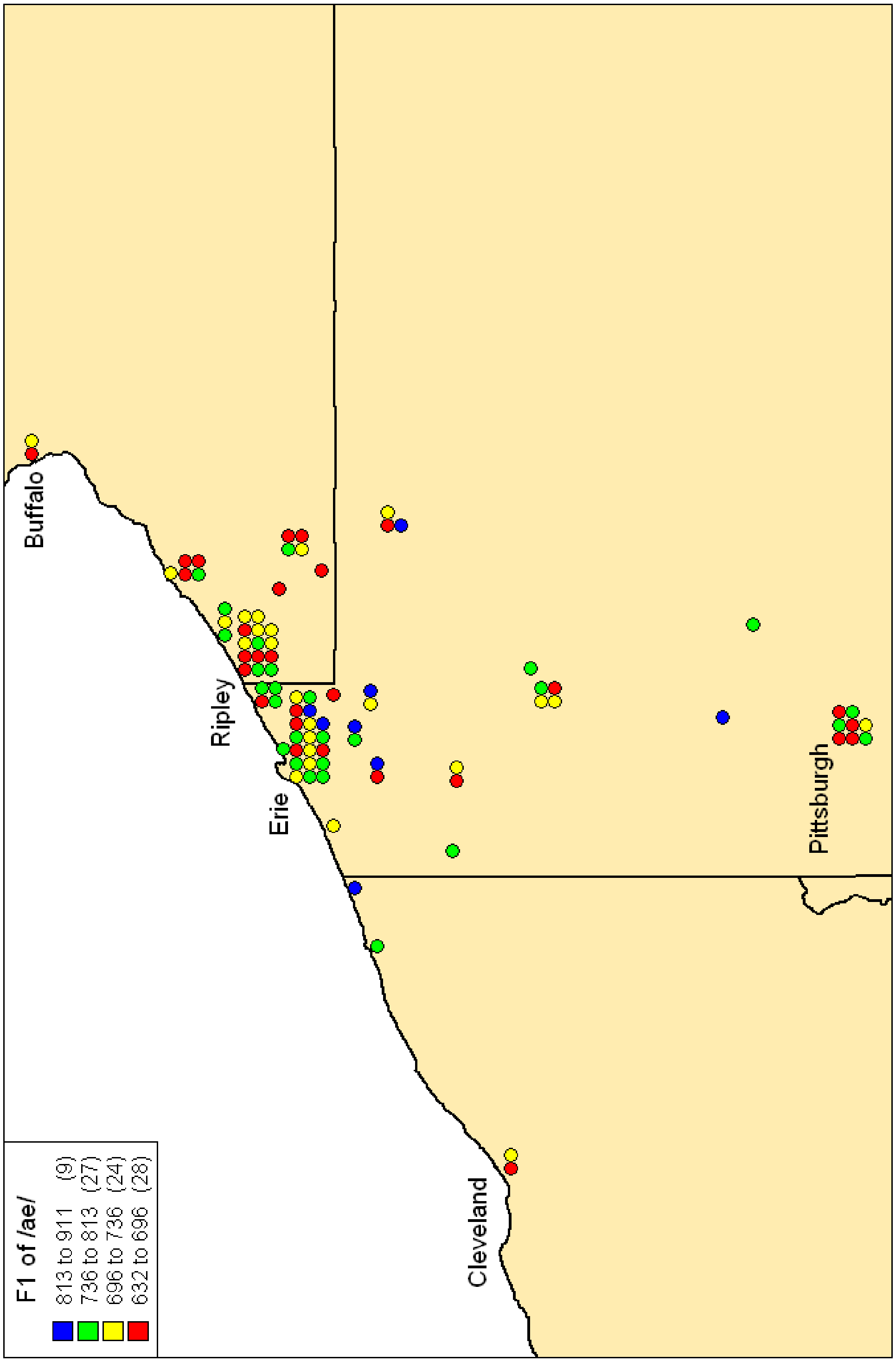

Figure 5.5: Natural break map for F1 of /æ/

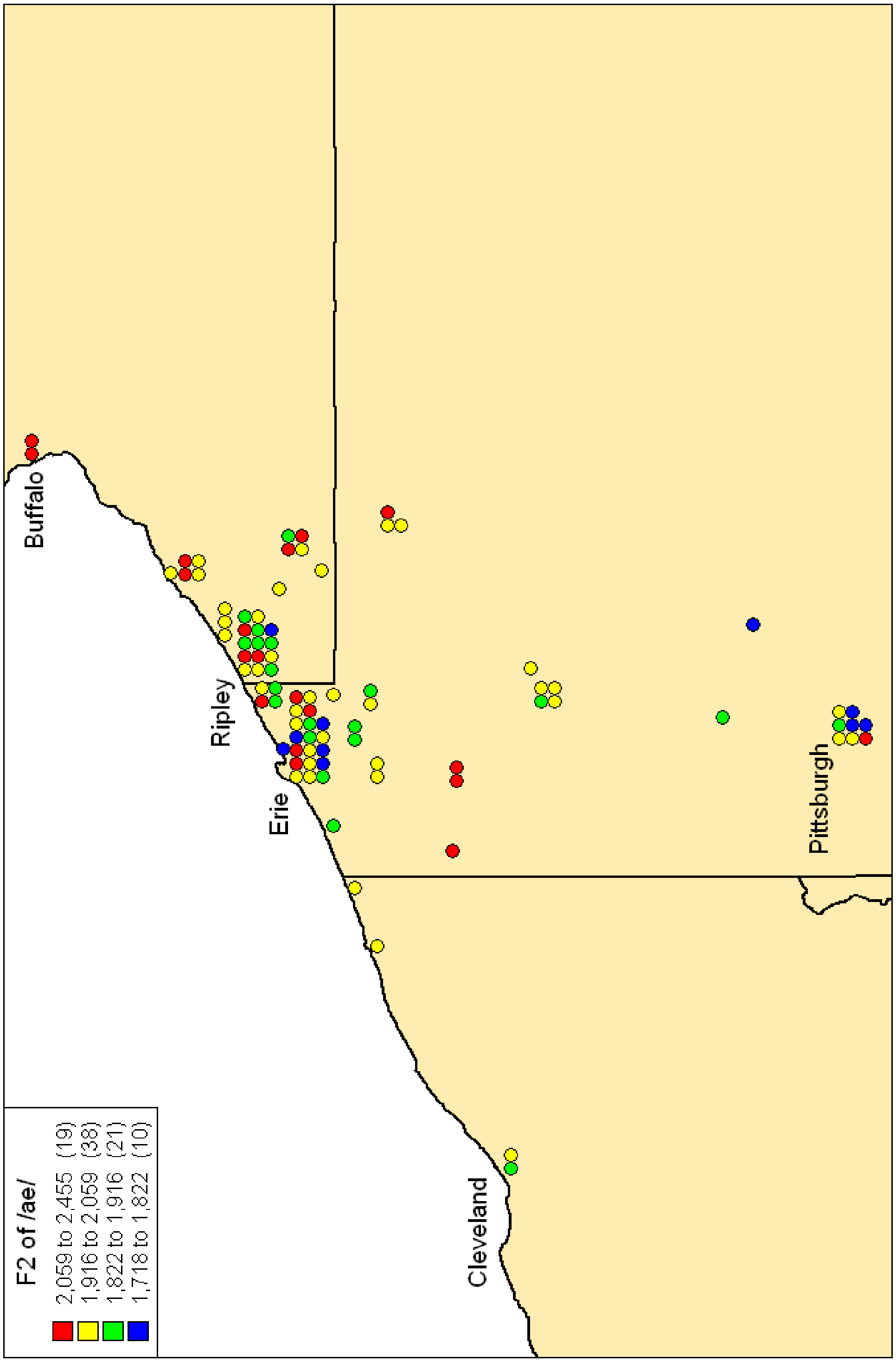

Figure 5.6: Natural break map for F2 of /æ/

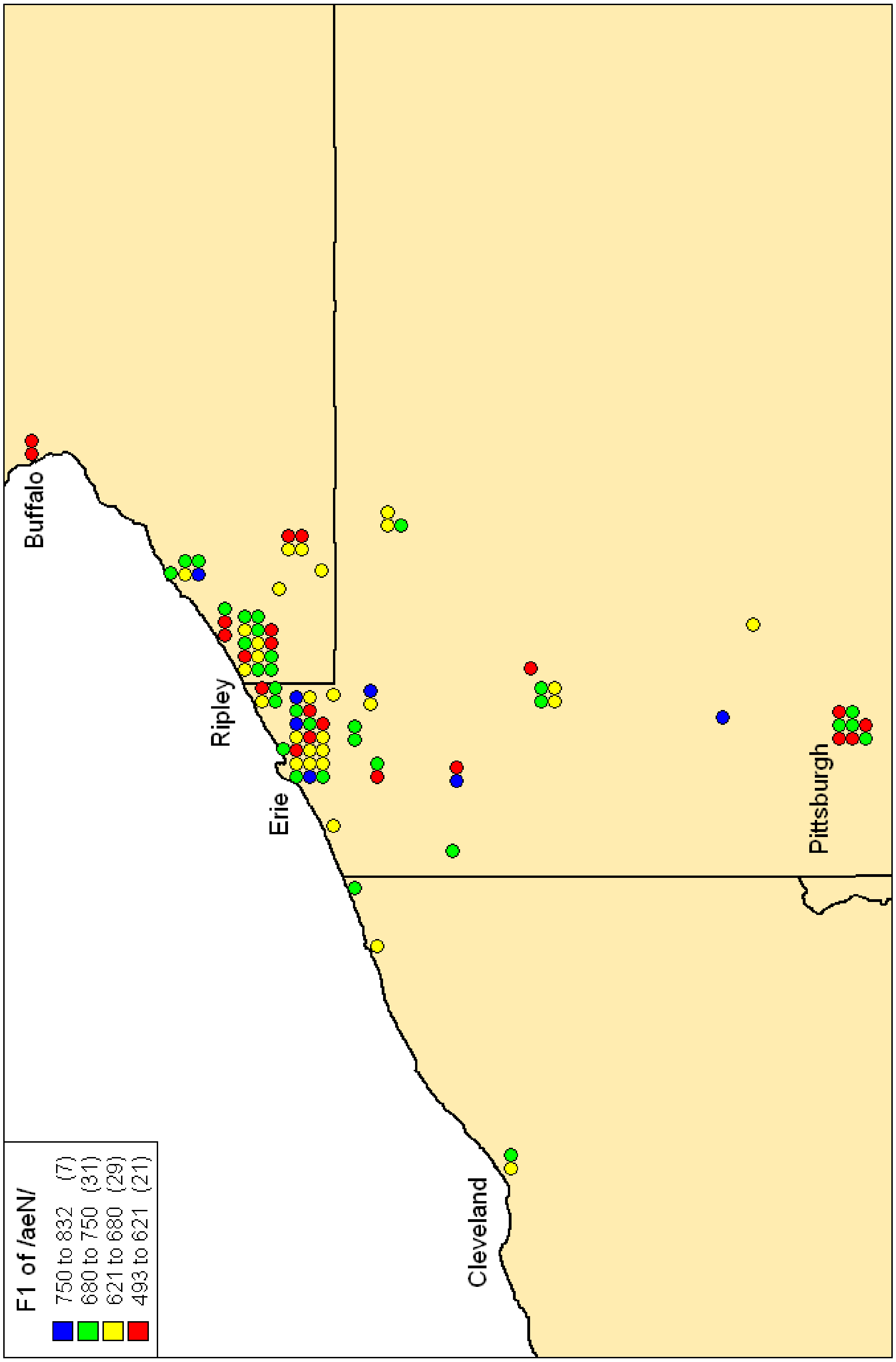

Figure 5.7: Natural break map for F1 of /æN/ (/æ/ before nasals)

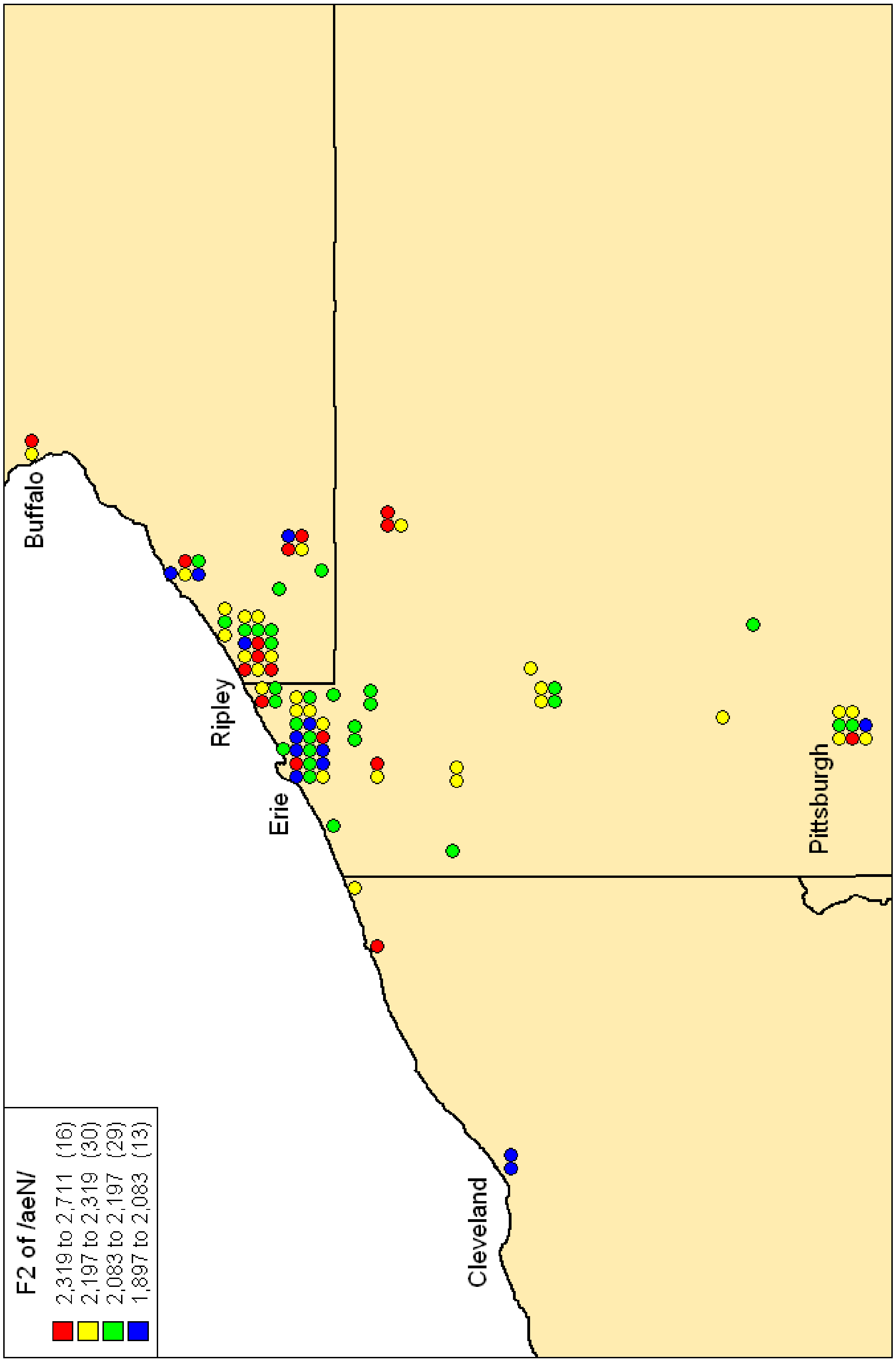

Figure 5.8: Natural break map for F2 of /æN/ (/æ/ before nasals)

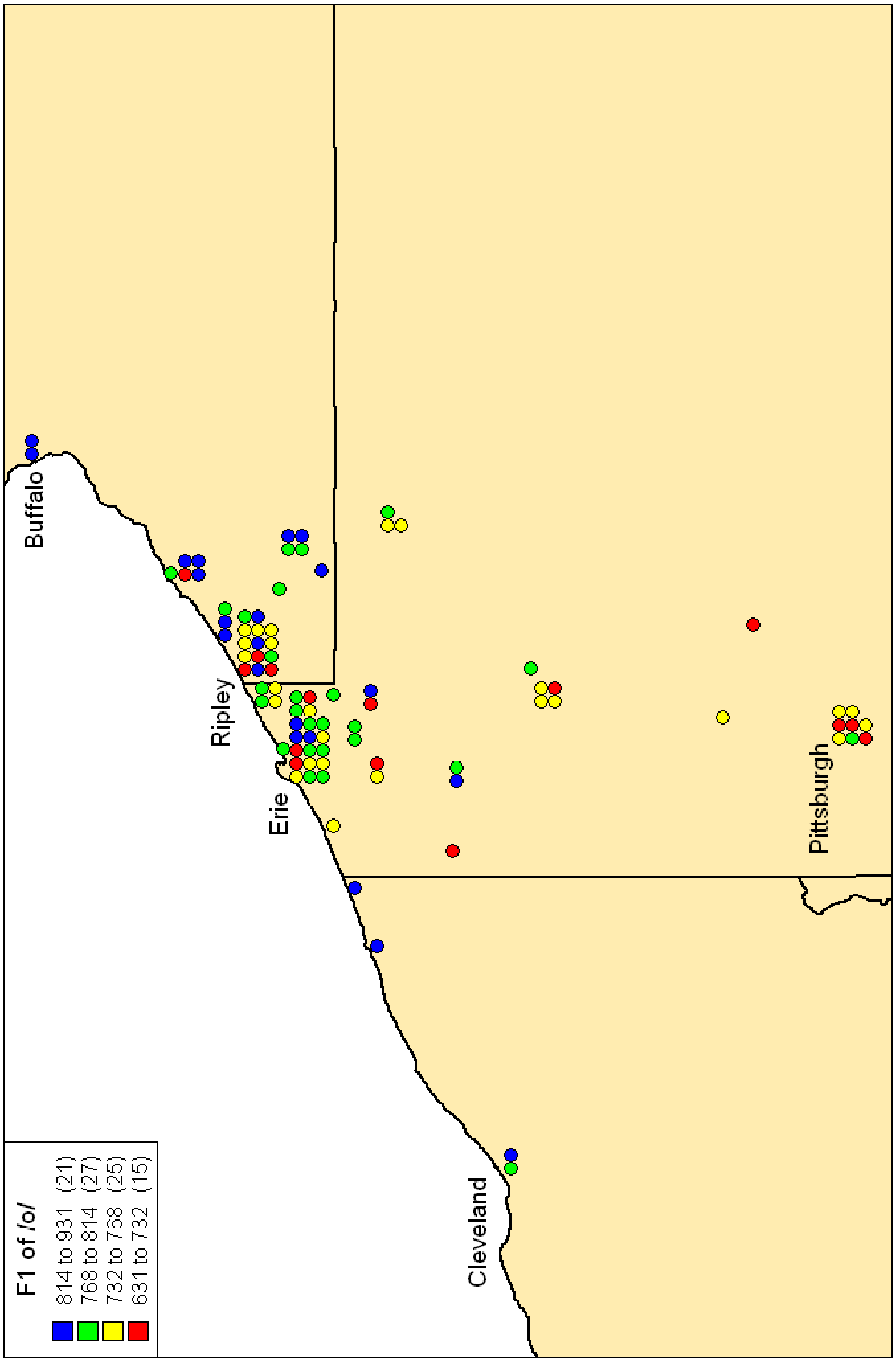

Figure 5.9: Natural break map for F1 of /o/

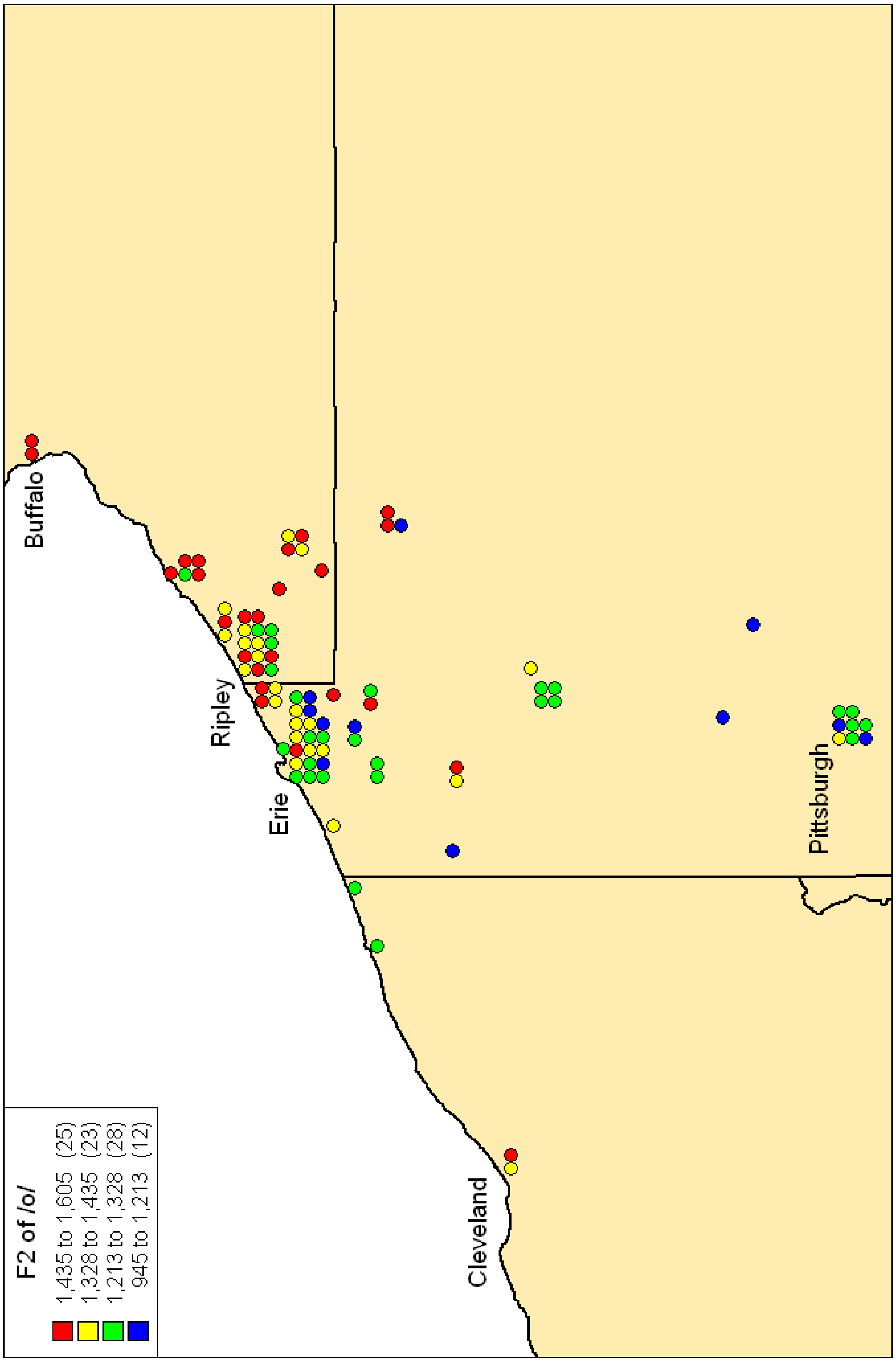

Figure 5.10: Natural break map for F2 of /o/

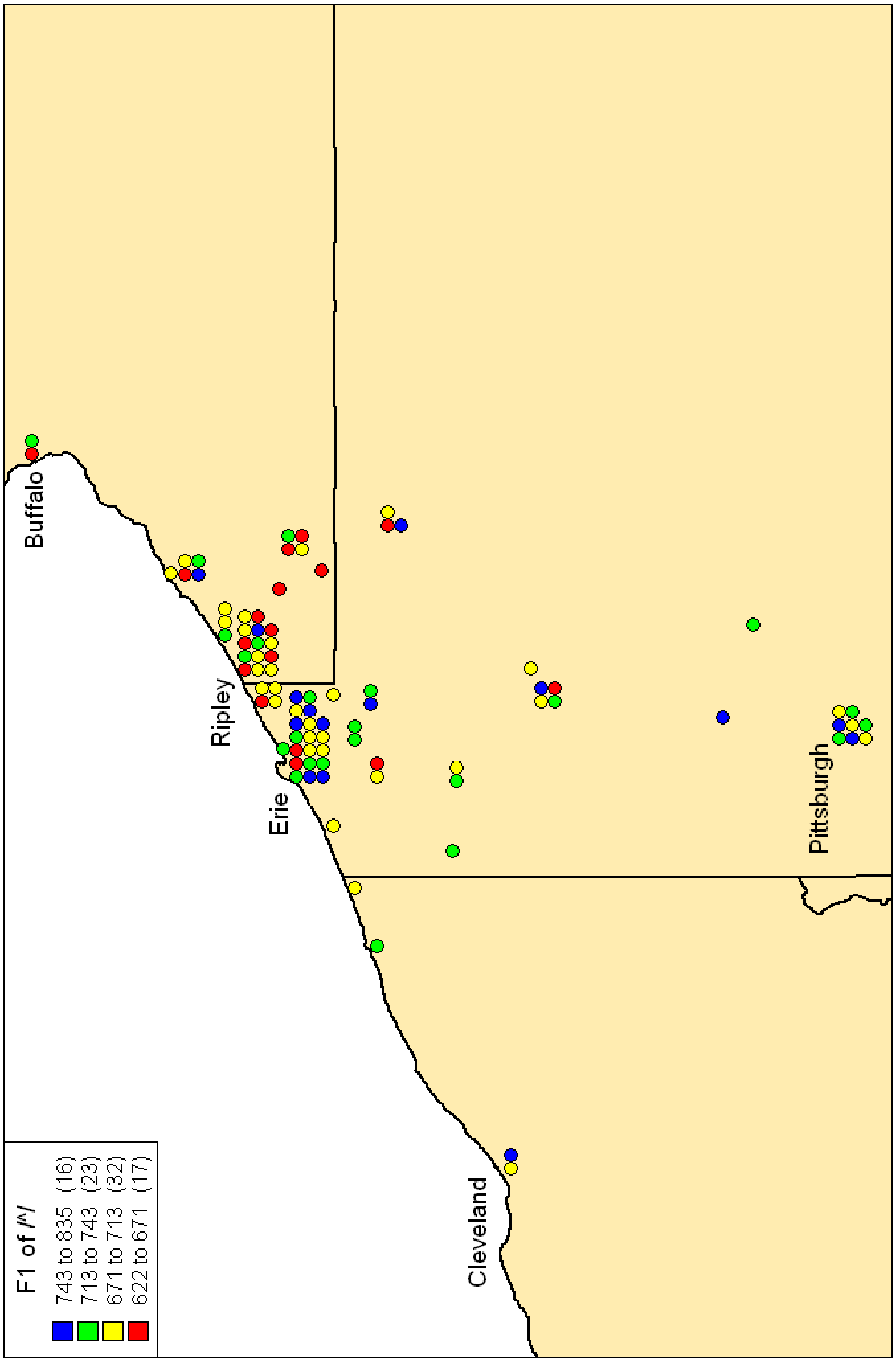

Figure 5.11: Natural break map for F1 of /ʌ/

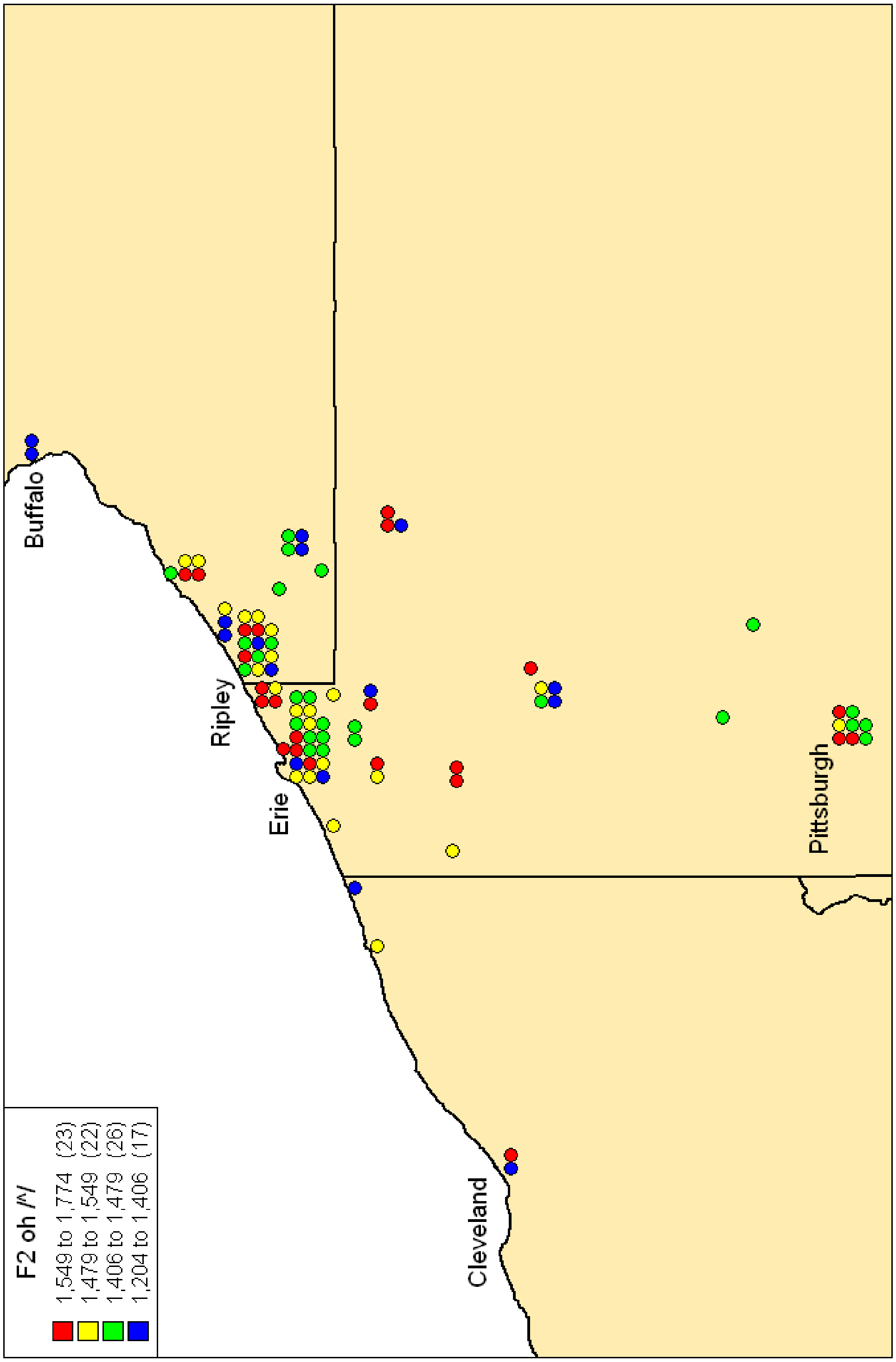

Figure 5.12: Natural break map for F2 of /ʌ/

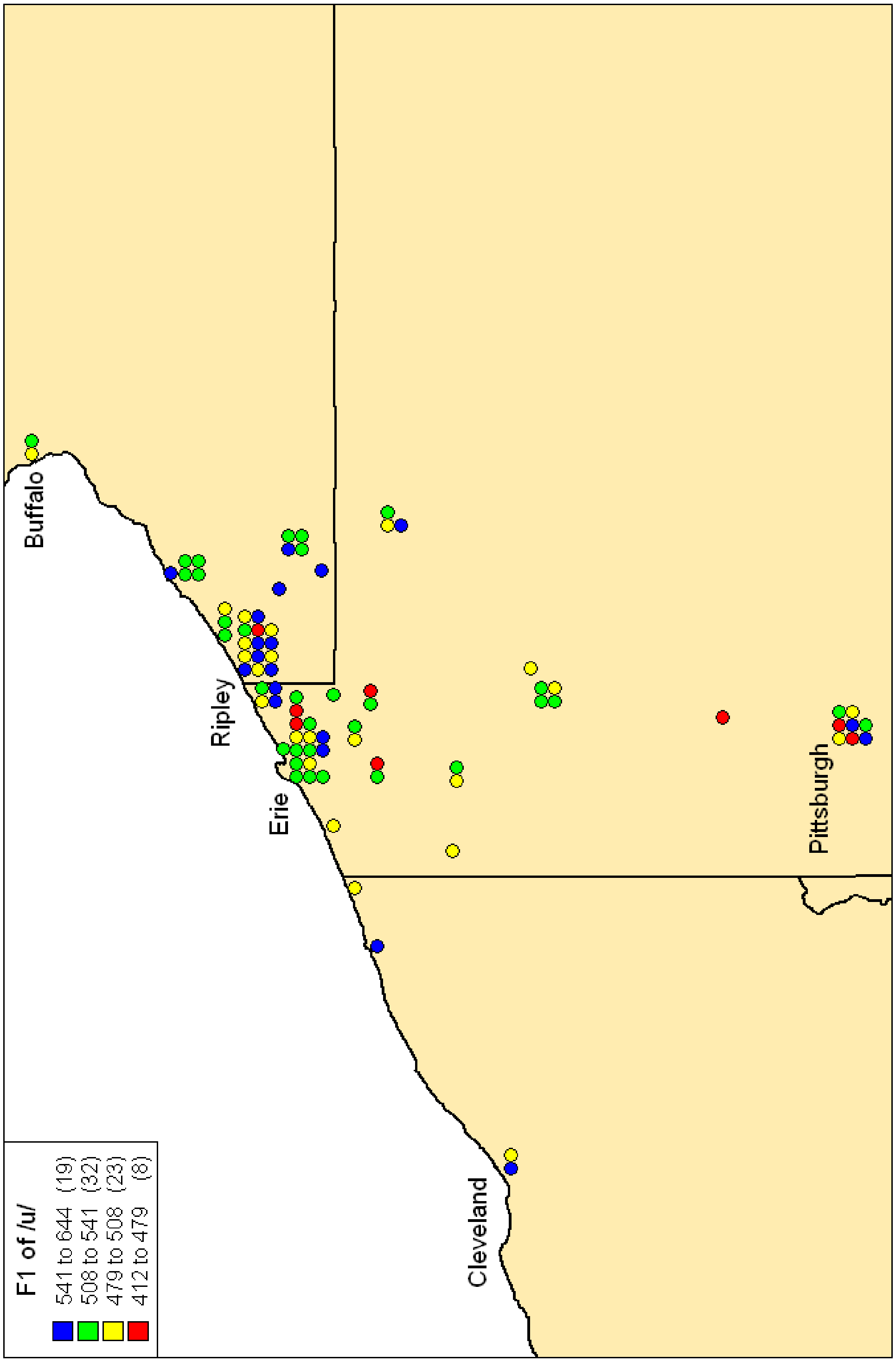

Figure 5.13: Natural break map for F1 of /u/

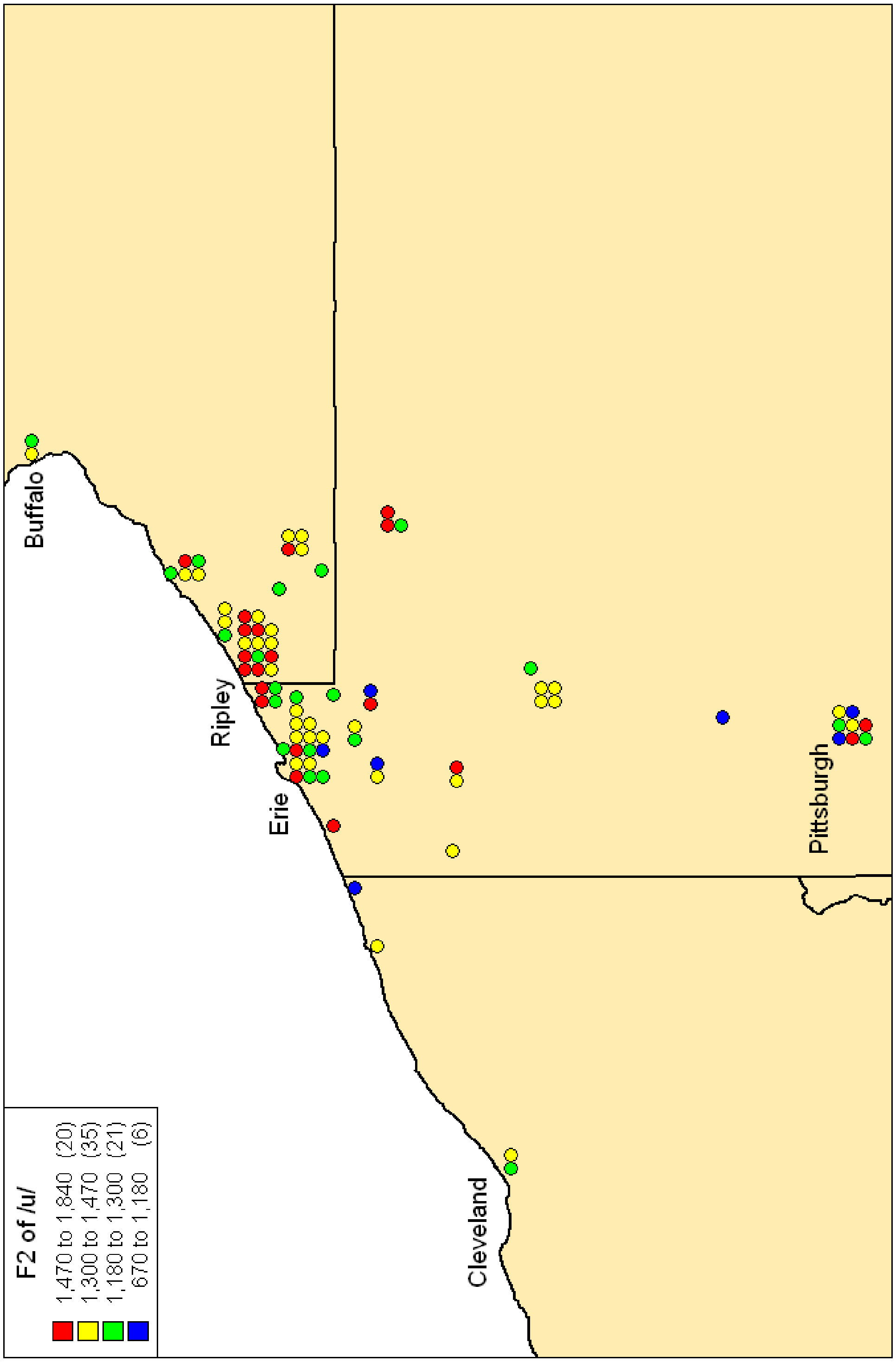

Figure 5.14: Natural break map for F2 of /u/

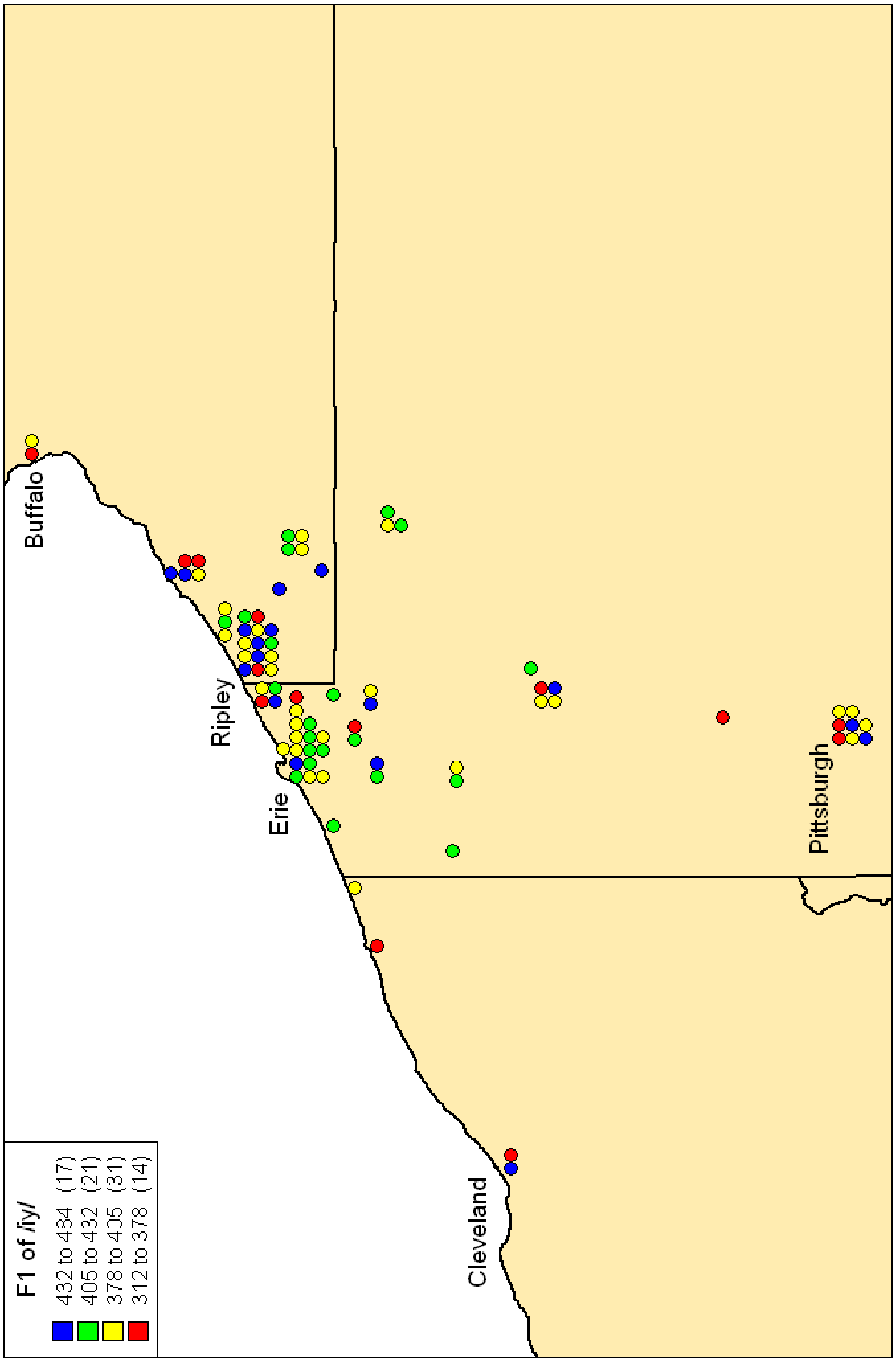

Figure 5.15: Natural break map for F1 of /iy/

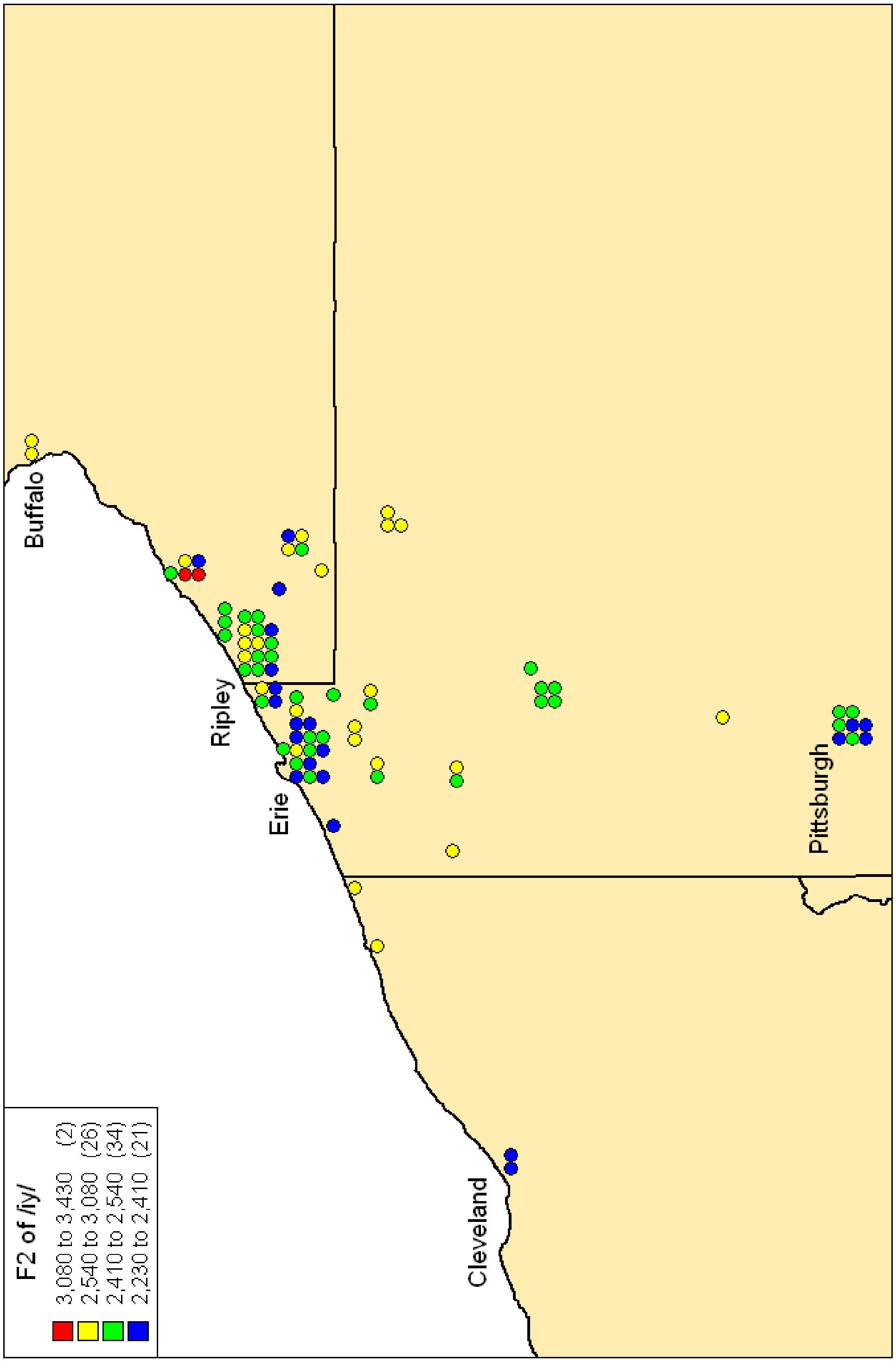

Figure 5.16: Natural break map for F2 of /iy/

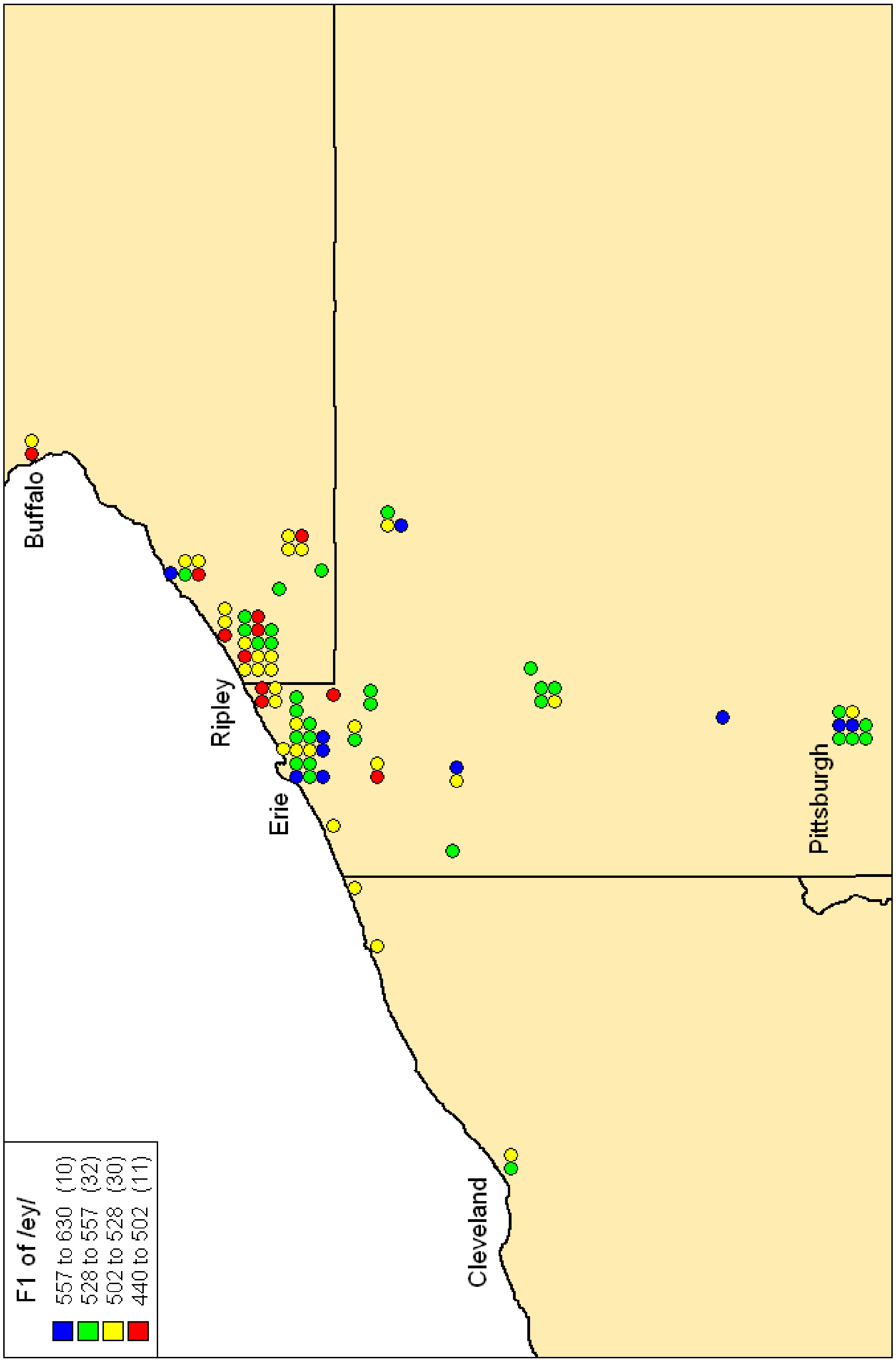

Figure 5.17: Natural break map for F1 of /ey/

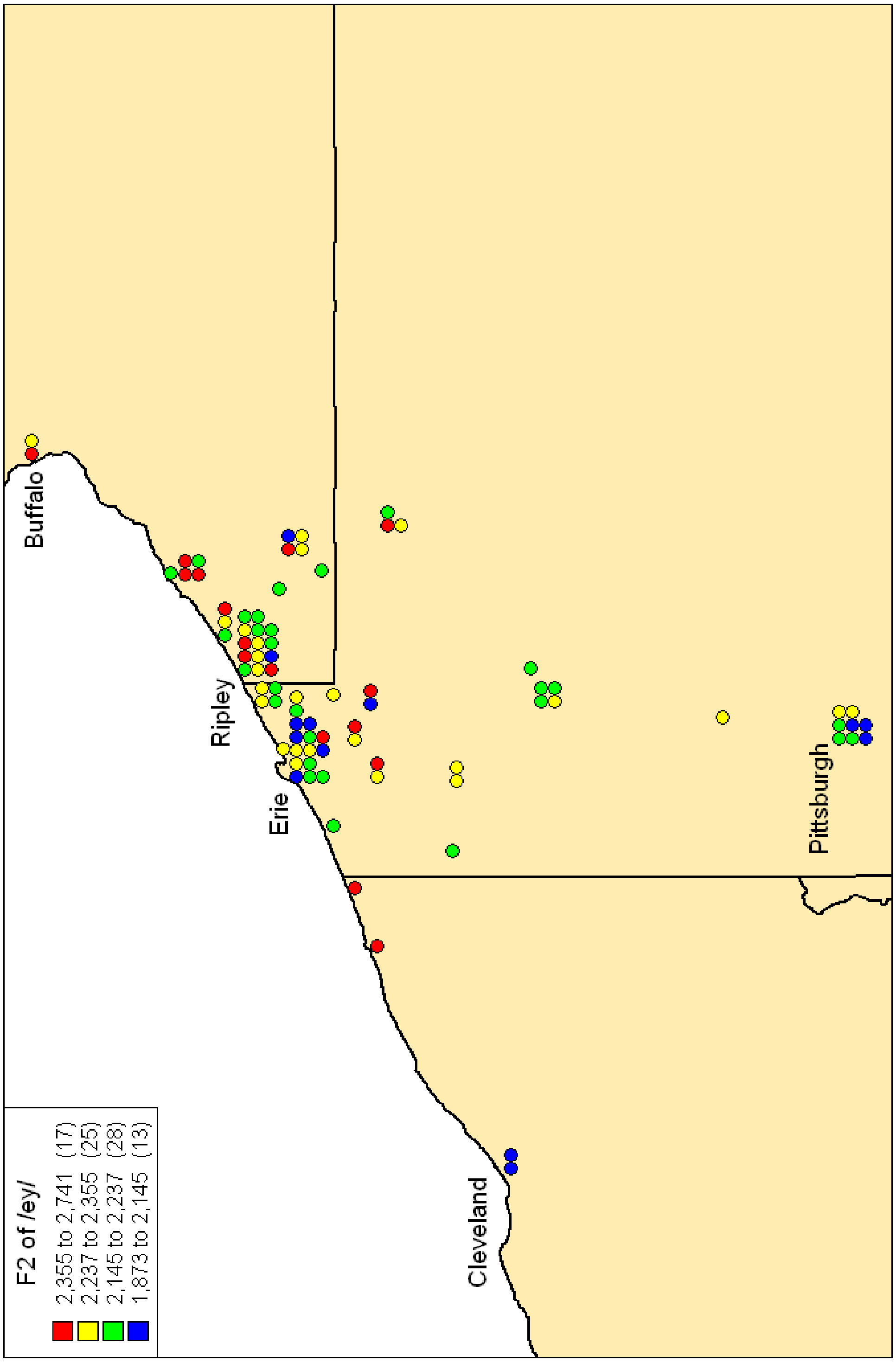

Figure 5.18: Natural break map for F2 of /ey/

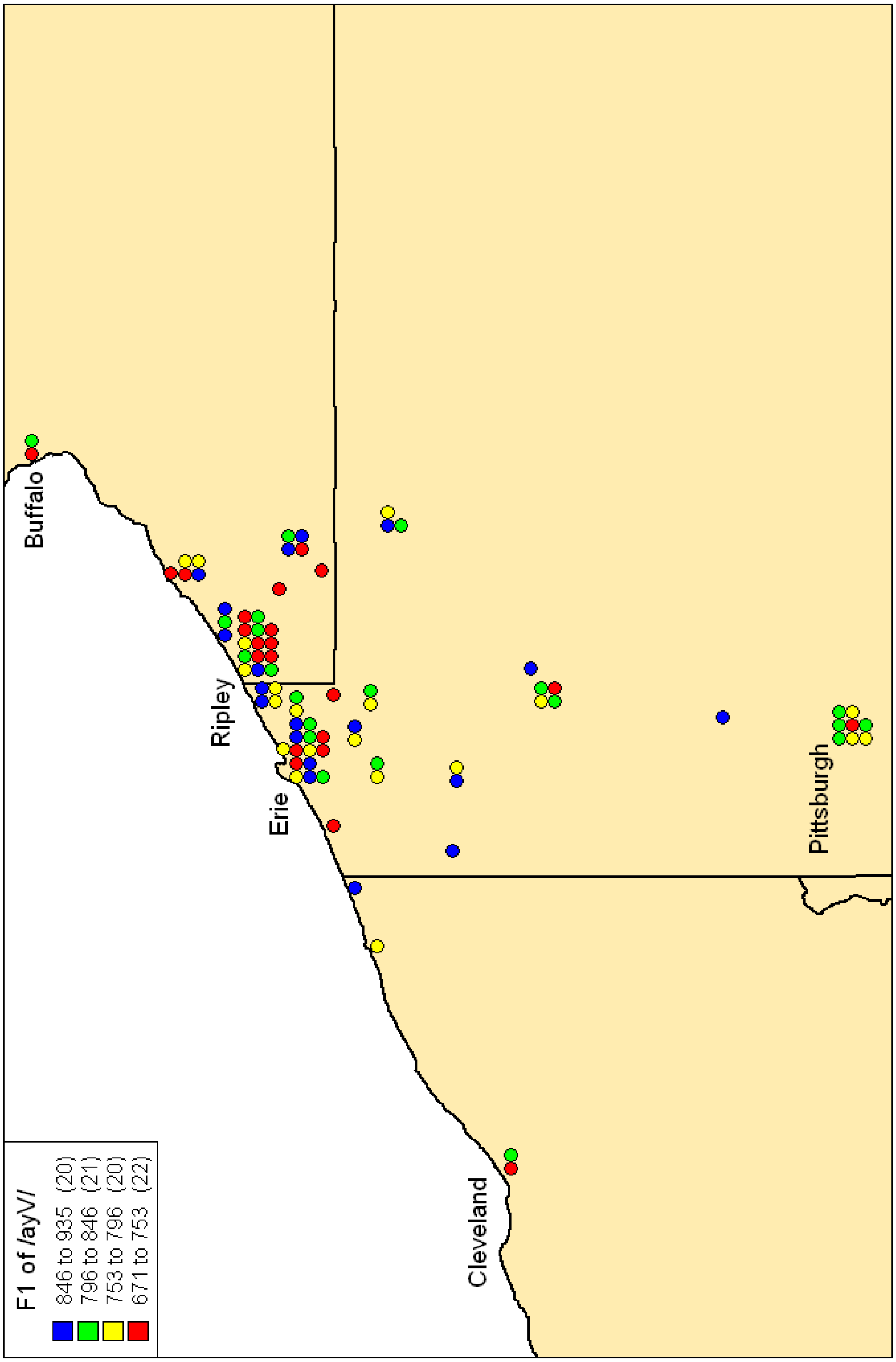

Figure 5.19: Natural break map for F1 of /ayV/ (/ay/ before voiced codas and word-finally)

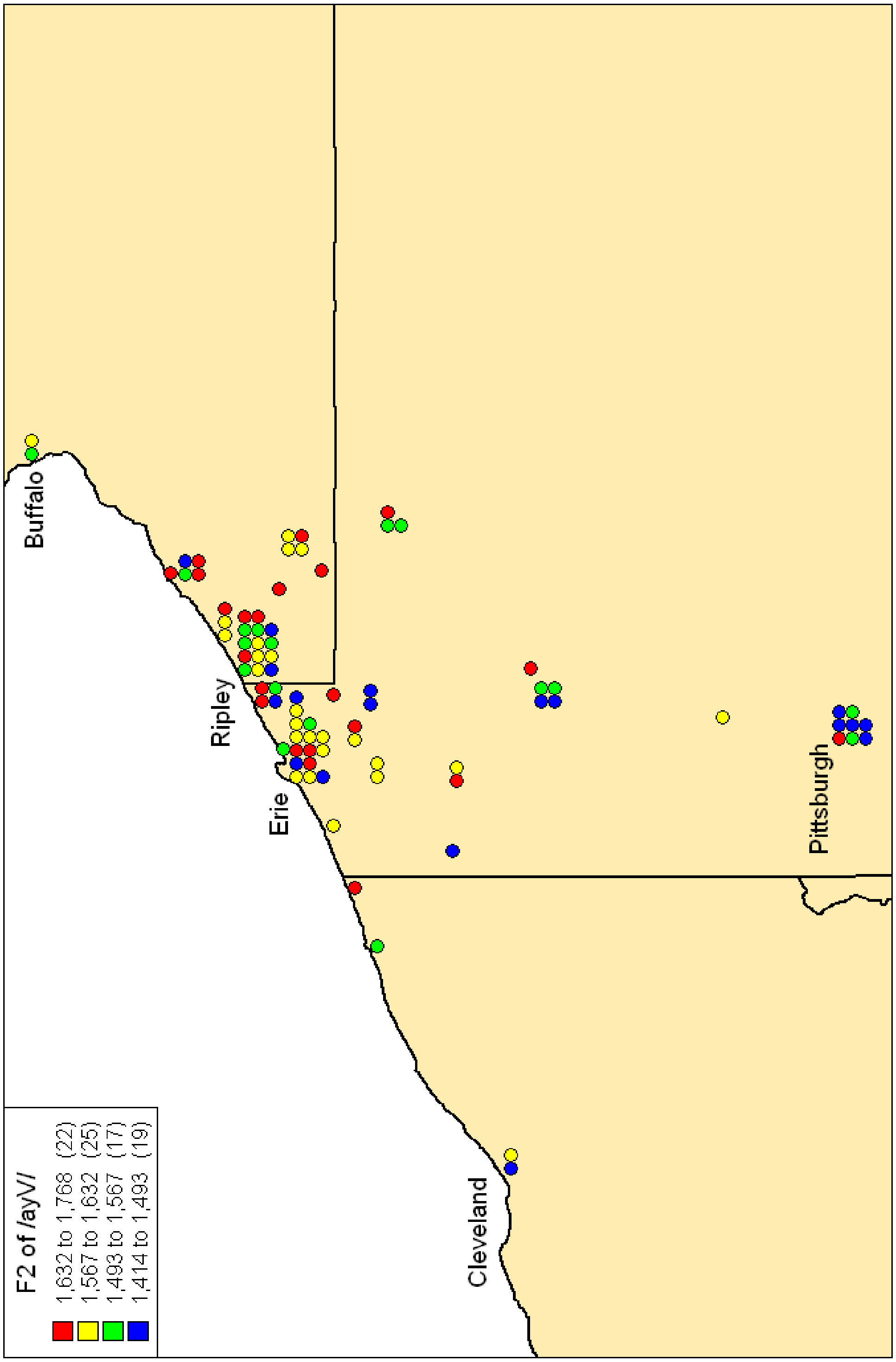

Figure 5.20: Natural break map for F2 of /ayV/ (/ay/ before voiced codas and word-finally)

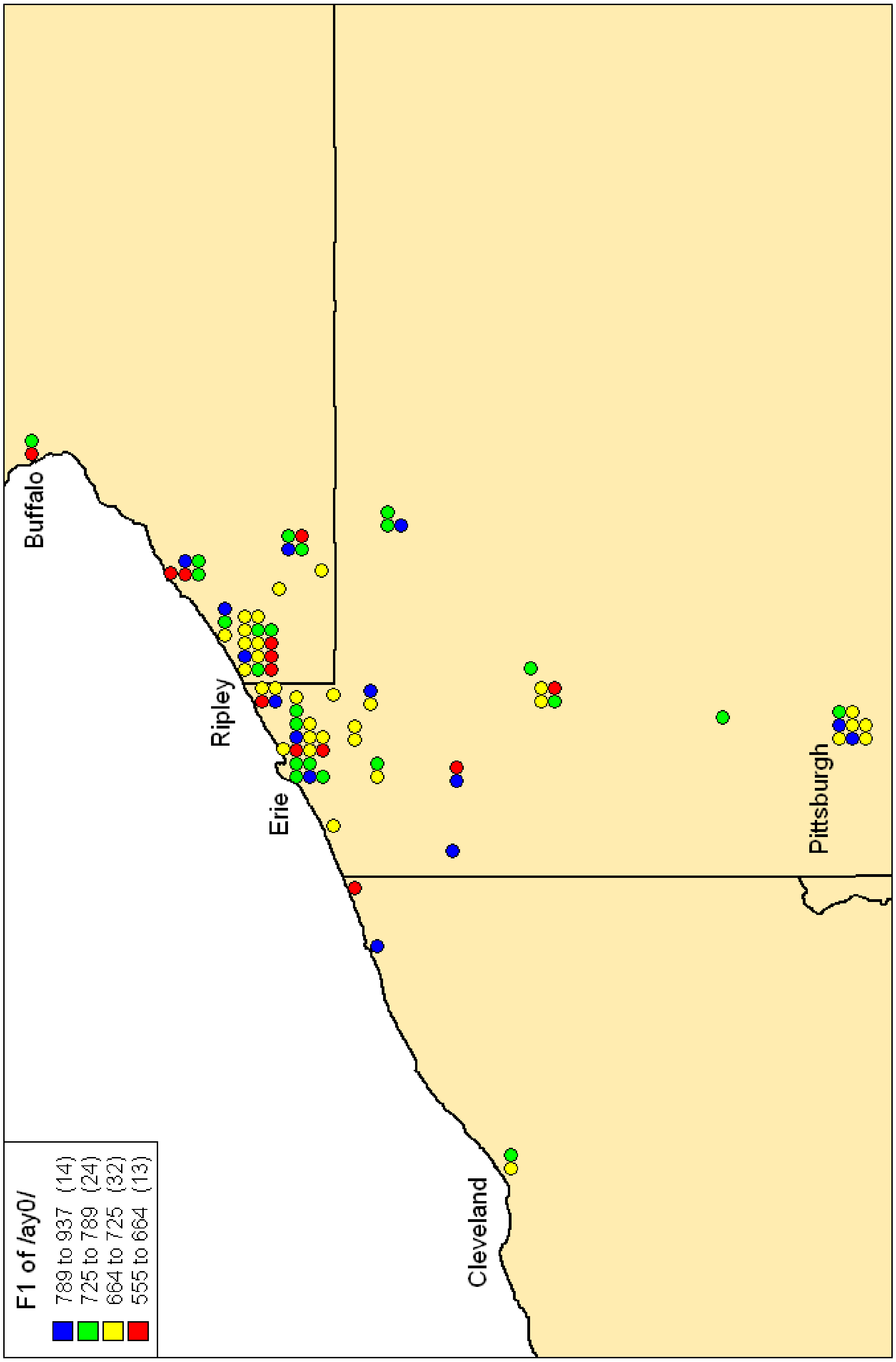

Figure 5.21: Natural break map for F1 of /ay0/ (/ay/ before voiceless codas)

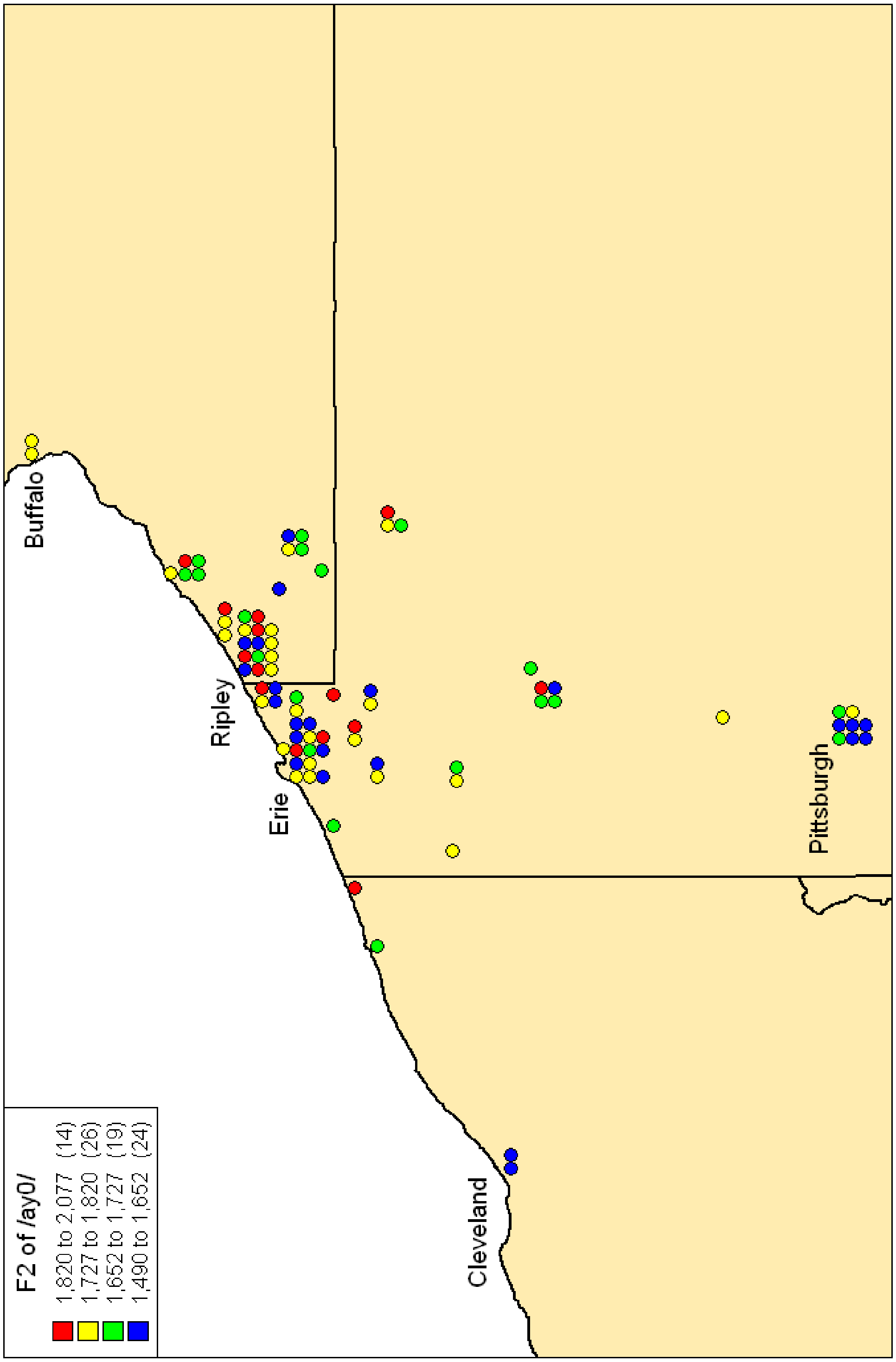

Figure 5.22: Natural break map for F2 of /ay0/ (/ay/ before voiceless codas)

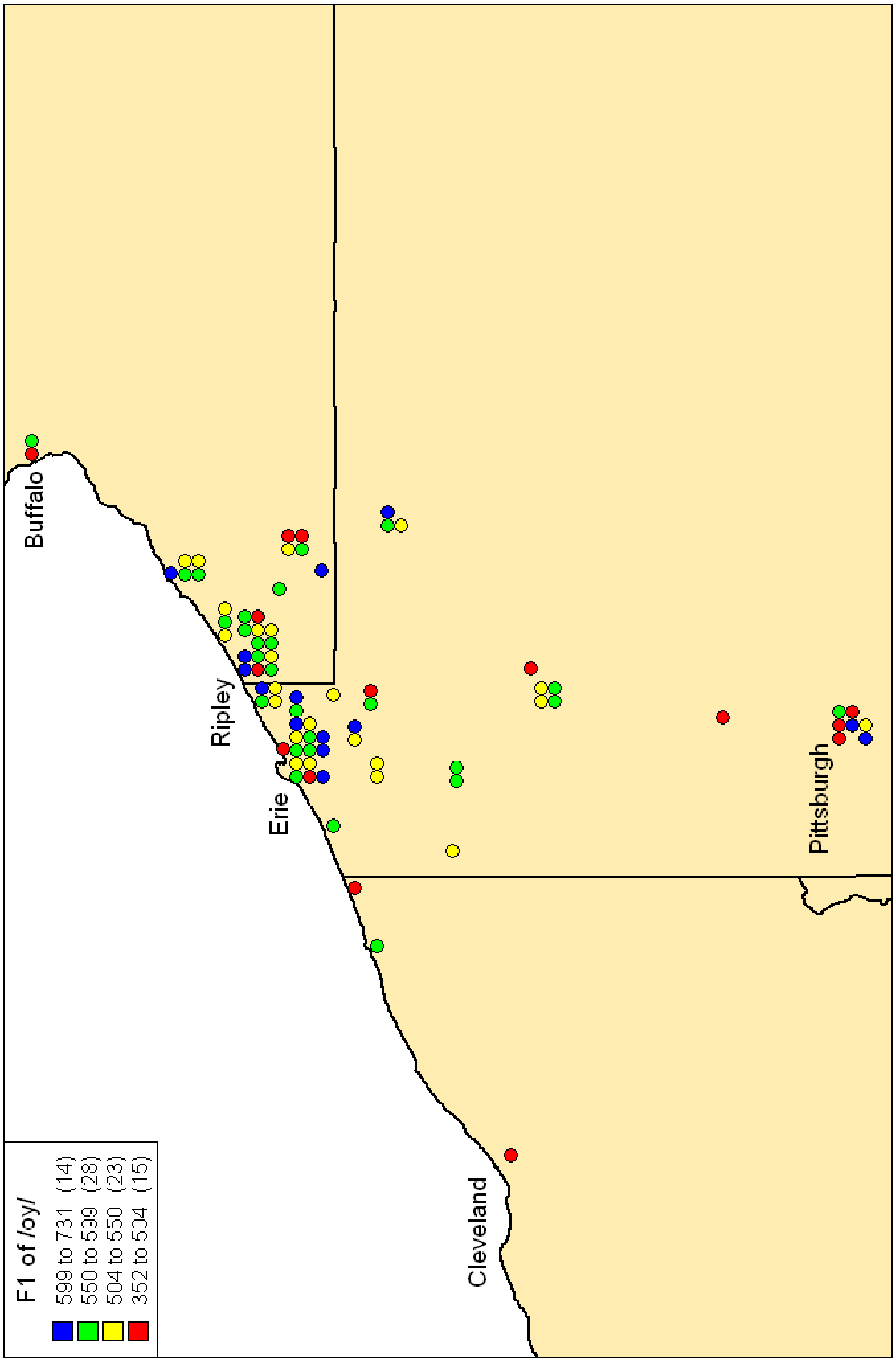

Figure 5.23: Natural break map for F1 of /oy/

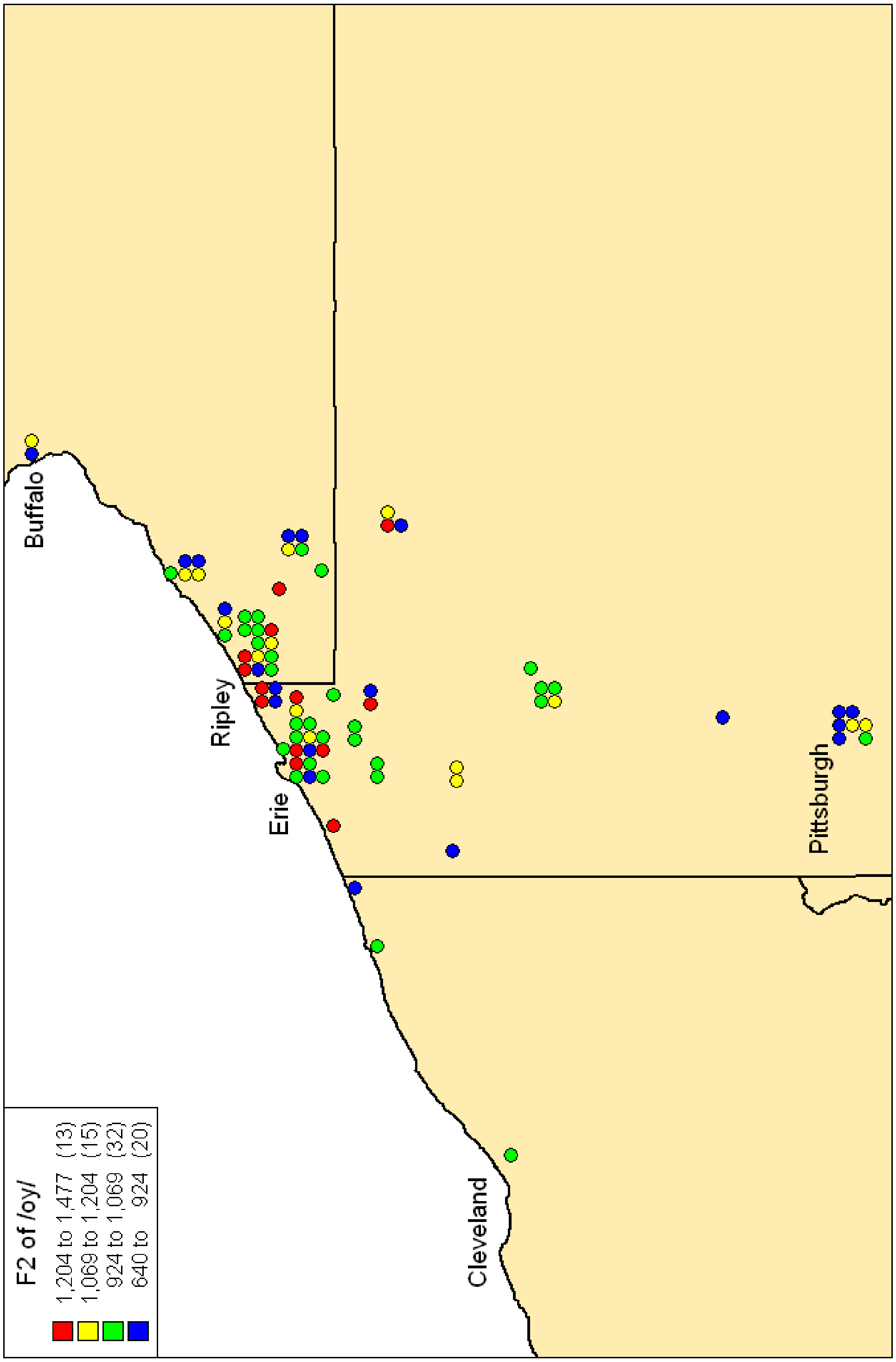

Figure 5.24: Natural break map for F2 of /oy/

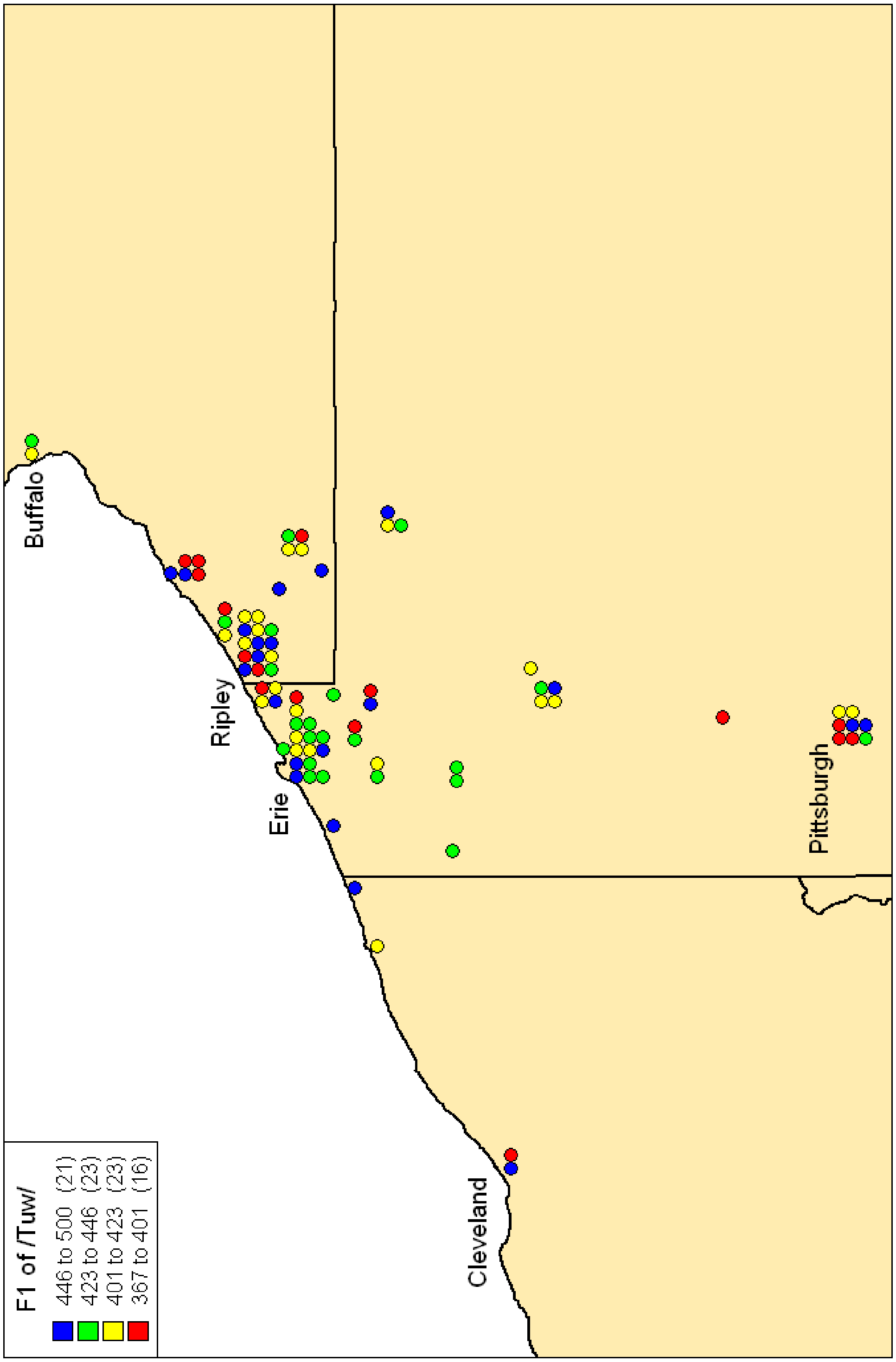

Figure 5.25: Natural break map for F1 of /Tuw/ (/uw/ after coronal onsets)

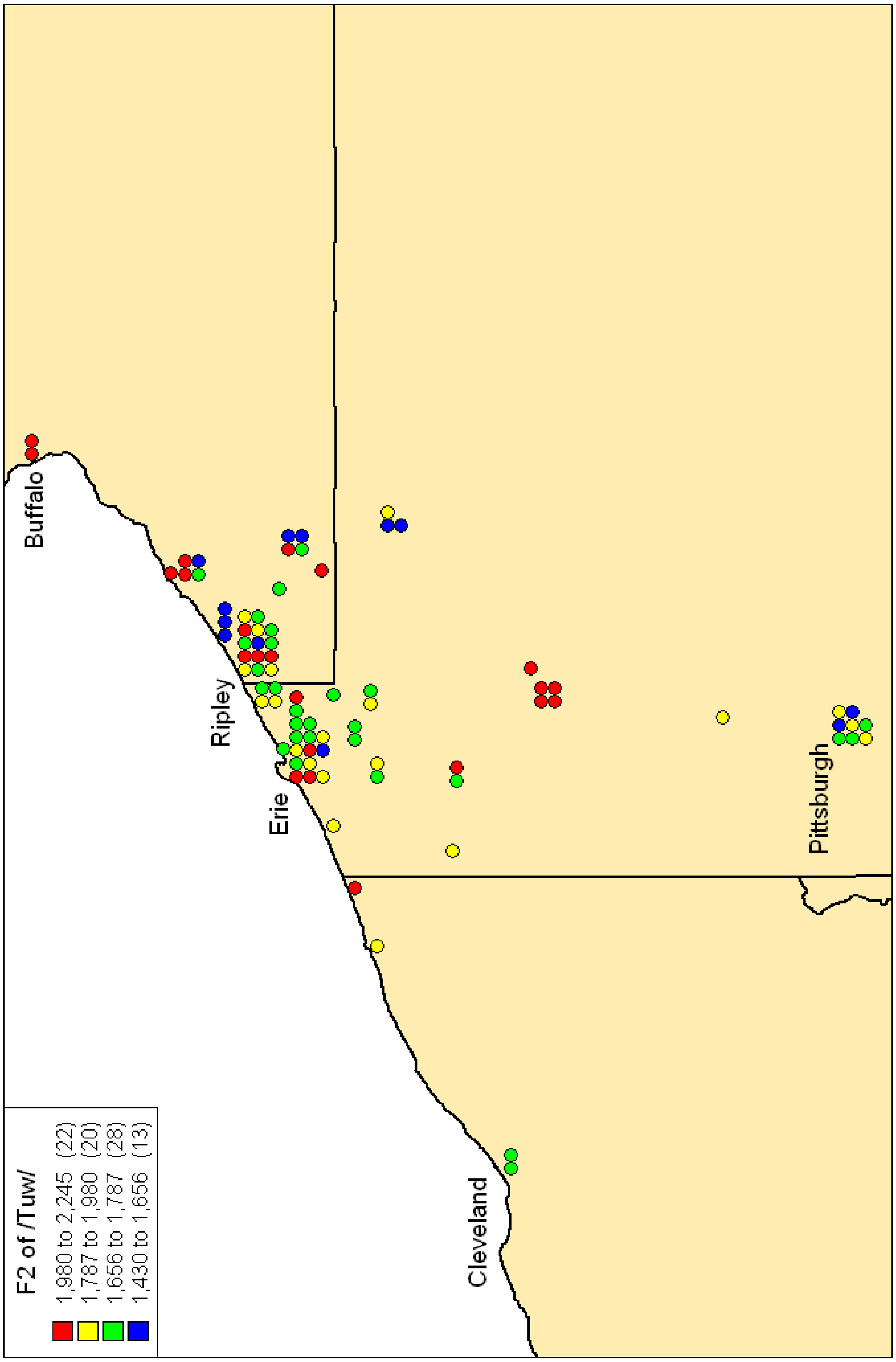

Figure 5.26: Natural break map for F2 of /Tuw/ (/uw/ after coronal onsets)

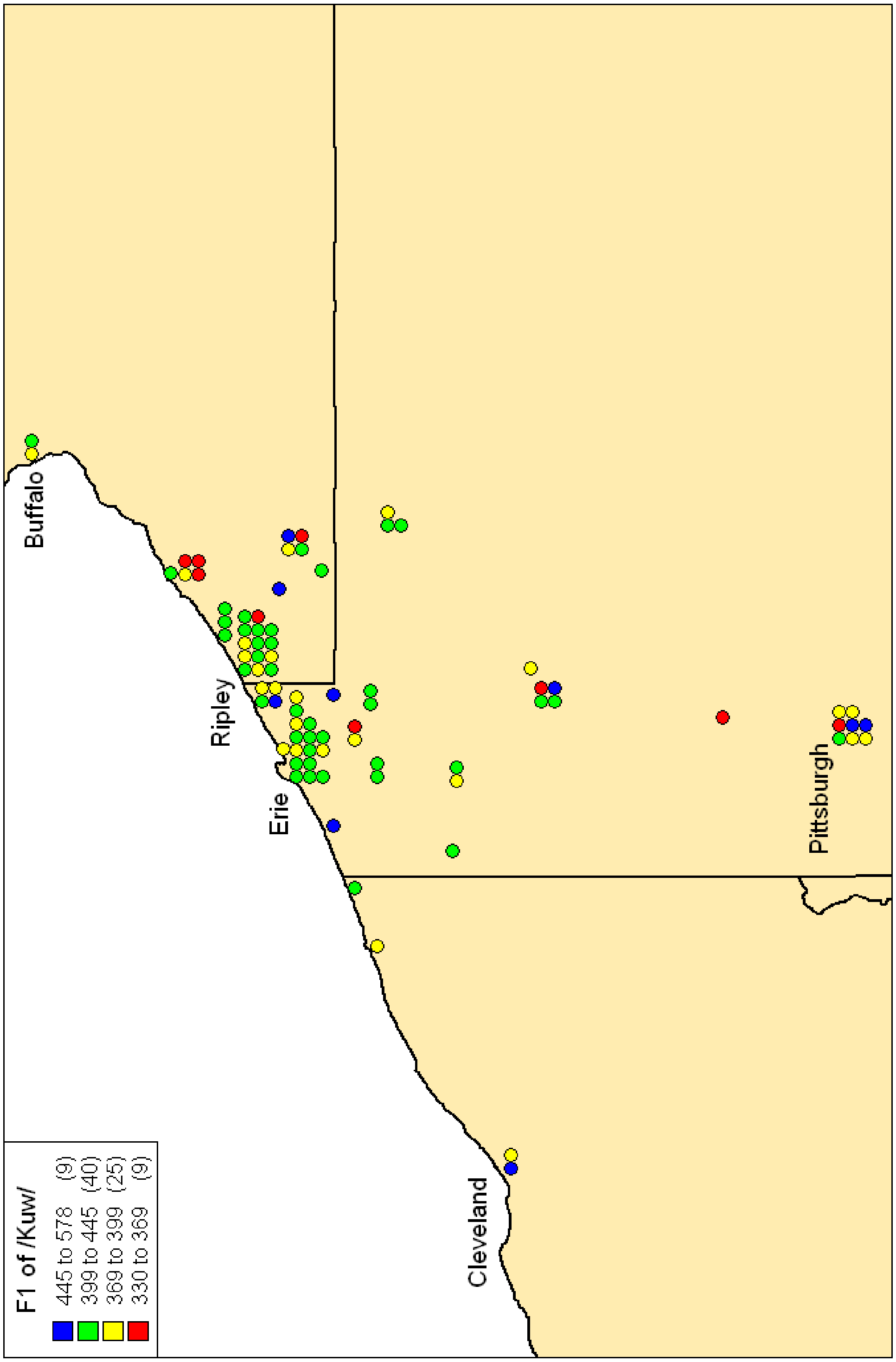

Figure 5.27: Natural break map for F1 of /Kuw/ (/uw/ after non-coronal onsets)

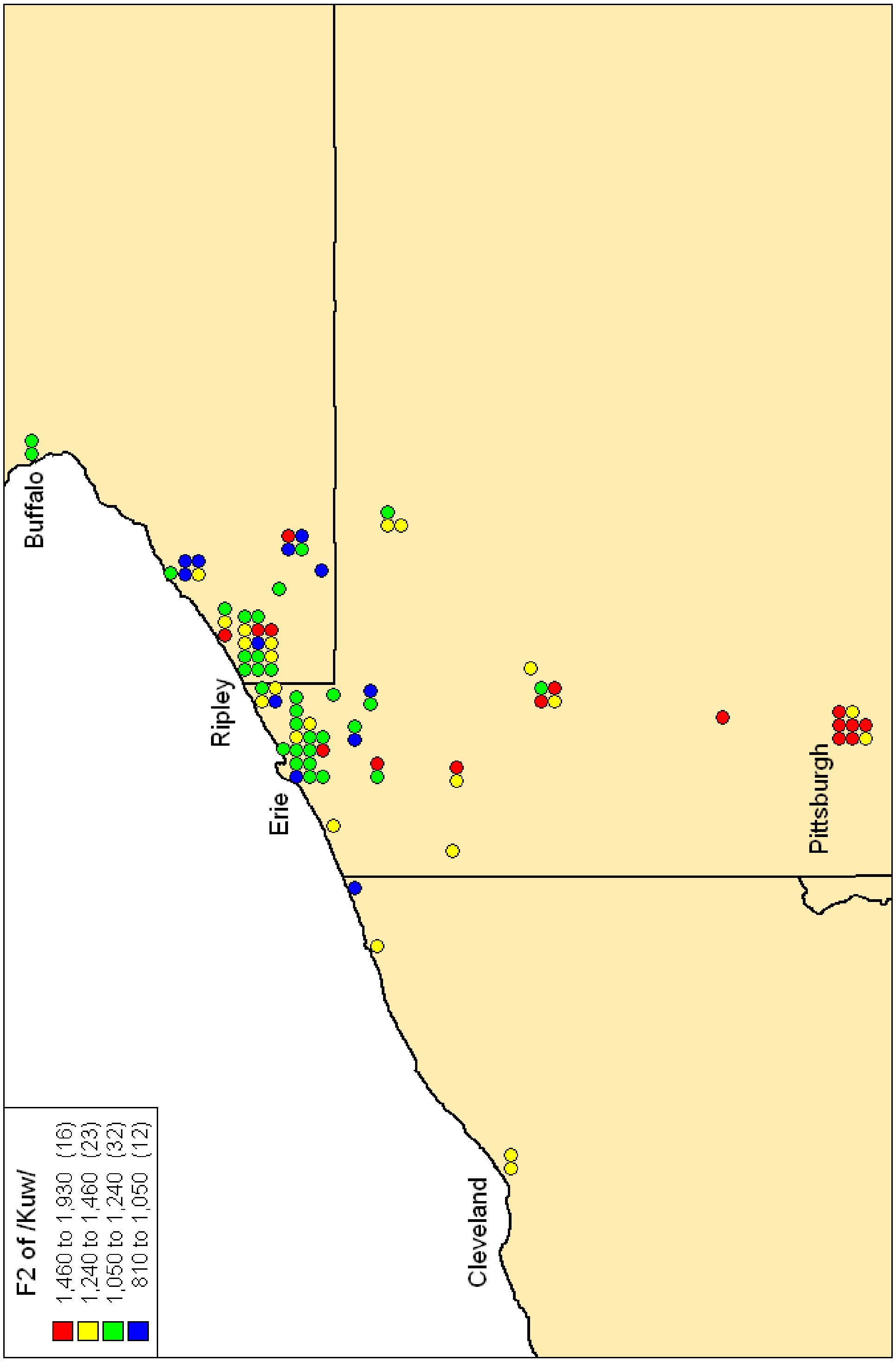

Figure 5.28: Natural break map for F2 of /Kuw/ (/uw/ after non-coronal onsets)

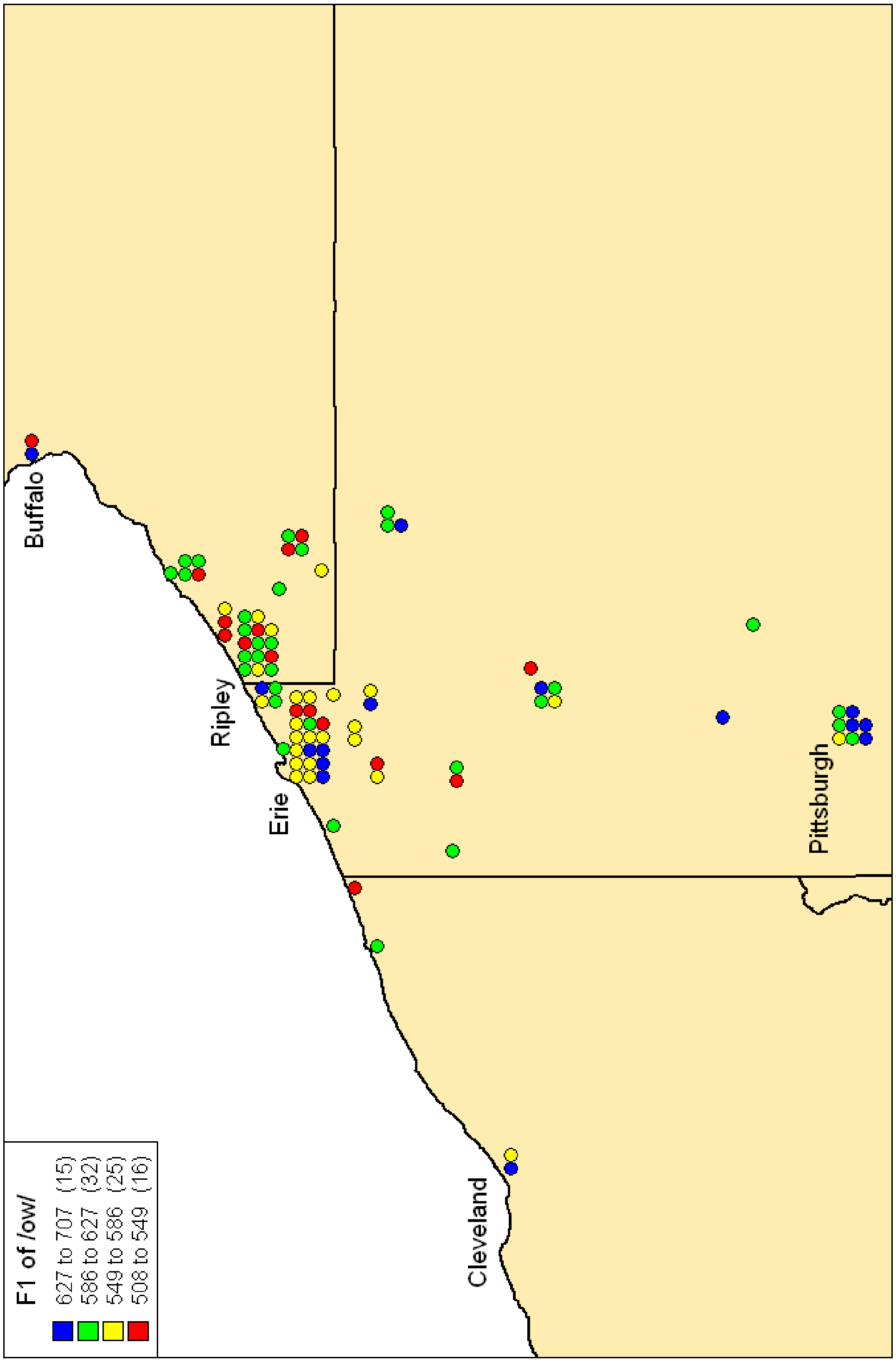

Figure 5.29: Natural break map for F1 of /ow/

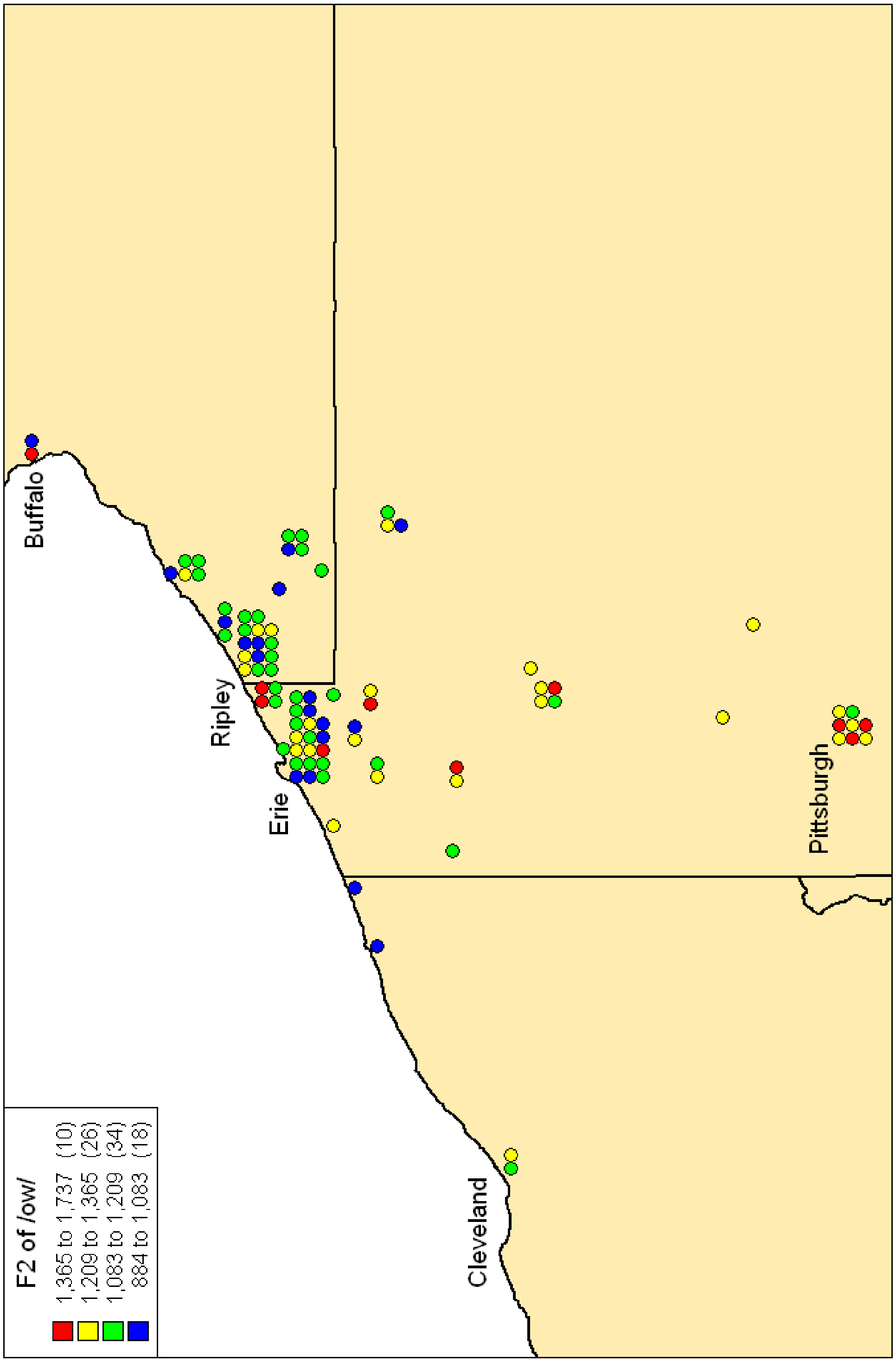

Figure 5.30: Natural break map for F2 of /ow/

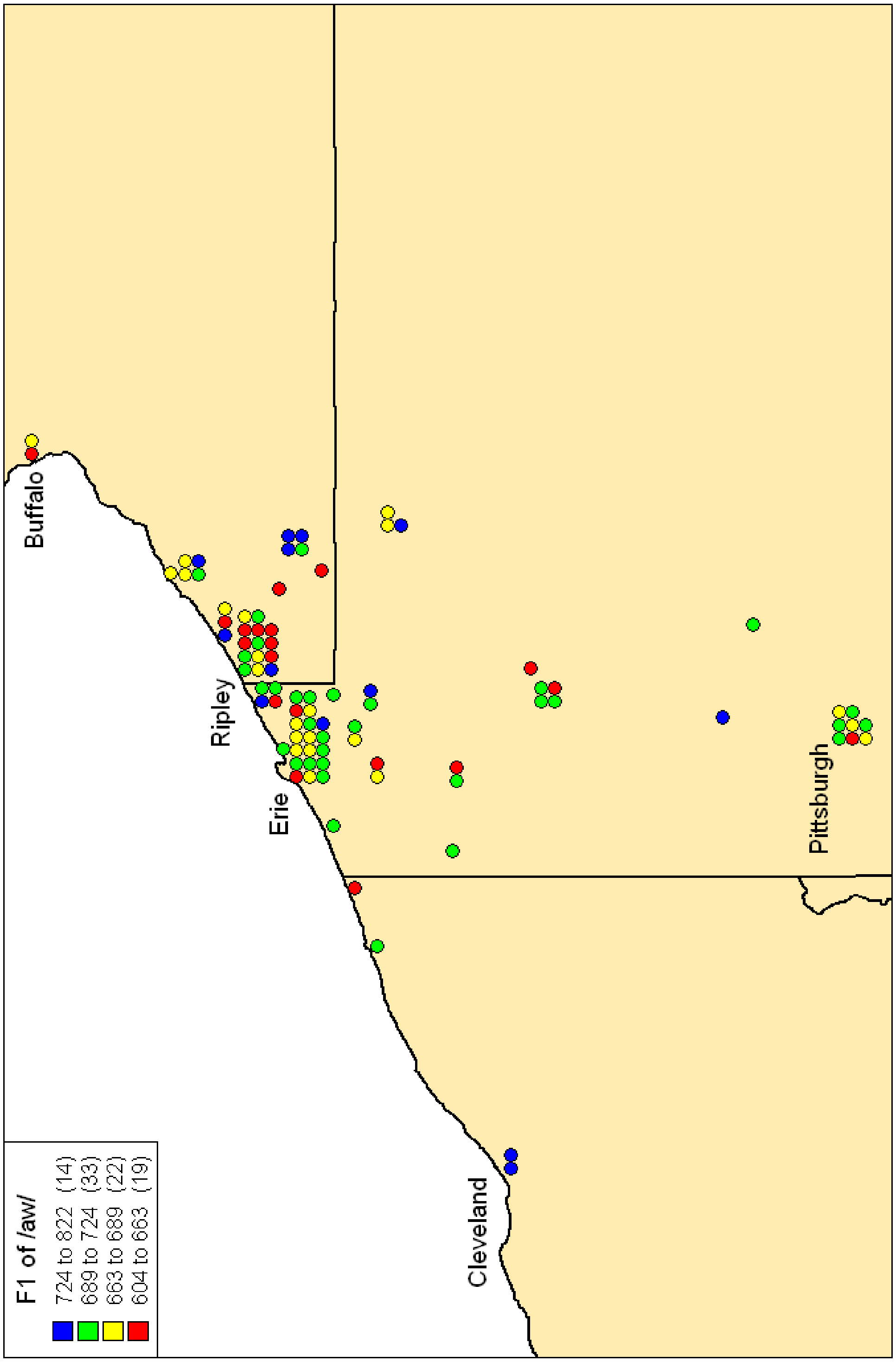

Figure 5.31: Natural break map for F1 of /aw/

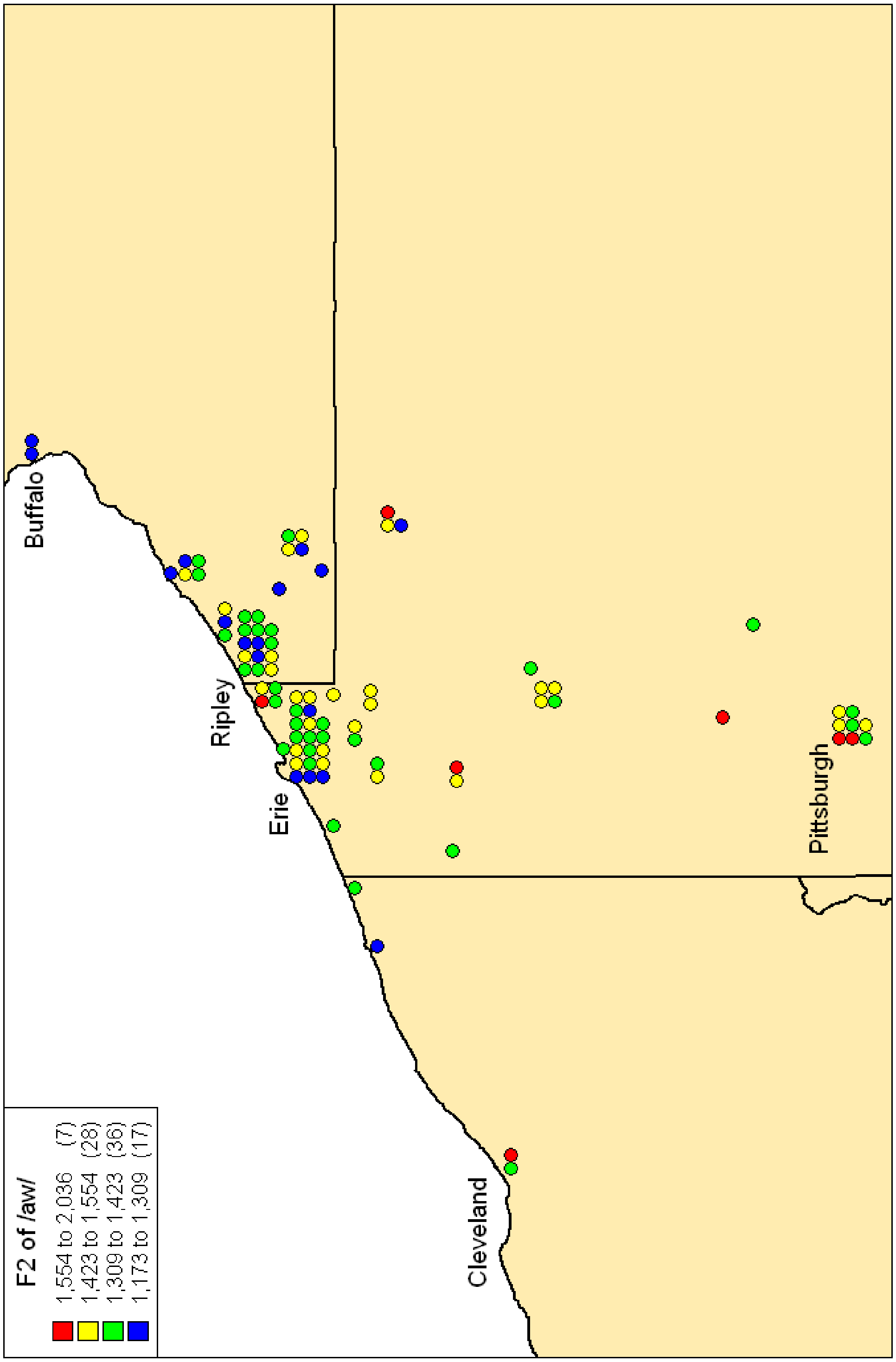

Figure 5.32: Natural break map for F2 of /aw/

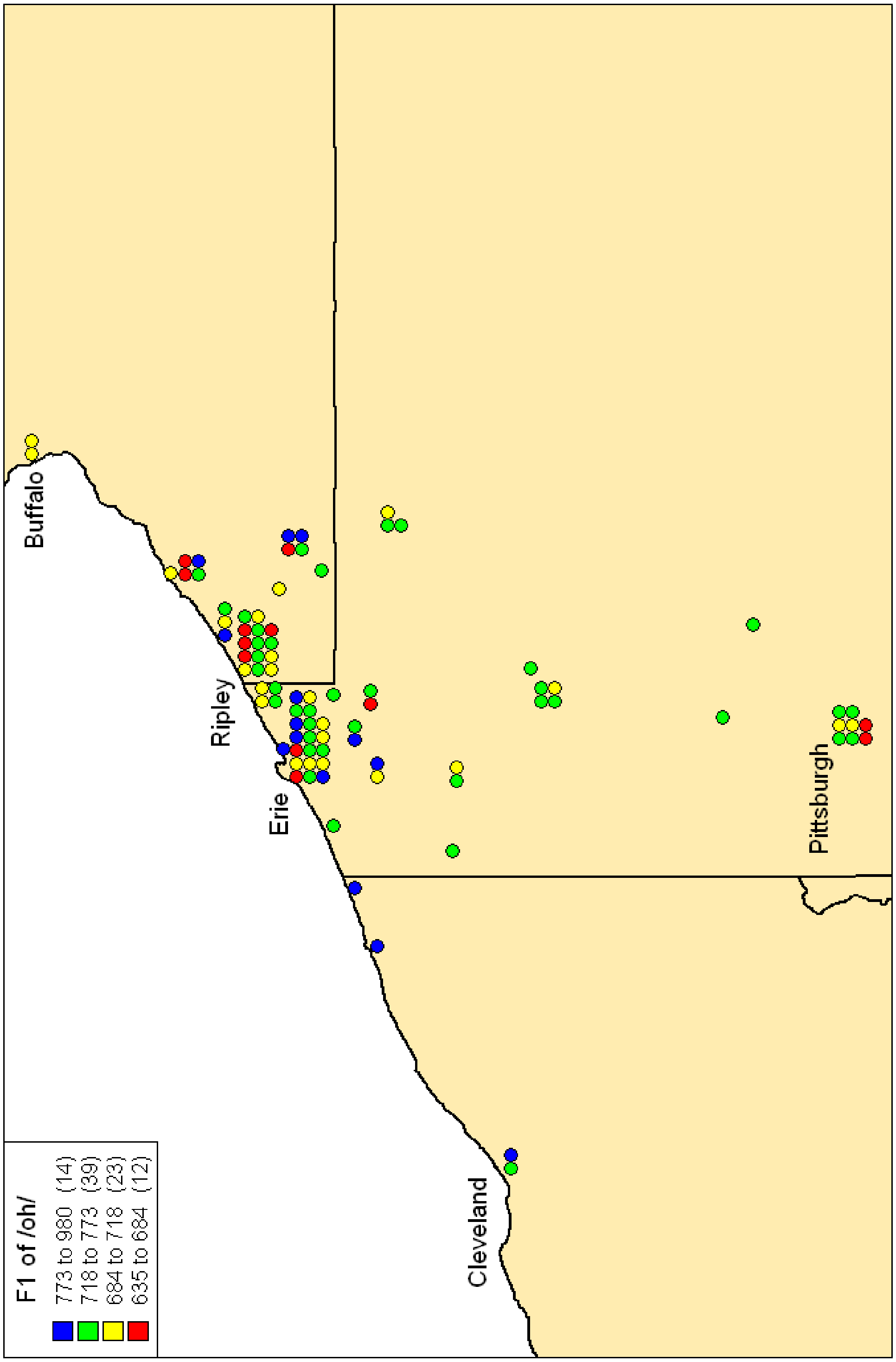

Figure 5.33: Natural break map for F1 of /oh/

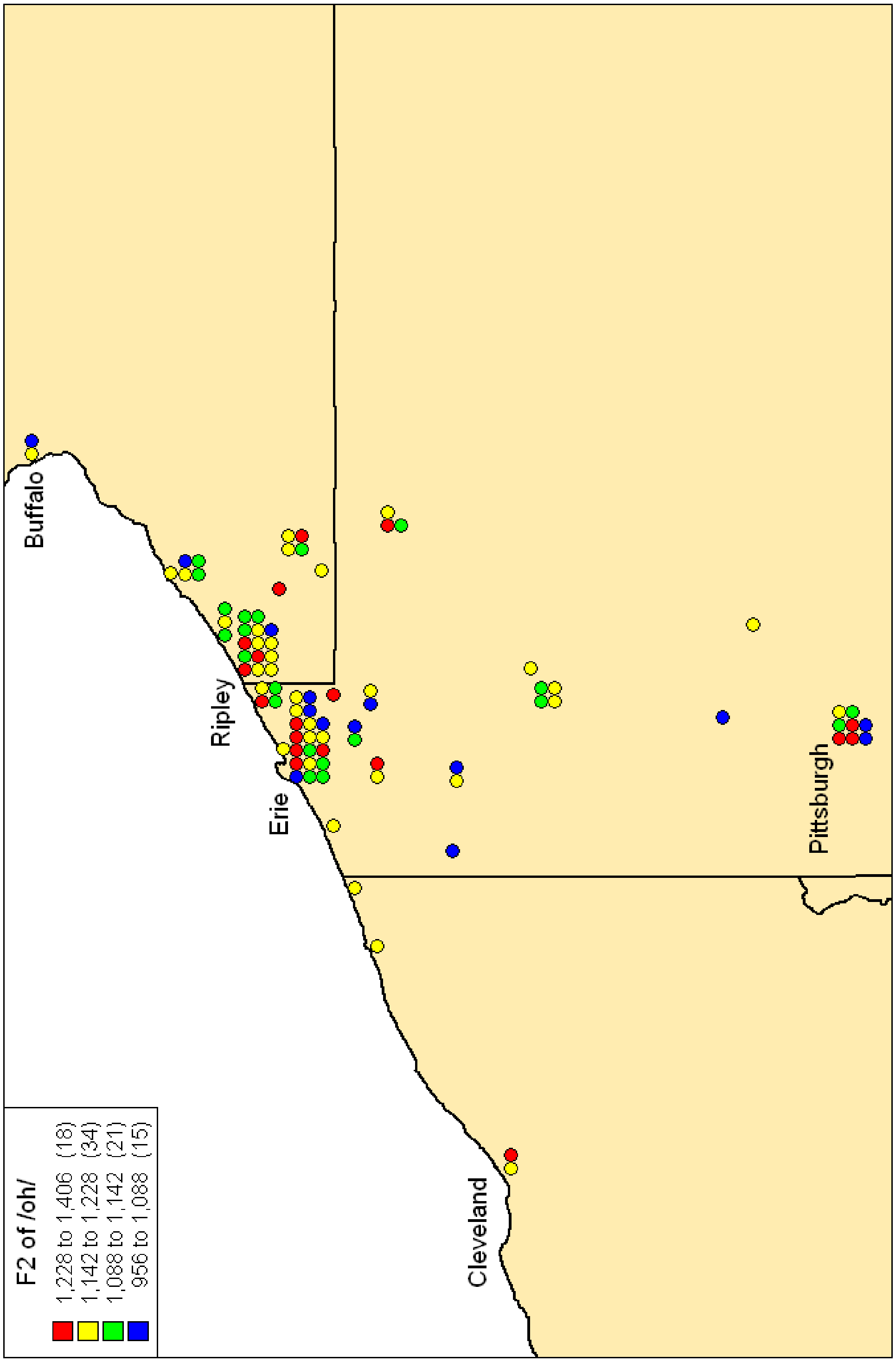

Figure 5.34: Natural break map for F2 of /oh/

Chapter 6

The Merger of /o/ and /oh/

6.1 Introduction

This chapter discusses the status of /o/ and /oh/ for all of the speakers included in the corpus for this dissertation. These two vowels are of utmost importance for the dialectological status of Erie; the spread of their merger to Erie is the clearest diagnostic for no longer including the city in the Northern dialect region. Furthermore, the merger of /o/ and /oh/ in Erie clearly aligns it with the area of the Midland directly to the south and centered around Pittsburgh, where the two vowels are also solidly merged. In this chapter, the apparent-time evidence from both the interviews I conducted and the archival sources I examined is analyzed to determine when and how the merger spread through Erie and beyond into Chautauqua County, NY. The analysis draws upon both acoustic data and experimental data (in the form of minimial pair tests), in an attempt to characterize each speaker in the corpus as merged, unmerged, or transitional.

This chapter is organized as follows: first, Section 6.2 provides an overview of the two vowel phonemes under consideration, /o/ and /oh/. Next, Section 6.3 describes the status of /o/ and /oh/ in the two regions neighboring Erie. Then, Section 6.4 discusses the question of how to determine whether a given speaker has the merger of /o/ and /oh/ or not. The information about the status of /o/ and /oh/ in Erie available from previously published sources is reviewed in Section 6.5. Then, Sections 6.6, 6.7, and 6.8 present the results from my fieldwork and archival research. Section 6.9 takes a closer examination at the data from Ripley, a town in Chautauqua County, NY, which acquired the merger in the middle of the 20th century. Finally, Section 6.10 reviews all of the available evidence for the chronology of the spread of the merger.

6.2 Overview of /o/ and /oh/

The short-*o* vowel is represented here by the symbol /o/, following the notation in the ANAE, and it corresponds to the LOT vowel class in Wells (1982).[1] It is descended primarily from short **o** in Middle English, and occurs in nearly all segmental environments. Some examples of words with /o/ include *lock*, *pot*, *god*, and *stop*.

In most dialects of North American English, /o/ has been unrounded and lowered to [ɑ]. In many of these dialects, /o/ has moved towards the front, and is unrounded. In these dialects, the best phonetic representation would be [a]. This is especially the case in the North where the fronting of /o/ as the second stage of the Northern Cities Shift has caused /o/ to move close to the position formerly occupied by /æ/. In other dialects, /o/ has maintained its roundedness, merging with /oh/ in the low back position. This is the case for the Western Pennsylvania dialect centered around Pittsburgh.

The symbol /oh/ is used to represent the long open-*o* class, and corresponds to Wells' THOUGHT lexical set. It is derived primarily from the monophthongization of the Middle English diphthong **au**, which itself was derived from a variety of sources (such as Old

---

[1] All of the speakers in my dissertation corpus have the merger of /o/ and /ah/, as is evidenced by the near-minimal pair of *father* and *bother*. For speakers who maintain a distinction between /o/ and /ah/, the /o/ vowel corresponds to LOT, and the /ah/ vowel corresponds to Wells' PALM.

English /aw/, OE /a/ + /x/, as in *fought*, vocalization of OE coda /g/, as in *draw*, and Middle French loan words, as in *applaud*). Another large source for /oh/ words was the lengthening of /o/ to /oh/ before voiceless fricateves, as in *lost*, and the velar nasal, as in *strong*. The distribution of /oh/ is severely restricted, and it occurs before only a small number of consonants, mainly before /t/, /d/, /k/, /z/, /n/, /l/, and word-finally. Some examples of words with /oh/ include *thought*, *hawk*, *caught*, and *law*.[2]

In dialects of North American English where /o/ and /oh/ have not merged, /oh/ has changed in three different directions: 1) In the Mid-Atlantic region and New York City it has raised substantially and developed a central offglide, 2) In many areas of the South, it has developed a back upglide, and 3) In the North, it has lowered and fronted as Stage 3 of the Northern Cities Shift. In dialects where /o/ and /oh/ have merged, /oh/ can become unrounded and rather front, especially in the West.

6.3 /o/ and /oh/ in the Midland and the North

In this section, the distributions of /o/ and /oh/ for two typical speakers from the dialect regions neighboring Erie will be examined. First, a speaker from Buffalo will display a clear distinction between the two vowel classes along with the Northern fronting of /o/. Next, a speaker from Pittsburgh will demonstrate the low-back realization of the phoneme resulting from the merger of /o/ and /oh/ in Western Pennsylvania.

Figure 6.1 displays all of the individual tokens of /o/ and /oh/ (including both word list and interview data) produced by Walter K., an 82-year-old Sun Valley resident from Buffalo. Walter K.'s distributions of /o/ and /oh/ show an almost perfect separation between the two classes. Only a single token of /oh/ (the token of *talk* labeled in Figure 6.1) overlaps with the distribution of /o/, and no tokens of /o/ overlap with the /oh/ distribution.

---

[2] See Labov et al. (2006:58) for a more complete description of the historical sources of /o/ and /oh/, as well as words exemplifying all possible segmental environments for the two vowels.

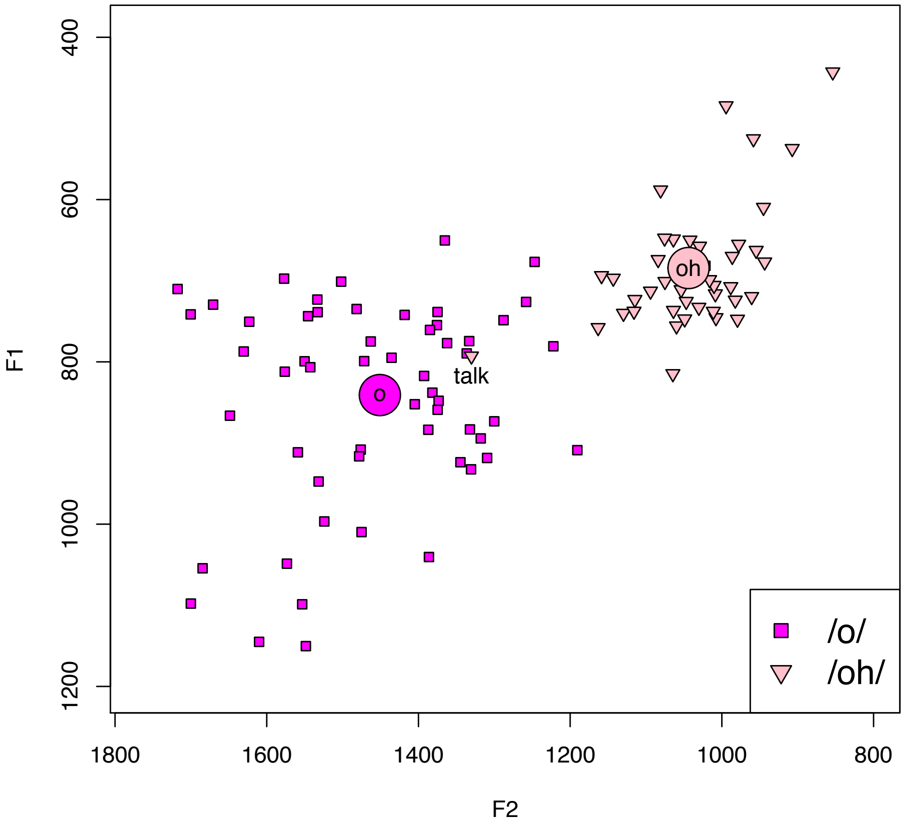

Figure 6.1: /o/ and /oh/ from Walter K., born 1927 in Buffalo,
Mean(/o/) = (841, 1451), N=56; Mean(/oh/) = (684, 1044), N=24; Dist(/o/, /oh/) = 436

Walterk K.'s /o/ is fronted (although not as extremely as it is for many younger speakers from the North) with a mean F2 value of 1451 Hz. The Euclidean distance between the two vowel means is quite large at 436 Hz.

The acoustic evidence for Walter K. demonstrates clearly that he maintains a distinction between /o/ and /oh/, and presents a typical distribution for a Northern speaker. Experimental evidence from minimal pair tests confirms that Walter K. maintains the distinction:

he produced the minimal pairs *cot* / *caught* and *Don* / *dawn* as clearly distinct, and also judged them to be different perceptually.

The caption for Walter K.'s plot of /o/ and /oh/ in Figure 6.1 illustrates a few notational conventions that I will use when displaying vowel plots. First, the notation Mean(V) will be followed by a tuple containing the F1 and F2 mean values for the vowel V, along with the number of tokens that were used to calculate the mean values.[3] Secondly, the notation Dist(V1, V2) will be used to represent the two-dimensional Euclidean distance between the F1 and F2 means for the two vowels V1 and V2, as calculated in Equation 6.1.

As an example of a Midland speaker from Western Pennsylvania, Figure 6.2 displays a plot of /o/ and /oh/ for Gwen S., a speaker from the ANAE (TS # 355)[4]. Gwen S. was born in 1929 in Pittsburgh, and provides a clear example of the solid merger of /o/ and /oh/ in that region. Figure 6.2 shows almost complete overlap between the tokens of /o/ and /oh/: the F1 and F2 ranges for both vowels are very similar, and both vowel classes have tokens distributed throughout their entire ranges. Gwen S.'s /o/ has remained a low back vowel, with a mean F2 value of 1149 Hz. Additionally, most of her tokens of /o/ are also clearly rounded perceptually. The Euclidean distance between her means of /o/ and /oh/ is 66 Hz, also indicating that she has a solid merger. Furthermore, the experimental evidence corroborates this acoustic evidence: all ANAE minimal pair tests involving /o/ and /oh/ show Gwen S. to be merged in both production and perception.

Walter K. and Gwen S. occupy two ends of a continuum representing potential realizations of /o/ and /oh/ in the North and the Midland. Additionally, since they are from

---

[3] All mean values for individual vowel classes reported in this section were calculated using the exclusions described in Section 3.8. Thus, these tokens are not counted in the N values associated with each mean value. However, all tokens, even the ones not contributing to the mean values, are displayed in the vowel plots. The only tokens not displayed in the plots are extreme outliers caused by measurement errors. These are omitted from the visual display so that the distributions of the correct measurements can be observed more clearly. However, they are not excluded from the calculation of the mean values.

[4] As described in Section 3.3.2, my analyses of the ANAE speakers are not based on the publicly released Plotnik files containing the annotations used for the ANAE, but rather my own re-analysis using the methodology of forced alignment and automatic vowel analysis

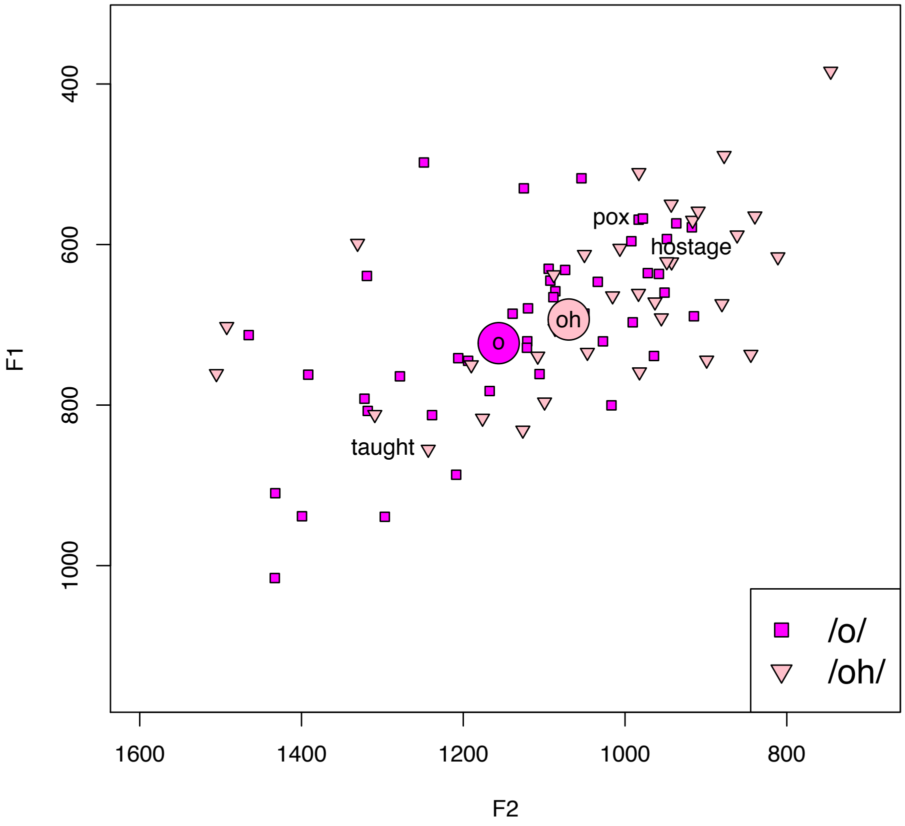

Figure 6.2: /o/ and /oh/ from Gwen S., born 1929 in Pittsburgh,
Mean(/o/) = (723, 1149), N=45; Mean(/oh/) = (689, 1092), N=10; Dist(/o/, /oh/) = 66

Buffalo and Pittsburgh, they also represent two ends of the geographic continuum that the data for this dissertation was drawn from. The evidence for the status of /o/ and /oh/ is not always as clear as it is in the cases of Walter K. and Gwen S., especially if the analysis involves speakers who did not provide experimental evidence in the form of minimal pair tests (for example, speakers who were drawn from archival sources). In cases where such experimental evidence is lacking, however, it is still useful to be able to distinguish merged from unmerged and transitional speakers based on the acoustic evidence alone. The following section will review the techniques that have been used to characterize a speaker's status with respect to the low-back merger (especially speakers whose distributions are not as clear-cut as Walter K. and Gwen S.) and describe my reasons for choosing the ones that I use for the analyses in this chapter.

6.4 Determining whether a speaker has the merger of /o/ and /oh/

As the previous section demonstrated, the status of some speakers with respect to the merger of /o/ and /oh/ is clear; for other speakers, however, the evidence can be less conclusive. On the other hand, in the case of merged speakers, the acoustic evidence often does not show a complete overlap between the distributions for the merged vowels, due to different allophonic constraints in the words that belong to each class. In the case of unmerged speakers, the distributions for the two classes are often not completely discrete. For such speakers, it is often difficult to determine from the acoustic evidence alone whether they maintain a complete distinction, or whether they are transitional; in these cases, experimental methods provide a more reliable means of categorizing speakers. However, in the data for this dissertation, speakers are drawn from both my own interviews and archival recordings, where experimental evidence is necessarily lacking. Therefore, it will be useful to use a metric that is applicable to all speakers in the corpus.

The experimental method that provides the most reliable characterization of a speaker's phonemic status is the commutation test. In this test, the speaker is recorded uttering several tokens of each item in a minimal pair (these tokens can come either from spontaneous speech or in an elicitation context). Then, the tokens are played back in random order and the speaker is asked to label each of the utterances. Speakers who maintain a phonemic distinction between the two words in the pair generally attain 100% accuracy on this task whereas speakers with a merger perform at chance level.[5] The commutation test methodology, however, was impractical given the logistical constraints of the fieldwork setting for this dissertation, and was not used.

A simpler experimental task that also generally provides reliable results about the status of a speaker's phonemic contrast between two vowels is the minimal pair test. In this test, the speaker is asked to pronounce the two words in a minimal pair in direct succession. Then, he tells the analyst whether they sound the same or different to him, and pronounces the pair a second time. This elicitation technique thus quickly obtains two sources of information about the two words in the minimal pair: the speaker's own *perception* about whether they are the same or different and two acoustic records of the speaker's *production* of each item in the minimal pair. In general, the results obtained from the perception and production parts of the test are identical for each minimal pair, and the speaker's status is easy to determine. The main cause of a discrepancy between the two is when a speaker is from a transitional area; in these cases, the two items are often pronounced only slightly differently. These tokens are provided with the intermediate label of "close". Also, the judgments of a speaker in a transition area may not match the production data. In general, though, the results of the minimal pair tests are reliable and consistent, and they will be

---

[5] In cases of a near-merger, the interpretation of the results can be more complicated. This is the case with the distinction between *ferry* and *furry* in Philadelphia (Labov et al. 1991).

given primary importance for determining the phonemic status of /o/ and /oh/ for a given speaker, when available.[6] However, this source of information does not exist for speakers drawn fromarchival recordings. For those speakers, a metric based on the acoustic evidence alone is required.

A statistical method that has frequently been used to determine whether a speaker maintains a distinction between two vowel classes on the basis of formant measurements is the unpaired *t*-test. Herold (1990) applied this methodology in an attempt to determine whether a speaker's F1 and F2 means for /o/ and /oh/ were significantly different or not. However, she ran into the difficulty that a *t*-test comparing the means of /o/ and /oh/ for several speakers' interview data produced statistically significant differences, even though the distributions for the two vowels overlapped substantially. Furthermore, these speakers were clearly judged as merged based on perceptual tests (Herold 1990:73). She argued that this apparent paradox was due to an imbalance in the types of consonants that can follow each of the vowels, and concluded that this uneven allophonic distribution thus makes the unpaired *t*-test unsuitable for comparing mean values of /o/ and /oh/.

Johnson (2007:284–289) introduced the paired *t*-test as a technique in overcoming this imbalance. In this test, items from minimal pairs (or near-minimal pairs containing similar segmental environments) are directly compared when the *t*-statistic is computed for the means of the two classes. Thus, any potential differences due to an imbalance in the segmental environments between the two classes is factored out. This method thus produces more reliable results than a simple unpaired *t*-test of tokens from spontaneous speech. However, it requires a large number of minimal pairs to be recorded. In Johnson's study, this was not a problem, since the low back vowels were the only target for analysis. For the

---

[6] The case of Bill Peters from Duncannon, PA is a notable exception to the accuracy of the minimal pair test in determinining the phonemic status of two vowels. He produced all minimal pairs involving /o/ and /oh/ identically and also judged them to be the same perceptually. However, his spontaneous speech showed a large and consistent difference between tokens from the two classes (Labov et al. 1972:235–236).

present study, however, the interview procedure required time for several other elicitation tasks; thus, it was not possible to include enough minimal pairs to produce statistically reliable results from a paired *t*-test.

Finally, various distance metrics between the formant values for the two vowel classes can be examined. The Euclidean distance between the F1 and F2 mean values is the most straightforward and most commonly used method. Equation 6.1 provides the formula for calculating the Euclidean distance between two vowel means in F1 and F2 space, where V1 and V2 are the means of the two vowels under comparison.

$$Dist(\mathrm{V1}, \mathrm{V2}) = \sqrt{(\mathrm{F1}_{\mathrm{V1}} - \mathrm{F1}_{\mathrm{V2}})^2 + (\mathrm{F2}_{\mathrm{V1}} - \mathrm{F2}_{\mathrm{V2}})^2} \tag{6.1}$$

Speakers with the merger of /o/ and /oh/ will usually have a lower value for the Euclidean distance than unmerged speakers. Transitional speakers often have intermediate values.

This method of comparing the means for two vowels can also be influenced, as is the unpaired *t*-test, by an unbalanced distribution of the segmental environments for the tokens from two vowel classes. However, it does provide a useful way of comparing speakers whose acoustic data comes from different sources and was obtained through different methods (as is the case for this dissertation's corpus). In the case of the "Arthur the Rat" recordings from the DARE corpus, it is an ideal means for comparing the DARE speakers with each other. Since these speakers all uttered the same words, the Euclidean distances can be calculated over sets of tokens with the same environmental contexts.

The analyses in this section will combine the results of minimal pair tests, when available, with the Euclidean distance metric in determing the phonemic status of /o/ and /oh/ for a given speaker. The Euclidean distance rarely approaches 0, even for a completely merged speaker, due to the different historically-derived allophonic distributions of the two classes.[7] The results from the ANAE show that the average Euclidean distance for speakers

---

[7] The mean F2 value of /o/ is usually greater than the mean F2 value of /oh/ for both merged and un-

from three regions with the merger (West, Canada, and Western Pennsylvania) is less than 100 Hz, and it is less than 200 Hz for speakers from Eastern New England. On the other hand, the average Euclidean distance for speakers from the Inland North is around 300 Hz.

In addition to the minimal pair tests and the Euclidean distance metric, vowel plots showing the locations of all tokens of /o/ and /oh/ will be analyzed for several speakers. Speakers who have a clear distinction between the two classes usually have very little overlap between the clouds of tokens for /o/ and /oh/. On the other hand, speakers with the merger have a substantial overlap between the two classes. Speakers who are transitional may have a smaller number of tokens overlapping, in addition to having many indeterminate tokens in the intermediate area between the two classes.[8]

6.5 Previous sources of information about the merger

Before describing the results from my own research pertaining to the status of /o/ and /oh/ in Erie, I will review the prior sources of information that are available. There are three previously published sources that provide evidence for the status of /o/ and /oh/ in Erie:

- the field surveys for LAMSAS (as published in Wetmore (1959) and Kurath and McDavid (1961)), which included two informants from rural areas of Erie County
- the telephone survey of the state of Pennsylvania conducted by Herold (1990), which included one speaker from the city of Erie
- the interviews for ANAE, which included two speakers from the city of Erie

---

merged speakers. Likewise, the mean F1 value of /o/ is usually greater than the mean F1 value of /oh/ for both groups of speakers.

[8] This discussion assumes that the merger is proceeding by approximation, not transfer or expansion (Labov 1994:321–323). The vowel plots shown for the transitional archival speakers in Section 6.7 tend to support this view.

Speaker #	Township	County	Year of Birth	Type	Sex	Occupation
NY64a	Westfield	Chautauqua	1868	I	M	village clerk
NY64b	Westfield	Chautauqua	1869	II	F	florist
NY64c	Westfield	Chautauqua	1884	III	F	librarian
PA55a	Springfield	Mercer	1855	I	M	farmer
PA55b	Findley	Mercer	1900	II	M	farmer
PA56a	Canal	Venango	1857	II	M	teacher
PA56b	Richland	Venango	1896	II	M	farmer
PA65a	Sugar Grove	Warren	1866	I	M	farmer
PA65b	Triumph	Warren	1889	II	M	farmer
PA66a	East Fallowfield	Crawford	1859	I	M	farmer
PA66b	Conneaut	Crawford	1890	II	M	farmer
PA67a	Venango	Erie	1864	I	M	farmer
PA67b	Amity	Erie	1903	II	M	farmer

Table 6.1: Demographic information for the 13 LAMSAS informants shown in Figure 6.3

Section 2.3 presented the evidence from Kurath and McDavid (1961) and Wetmore (1959) that demonstrates that the two LAMSAS informants from Erie County maintained a consistent distinction between /o/ and /oh/. This conclusion is based on over 20 lexical items from the LAMSAS survey.[9] In order to provide a better context for interpreting the evidence from these two LAMSAS speakers, the status of /o/ and /oh/ for the other LAMSAS speakers from the neighboring regions will be investigated.

Table 6.1 presents the demographic information for the LAMSAS speakers from Erie County and the neighboring regions. The "Type" column displays a subjective classification of each speaker's social characterists according to the following three-way scheme: "folk speakers" are Type I, "common speakers" are Type II, and "cultivated speakers" are Type III (Kretzschmar et al. 1993:25).

---

[9] These two speakers behave consistently for the 7 relevant lexical items mapped in Kurath and McDavid (1961). Wetmore (1959) had access to the original, unpublished field notes for LAMSAS, and was able to examine several additional lexical items. Based on this, he concluded that both speakers from Erie County had the distinction, although he doesn't provide information about the specific transcriptions used for these speakers in the additional lexical items he examined.

The status of /o/ and /oh/ for these 13 LAMSAS speakers is displayed in Figure 6.3. In most cases, the decision about whether to classify a LAMSAS speaker as merged or not was based on the classifications provided by (Wetmore 1959:113).[10] The only speakers that were not included in Wetmore's study are the three from Chautauqua County, NY. Their status was determined by examining the same maps from PEAS that were analyzed in Table 2.9 for the two speakers from Erie County.

Figure 6.3 suggests that the merger of /o/ and /oh/ was in the process of spreading through Crawford County in the last few decades of the 19th century, but had not yet reached Erie or Warren Counties at this time. The fact that the younger of the two LAMSAS informants from Erie County was born in 1903 indicates a time around the first decade of the 20th century as the *terminus post quem* for the merger of /o/ and /oh/ in Erie.[11]

The earliest study to document the merger of /o/ and /oh/ in Erie County is Herold (1990). She conducted a telephone survey in 1987–1988 of all of the counties in Pennsylvania that were reported as distinct based on the interpretation of the LAMSAS data in Wetmore (1959) in order to track the progress of the merger in the state. She interviewed one speaker from Erie County: a 63-year-old female from the city of Erie. Through a series of elicitations and minimal pair tests she concluded that this speaker had the merger of /o/ and /oh/. This evidence thus suggests a time around 1925 as the *terminus ante quem* for the merger of /o/ and /oh/ in the city of Erie. Combining the chronology from LAMSAS and Herold (1990) suggests a window of about a generation in the second two decades of the 20th century for the completion of the merger in Erie County.

The results from Herold's survey also provide similar information about Warren County. She interviewed a 59-year-old male from the city of Warren who she judged to have the

---

[10] Additionally, the LAMSAS field notes describe speaker PA66a as having the merger and speaker PA66b as maintaining the distinction (Kretzschmar et al. 1993:262).

[11] This line of reasoning assumes that the merger spreads uniformly throughout all towns in a county. This abstraction is clearly an over-simplification—a more detailed analysis that takes into account the specific locations of the LAMSAS speakers in Erie County will be presented in Section 6.10.

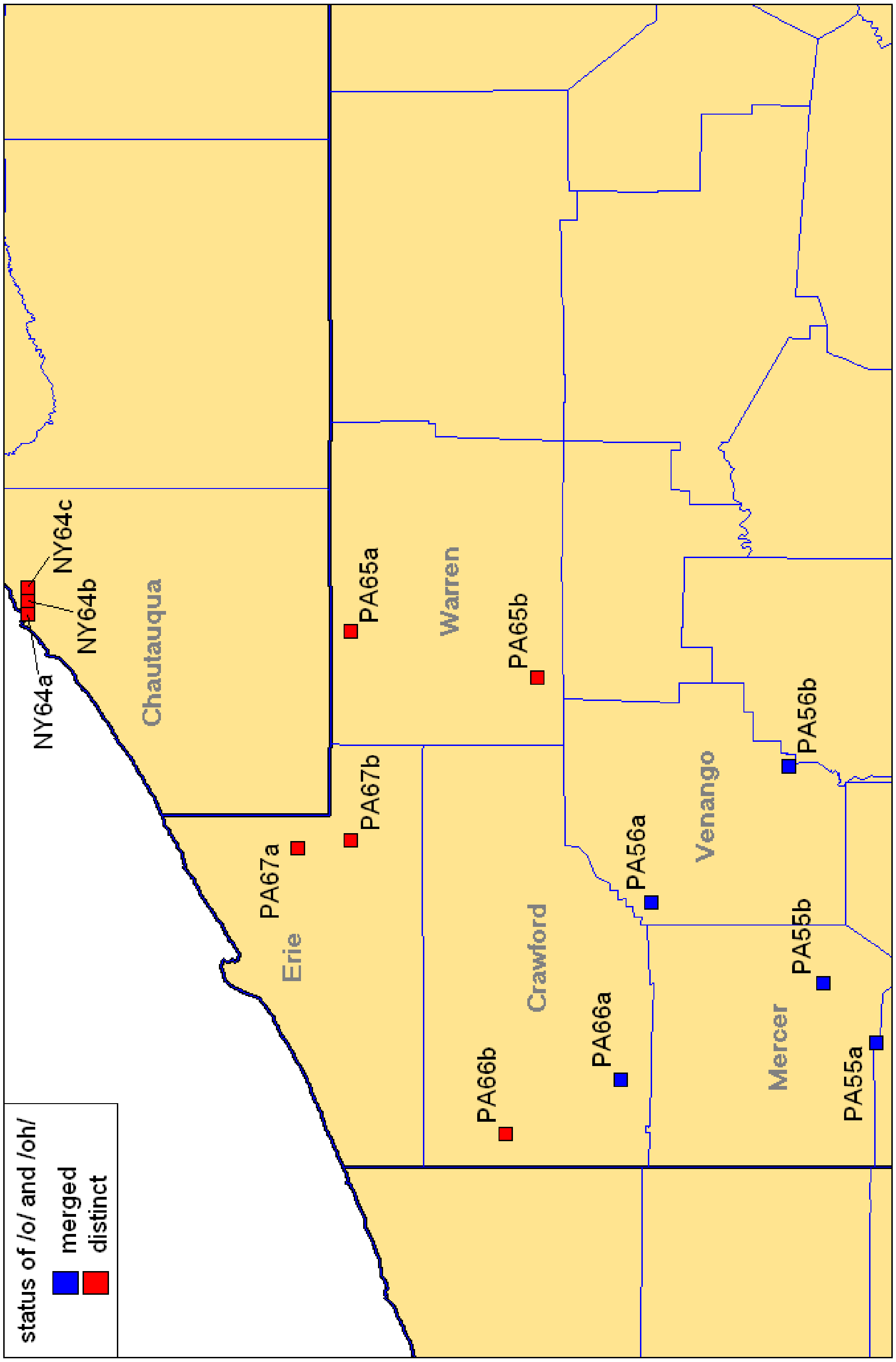

Figure 6.3: The status of /o/ and /oh/ for 13 LAMSAS speakers from the region around Erie

merger. Given that the youngest LAMSAS speaker from Warren County was born in 1889, this would again suggest that the merger spread to Warren County in the first few decades of the 20th century.

Subsequently, two female Erieites were interviewed in 1995 for the ANAE survey. At the time, they were 31 and 39 years old, and both had a solid merger of /o/ and /oh/ in perception and production. This finding is not surprising, since these speakers were born several decades after the merged Erieite interviewed by Herold (1990).

The following section will supplement these previous studies with apparent time evidence from my own fieldwork conducted in the region around Erie as well as archival recordings of speakers from the area.

6.6 The city of Erie: an apparent time study

If the *terminus post quem* for the merger of /o/ and /oh/ in Erie suggested by the LAMSAS data is correct, then it might still be possible to find some elderly Erieites who were born before the merger took place in Erie. In order to test this hypothesis, I conducted an apparent time study at a retirement center in Erie, which I will call Sun Valley. In total, I conducted interviews with 12 senior citezens at the center, 9 of whom are life-long residents of the city of Erie (see Section 3.3.1 for more details about these speakers).

Despite the fact that I was able to interview several elderly Sun Valley residents, including four who were born before 1920, none of them showed any trace of a distinction between /o/ and /oh/. It is clear from both minimal pair tests and the acoustic measurements taken from interviews and word lists that all of the native Erieites interviewed at Sun Valley have a complete merger between /o/ and /oh/. None of them had a difference in production of perception for any of the minimal pairs, and the vowel plots show almost total overlap between the two classes.

For example, consider the vowel plot shown in Figure 6.4 for Dan R. He was born in 1912, and is the oldest speaker I interviewed at Sun Valley. His means for /o/ and /oh/ are only separated by 10 Hz in the F1 dimension and 78 Hz in F2. The two vowel clouds show considerable overlap throughout their entire ranges. To complement this acoustic evidence, the minimal pair data from Dan R. also point to a complete merger. He produced the pairs *cot* / *caught* and *Don* / *dawn* identically and judged them both to be the same.

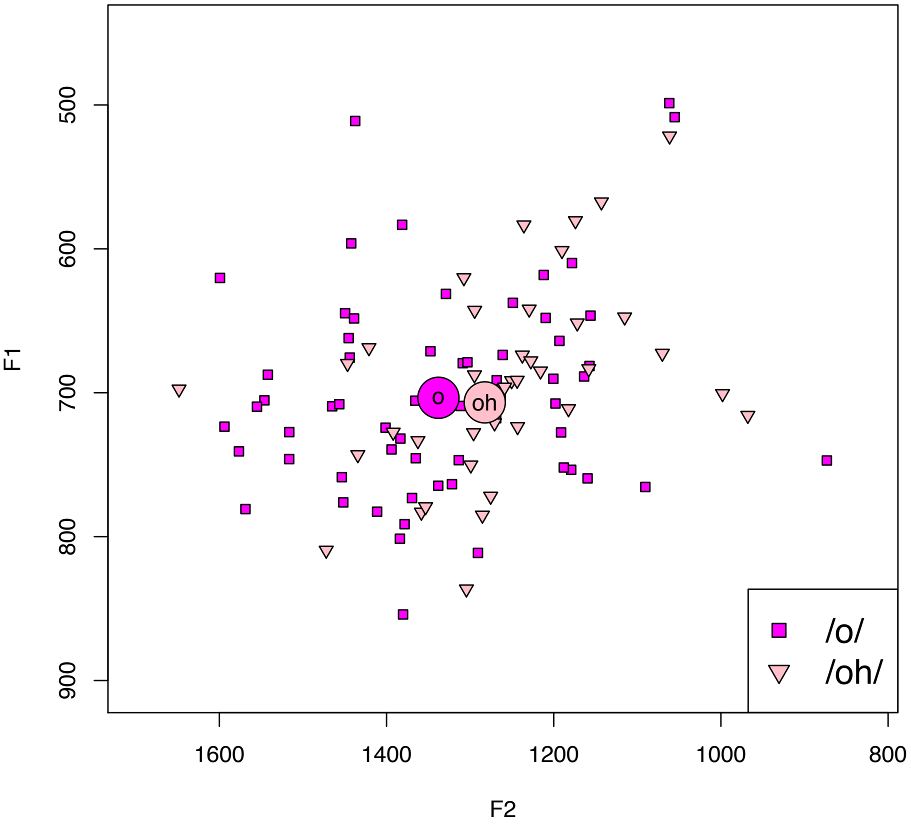

Figure 6.4: /o/ and /oh/ from Dan R., born 1912 in Erie,
Mean(/o/) = (704, 1338), N=55; Mean(/oh/) = (707,1283), N=31; Dist(/o/, /oh/) = 55

Table 6.2 displays the mean F1 and F2 values for /o/ and /oh/, as well as the Euclidean distances between the two vowels, for all of the Sun Valley residents from Erie. All of the

Euclidean distances between /o/ and /oh/ for these nine speakers are around 200 Hz or below, much lower than the value of 436 for Walter K., the unmerged Sun Valley resident from Buffalo shown in Figure 6.1. Furthermore, the vowel plots for all speakers are similar to Dan R.'s in Figure 6.4, and show that the clouds for the two classes overlap substantially. Finally, the minimal pair results for these nine Sun Valley speakers agree with the acoustic evidence and confirm that all these speakers have a solid merger of /o/ and /oh/: all nine speakers had the pairs *cot* / *caught* and *Don* / *dawn* merged in both production and perception.

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Dan R.	1912	(717, 1361), N=56	(707, 1283), N=31	78
Robert E.	1916	(693, 1463), N=34	(674, 1261), N=17	203
Flora R.	1919	(834, 1386), N=71	(811, 1300), N=30	89
Mary D.	1919	(856, 1402), N=36	(793, 1255), N=23	160
Charles B.	1925	(790, 1222), N=36	(820, 1168), N=11	62
Eloise B.	1925	(778, 1331), N=48	(758, 1155), N=27	177
Dottie A.	1926	(771, 1284), N=58	(758, 1120), N=22	165
Sally W.	1928	(733, 1285), N=31	(713, 1157), N=18	130
Dana W.	1941	(814, 1287), N=67	(745, 1166), N=38	139

Table 6.2: /o/ and /oh/ or 9 Sun Valley residents from Erie

The clear evidence for the merger of /o/ and /oh/ among several Sun Valley residents aged 80 and above indicates a time around 1915 as the *terminus ante quem* for the merger of these two vowels in Erie (pushing this date back by about 10 years from what was suggested by Herold's telephone survey). This evidence, along with the LAMSAS data discussed in Section 6.5, would seem to indicate a short window in the second decade of the 20th century for the merger's occurrence.

In order to shed more light on this chronology, I attempted to find older recordings of Erieites born before the Sun Valley residents. The results from these archival materials will be presented in the following section.

6.7 Archival evidence

This section presents evidence for the chronology of the merger of /o/ and /oh/ in Erie drawn from the archival sources described in Section 3.3.2. Most of these speakers were born over 100 years ago, and thus push the time depth of the acoustic evidence back a few decades earlier than the apparent time evidence from the elderly Sun Valley residents. Additionally, they are all approximately contemporaneous with the LAMSAS speakers described in Section 6.5. Thus, it will be possible to complement the impressionistic data provided by LAMSAS from that time period with acoustic data.

6.7.1 SWV corpus

The two speakers selected for analysis from the SWV corpus (see Section 3.3.2), Richard O. and Benjamin S., were chosen because they are the two oldest speakers in the corpus. Richard O. was born in 1906, and Benjamin S. in 1907. They both lived their entire lives in North East, PA.

Figures 6.5 and 6.6 show plots of the vowels /o/ and /oh/ for Richard O. and Benjamin S., respectively.[12] Richard O. has only a small amount of overlap between the two classes. This distribution suggests that he maintained a distinction between /o/ and /oh/. However, the distributions for two phonemes are quite close: the Euclidean distance between /o/ and /oh/ for Richard O. is only 201 Hz, compared to 436 Hz for the clearly unmerged Walter K. from Buffalo. This is the type of distribution that exists for many of the Midland ANAE speakers who are labeled as "transitional" with regard to the /o/ ∼ /oh/ merger (Labov et al. 2006:270).

Benjamin S. shows a much greater degree of overlap between the two classes, with several tokens from each class falling clearly within the cloud of the other class. Furthermore,

---

[12] As mentioned in Section 3.3.2, the vowel formant measurements for these two speakers are the only ones that were extracted manually for this dissertation, due to the poor sound quality of the archival recordings.

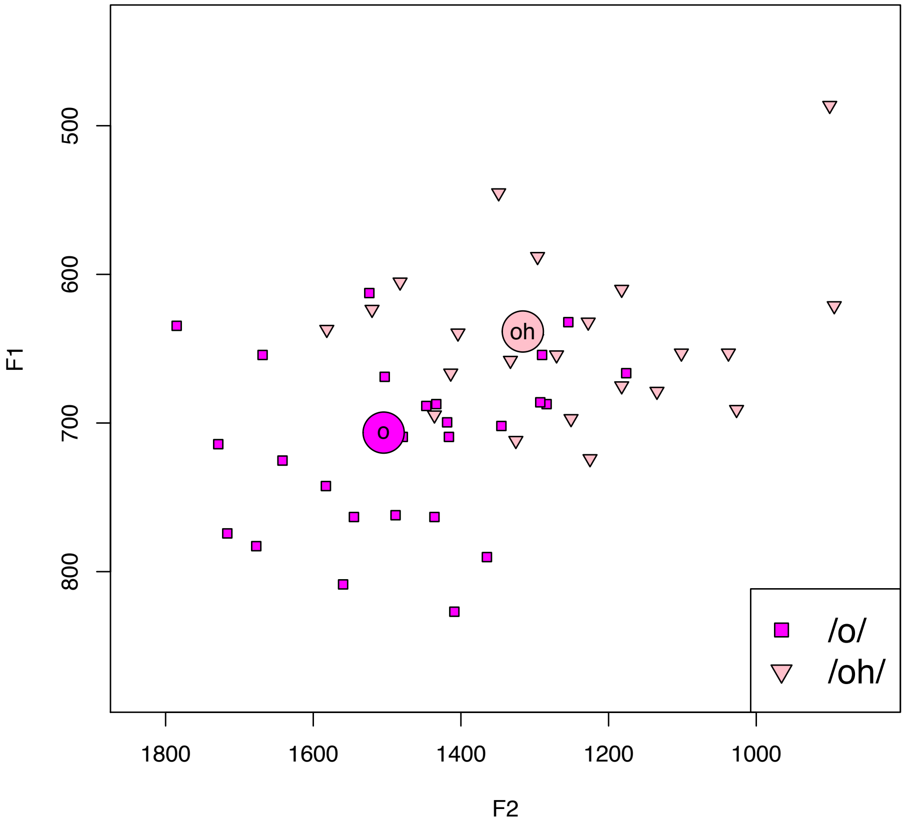

Figure 6.5: /o/ and /oh/ from Richard O., born 1906 in North East, from the SWV corpus, Mean(/o/) = (706, 1505), N=19; Mean(/o/ = (638, 1316), N=10; Dist(/o/, /oh/) = 201

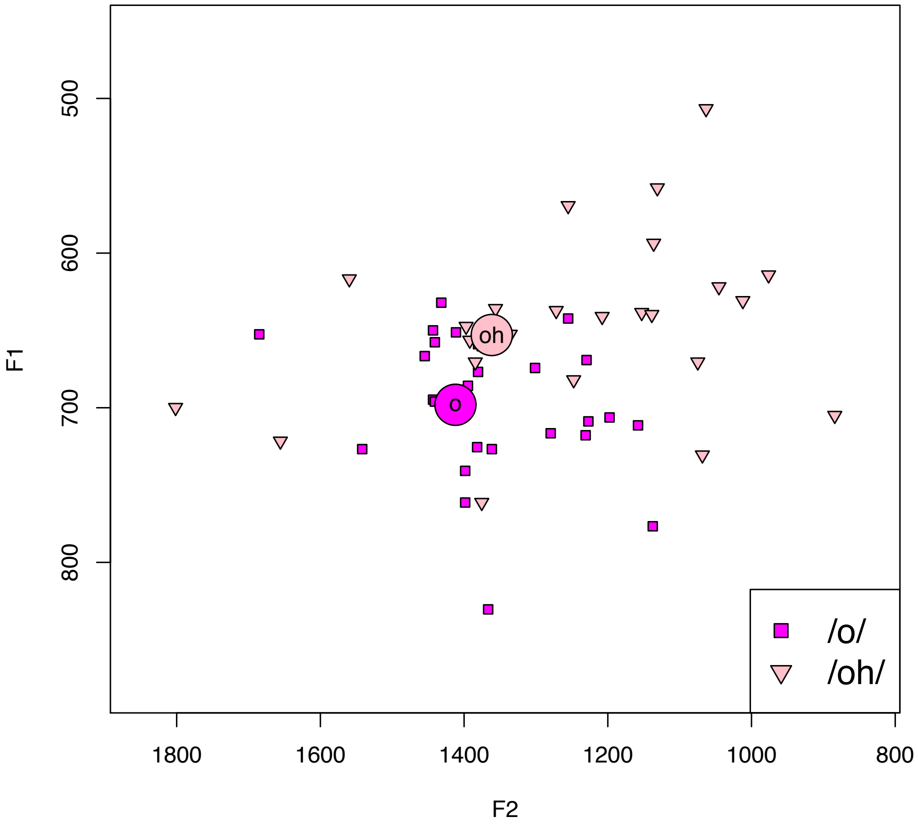

Figure 6.6: /o/ and /oh/ from Benjamin S., born 1907 in North East, from the SWV corpus,
Mean(/o/) = (698, 1412), N=15; Mean(/oh/) = (653, 1361), N=12; Dist(/o/, /oh/) = 68

the Euclidean distance between the means of the two classes is only 68 Hz. All of this evidence suggests that the merger of /o/ and /oh/ is quite advanced for Benjamin S., and has probably already reached completion for him.

Thus, the evidence from Richard O. and Benjamin S. suggests that the merger was already in transition in North East in the first decade of the 20th century. However, this would seem to contradict the LAMSAS evidence showing a distinction for the speaker from Erie County born in 1903. The following two sections will attempt to address this apparent contradiction by providing more evidence from the same time period from other areas of Erie County.

6.7.2 H. O. Hirt

H.O. Hirt is the oldest recorded speaker in the corpus from the city of Erie itself (see Section 3.3.2 for his demographic details). He was born in 1887, and thus pushes the time depth for our knowledge of the city of Erie back about 20 years from Dan R., the oldest Sun Valley resident. Figure 6.7 shows a plot of all tokens with /o/ and /oh/ from the interview with Hirt.

As the figure shows, there is a large amount of overlap between /o/ and /oh/, especially in the boundary area between the two distributions. However, the 250 Hz distance between the means is somewhat larger than would be expected for a completely merged speaker. As was the case for Richard O. from the SWV corpus, the distributions of /o/ and /oh/ for H.O. Hirt seem to indicate that the distinction was tenuous for him, and that he is transitional with respect to the merger. Thus, the merger was already in progress in Erie when Hirt was acquiring language in the early 1890's. Again, this seems to contradict the LAMSAS evidence, but is consistent with the evidence presented in Section 6.6 for the 9 elderly Erieites interviewed in 2007 at the retirement community and in Section 6.7.1 for the two elderly participants in the SWV corpus.

6.7.3 DARE

Another archival source that provides an early source of acoustic data for speakers in the region around Erie is DARE. While the DARE survey was primarily focused on elicitation of lexical items through a written questionnaire, interviews were recorded with 1,843 sub-

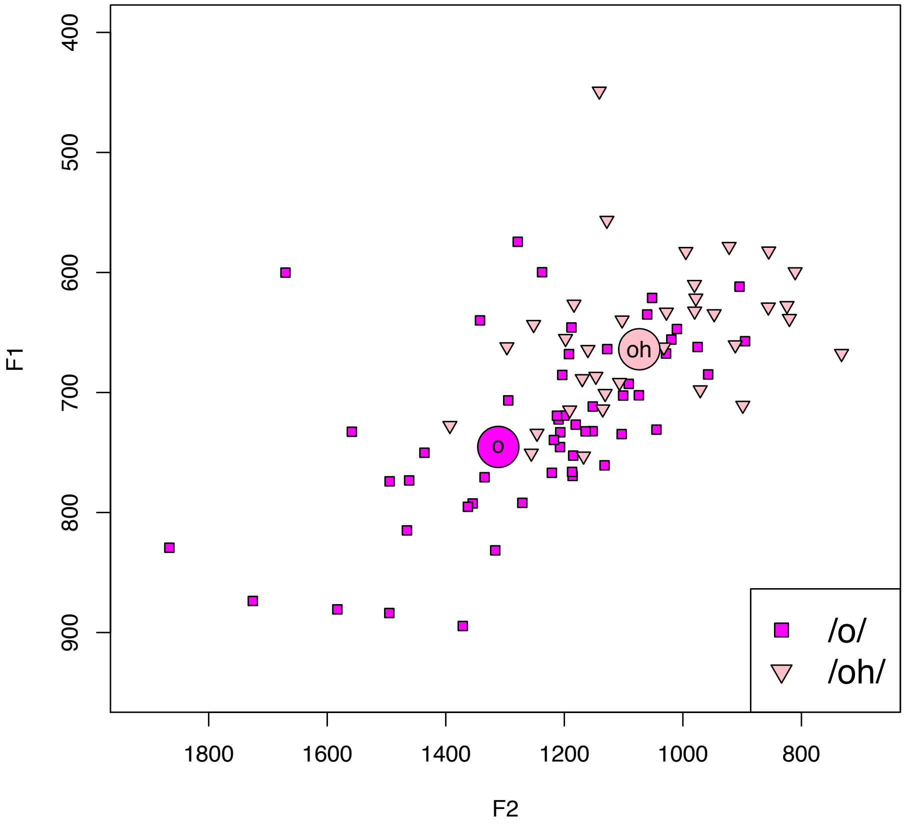

Figure 6.7: /o/ and /oh/ from H. O. Hirt, born 1887 in Erie,
Mean(/o/) = (745, 1311), N=36; Mean(/oh/) = (664, 1074), N=21; Dist(/o/, /oh/) = 250

/o/	/oh/
on (6x) *not* (3x) *foggy* *got* *odd* *rotted* *upon*	*long* (3x) *caught* (2x) *loft* (2x) *off* (2x) *coughed* *haunted* *log* *longer* *undaunted*
N(/o/) = 14	N(/oh/) = 14

Table 6.3: Words with the vowels /o/ and /oh/ contained in DARE's "Arthur the Rat" reading passage

jects, about 2/3 of the total number of DARE subjects (Cassidy and Hall 1985:xiv). These interviews are usually about 30 minutes in length, and cover topics such as regional history, agricultural practices, and local traditions. In addition to the conversational interview, most subjects were also recorded reading the story "Arthur the Rat". This story contains 593 words, and has a relatively balanced distribution of all English vowels in a variety of segmental environments. It thus provides an efficient means for comparing speakers from the DARE corpus.

Table 6.3 lists all of the words included in the "Arthur the Rat" passage that were used for calculating the means for /o/ and /oh/, including the number of occurrences for words that appear more than once in the passage. The word *on* is listed as belonging to the /o/ class for this study, despite the fact that its pronunciation varies between /o/ and /oh/ in North America. For the region under consideration, the ANAE shows near-categorical use of /o/ in *on* for Northern speakers (Labov et al. 2006:189). Also, the five unmerged speakers in my corpus from Chautauqua County and Buffalo who had *on* and *Don* in their list of minimal pairs all produced *on* with /o/; i.e., *on* and *Don* rhymed for all of them.

Since there is no evidence to expect *on* to be pronounced with /oh/ for any of the speakers in this region, it is assumed that it contains the vowel /o/ for the DARE speakers.

Additionally, the "Arthur the Rat" reading passage contains several words with /o/ and /oh/ that were excluded from the calculations of the mean values of /o/ and /oh/, since they match the exclusion criteria listed in Section 3.8. For /o/, these words are *watched* and *washing*; for /oh/, they are *all* (4x), *always*, *crawled*, *walk*, and *walls*. Although these tokens were excluded from the calculation of the mean values, they are still displayed in the vowel plots.

The individual token counts for /o/ and /oh/ for each speaker analyzed below often vary slightly from the 14 that would be expected based on the list in Table 6.3. These differences are caused either by the exclusion of suspected mis-measurements according to the procedure described in Section 3.8 or by individual deviations from the "Arthur the Rat" text in the DARE recordings.[13]

Unfortunately, no speakers from the city of Erie itself were interviewed for DARE. However, several speakers from the boundary areas around Erie were interviewed and recorded. In total, 14 of these speakers who read "Arthur the Rat" were analyzed for this dissertation. Since these recordings were made from a reading passage, a transcription already exists as potential input to the forced alignment system (the version that was given by the DARE fieldworkers to the subjects to read is printed in Appendix E). However, due to disfluencies, mis-readings, and background noise, the written text of the story does not provide a perfect transcription of each individual speaker's rendition of "Arthur the Rat." Therefore, in order to achieve optimal forced alignment, each speaker's transcription was amended beforehand to reflect any deviations from the original transcription. These improved transcriptions were then used for forced alignment and automatic vowel analysis

---

[13] For the sake of a standardized analysis, the word *log* is included in the /oh/ category for all speakers, despite the fact that its phonemic status is variable among unmerged North American speakers.

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Gladys T.	1899	(840, 1378), N=12	(737, 1163), N=13	238
Bill C.	1950	(727, 1339), N=11	(696, 1093), N=12	248

Table 6.4: /o/ and /oh/ from two DARE speakers from Meadville, PA

of the “Arthur the Rat” reading passages. In addition, two DARE speakers from the town of Ripley, NY were selected for more detailed analysis, because apparent time evidence from my interviews demonstrates that the merger spread to that town during the course of the 20th century (see Section 6.9 for a detailed discussion of the town of Ripley). Excerpts from their interviews were transcribed and analyzed using forced alignment and automatic vowel analysis.

The 14 DARE speakers that were analyzed for this dissertation come from the following 7 locations: North East and Union City (Erie Co., PA); Meadville (Crawford Co., PA); Warren (Warren Co., PA); Fredonia, Ripley, and Jamestown (Chautauqua Co., NY). Figure 6.8 summarizes the data from these speakers by displaying the Euclidean distances between /o/ and /oh/ for each speaker. The specific values for each speaker in this map are presented below in Tables 6.4 - 6.10, where a town-by-town analysis is conducted.

First, consider the two speakers from Meadville, PA, located in Crawford County. The LAMSAS evidence presented in Figure 6.3 showed that the northern boundary for the merger of /o/ and /oh/ crossed through the middle of Crawford County. The acoustic evidence from the oldest DARE speaker in Crawford County, Gladys T., suggests that she was transitional with regard to the merger (she was born nine years later than the younger, unmerged LAMSAS informant from Crawford County, PA66b). The younger DARE speaker from Crawford County appears to be merged from his vowel plot: a large number of tokens from the two classes overlap (although the Euclidean distance between his means of /o/

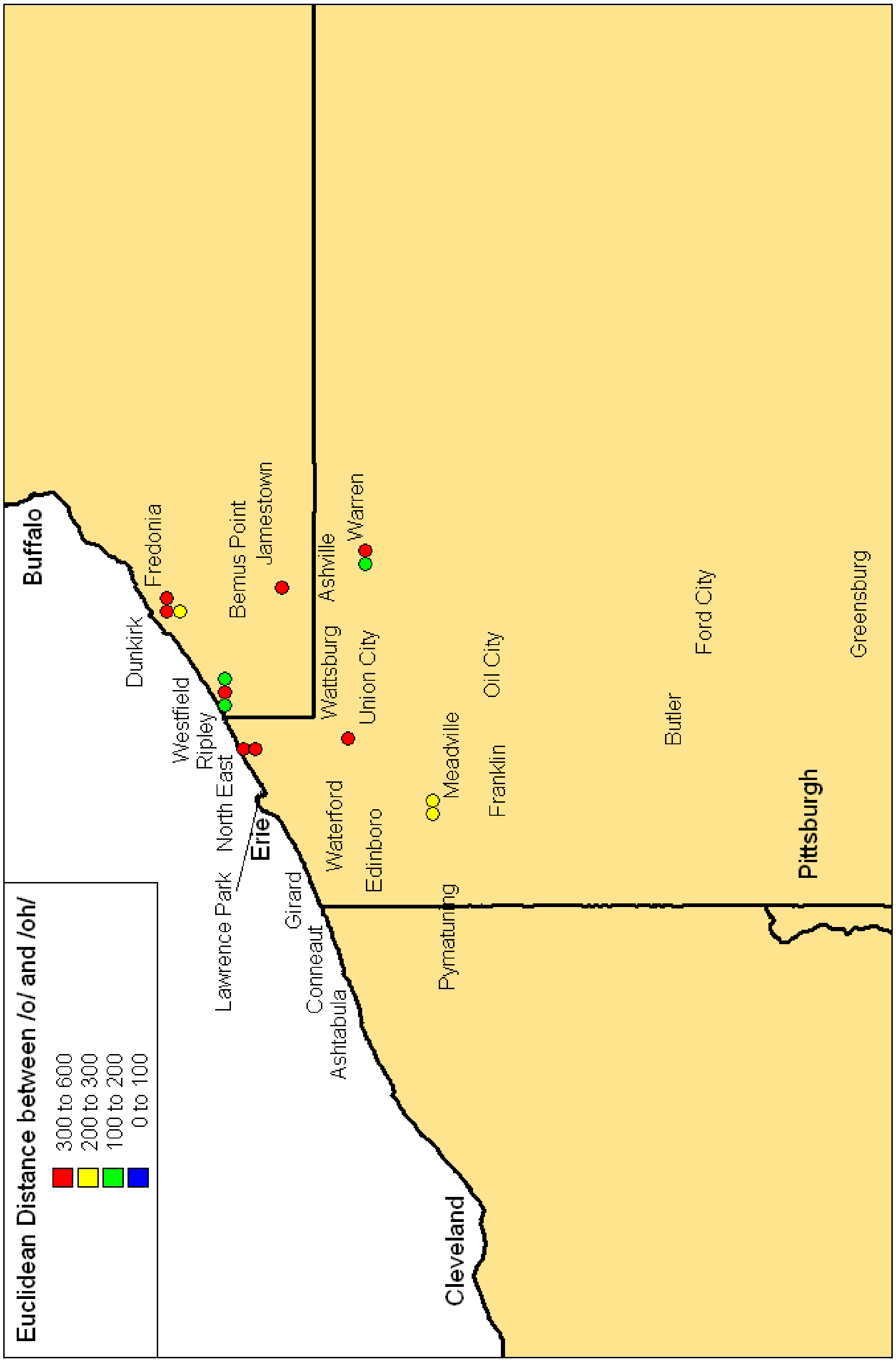

Figure 6.8: Euclidean distances between /o/ and /oh/ for 14 DARE speakers

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Maggie S.	1900	(724, 1524), N=12	(672, 1067), N=14	460

Table 6.5: /o/ and /oh/ from a DARE speaker from Union City, PA, born 1900

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Sarah N.	1897	(682, 1549), N=13	(574, 1130), N=8	433
Nancy S.	1908	(765, 1502), N=14	(687, 1203), N=13	309

Table 6.6: /o/ and /oh/ from two DARE speakers from North East, PA

and /oh/ is somewhat larger than would be expected from a completely merged speaker).

Next, the values for the DARE speaker from Union City are shown in Table 6.5. Her distributions of /o/ and /oh/ are quite clearly separated, and do not show signs of being transitional. Union City is located near Venango and Amity, the locations in Erie County of the two unmerged speakers from LAMSAS. Maggie S.'s data supports the conclusion that the merger had not yet spread to this part of Erie County at the turn of the 20th century.

However, Maggie S.'s results can be compared with Sharon N., the only speaker in Union City that I interviewed. Her acoustic data and minimal pair tests show that she has a solid merger of /o/ and /oh/, but she was born in 1931. This indicates that the merger spread to Union City in the first few decades of the 20th century.

Next, consider the two DARE speakers from North East, PA, Sarah N. and Nancy S. Their productions of /o/ and /oh/ are summarized in Table 6.6, and vowel plots showing the individual tokens are presented in Figures 6.9 and 6.10.

The Euclidean distances between /o/ and /oh/ for these two speakers from North East are both greater than 300 Hz, and there is little overlap between the two distributions for each speaker.[14] These two speakers from North East thus appear to still maintain a distinc-

---

[14] The token of *caught* displayed in the lower-left corner at (1130, 1777) in Nancy S.'s plot in Figure 6.10 is clearly a measurement error. This is an unfortunate case where the Mahalanobis distance metric for formant

tion between /o/ and /oh/. This evidence contrasts directly with transitional and merged vowel plots presented in Figures 6.5 and 6.6 for Richard O. and Benjamin S., the two speakers from the SWV corpus. These two speakers were born in 1906 and 1907, respectively, one year earlier than Nancy S. However, the two SWV speakers provide evidence that the merger was already spreading to North East in the first decade of the 20th century, while Nancy S. demonstrates that it had not yet spread to all speakers in North East yet at that time. There was thus a period of inter-speaker variation in the phonemic status of /o/ and /oh/ in North East at this time. The fact that the two speakers from this town who appear to be acquiring the merger, Richard O. and Benjamin S., are both male, whereas the two speakers who maintain a distinction, Sarah N. and Nancy S., are both female, suggests that males may have been in the lead in acquiring the merger in North East. A similar pattern in which the males appear to be leading the advance of the merger in Ripley will be presented in Section 6.9.

Table 6.7 displays the values for the two DARE speakers from Warren County. The distributions of /o/ and /oh/ for the older speaker, Agatha S., show some overlap, and the Euclidean distance between the mean values is just under 200 Hz. This suggests that she is transitional with regard to the merger. The other speaker, however, maintains a clear distinction. This inter-speaker variation in the city of Warren is consistent with the conclusion reached in Section 6.5 (based on evidence from two LAMSAS speakers and one speaker from Herold's telephone survey) that the merger spread to Warren County in the first few decades of the 20th century.

The results for the single DARE speaker from Jamestown, NY are displayed in Table 6.8. His acoustic evidence shows that he maintained a distinction between /o/ and /oh/,

---

prediction chose the wrong pair of poles and bandwidths. However, this error was not egregious enough to be excluded by the exclusion criteria described in Section 3.8. Since the methodological goal of this dissertation is to conduct the vowel analyses with no manual intervention in order to enable reproducibility, tokens like these were not removed.

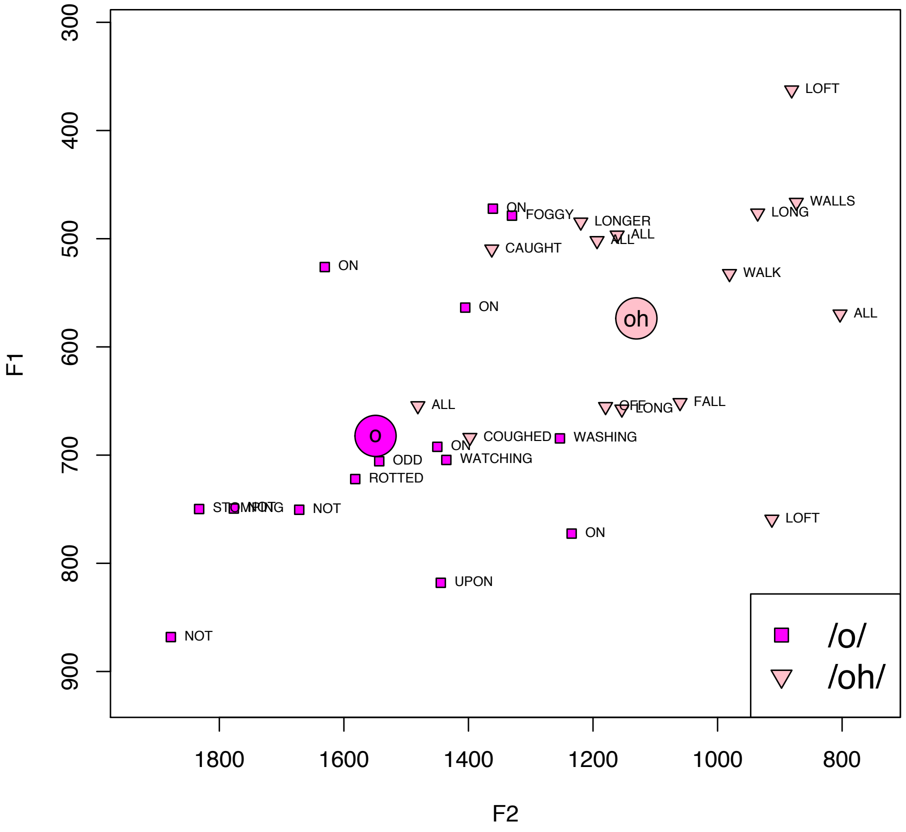

Figure 6.9: /o/ and /oh/ from Sarah N., born 1897 in North East, from the DARE corpus, Mean(/o/) = (682, 1549), N=13; Mean(/oh/) = (765, 1502), N=8; Dist(/o/, /oh/) = 433

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Agatha S.	1907	(757, 1490), N=14	(739, 1292), N=12	199
Steven G.	1915	(734, 1532), N=13	(690, 1128), N=13	406

Table 6.7: /o/ and /oh/ from two DARE speakers from Warren, PA

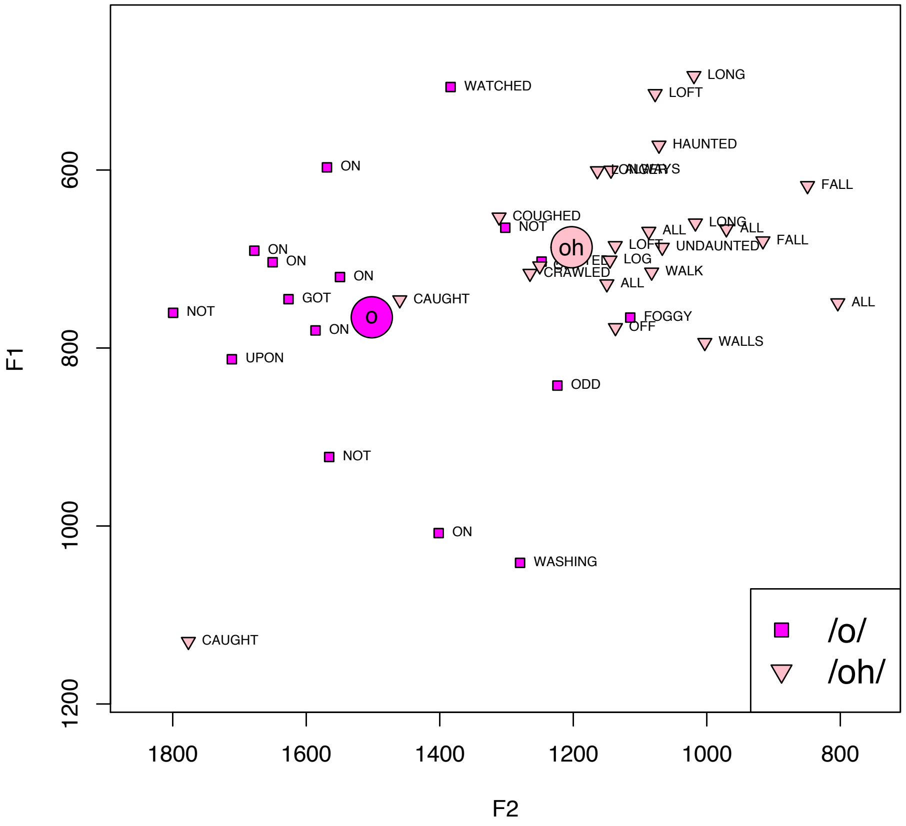

Figure 6.10: /o/ and /oh/ from Nancy S., born 1908 in North East, from the DARE corpus, Mean(/o/) = (765, 1502), N=12; Mean(/oh/) = (687, 1203), N=13; Dist(/o/, /oh/) = 309

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Ted L.	1904	(781, 1477), N=14	(642, 1157), N=11	349

Table 6.8: /o/ and /oh/ from a DARE speaker from Jamestown, NY, born 1904

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Wallace L.	1892	(627, 1250), N=13	(590, 1053), N=11	200
Leslie B.	1897	(849, 1519), N=13	(675, 984), N=13	568
Anne B.	1898	(917, 1619), N=13	(725, 1124), N=12	531

Table 6.9: /o/ and /oh/ from three DARE speakers from Fredonia, NY

as would be expected based on the current status of the vowels in Jamestown (see Section 6.8).

The results for the three DARE speakers from Fredonia, NY are displayed in Table 6.9. The two female speakers, Leslie B. and Anne B., maintain a clear distinction with no overlap between the two distributions, and a large distance of over 500 Hz between the mean values for /o/ and /oh/. The situation is different, however, for Wallace L., who was born a few years earlier. The 200 Hz distance between his mean values for /o/ and /oh/ is smaller than would be expected for a speaker with a clear distinction between the two vowels. To examine his status in more detail, a vowel plot for his individual tokens of /o/ and /oh/ is displayed in Figure 6.11.

There is actually little overlap between the two distributions for Wallace L. The measurement of *not* at (545, 772) is clearly a measurement error, and the position of *foggy* suggests that this lexical item actually contains the phoneme /oh/ for him. It thus appears that he still maintains a distinction between /o/ and /oh/. However, the mean values are much closer than they are for the other two speakers from Ripley; specifically, his mean F2 value for /o/, 1250 Hz, indicates that he has much less fronting of this vowel than most

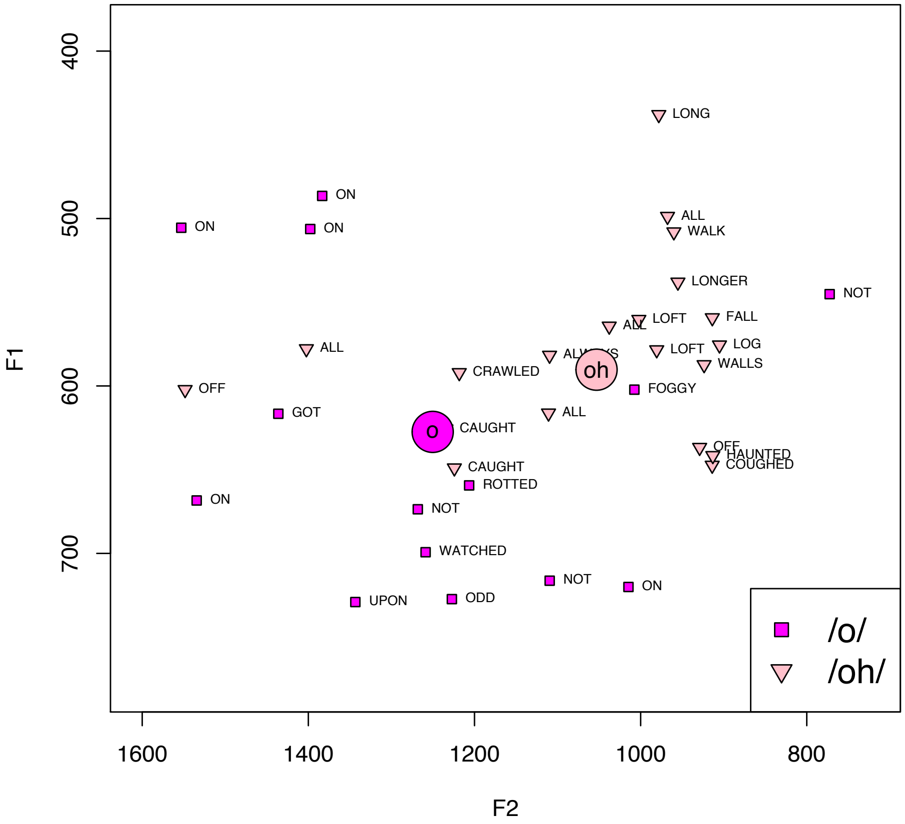

Figure 6.11: /o/ and /oh/ from Wallace L., born 1892 in Fredonia, from the DARE corpus, Mean(/o/) = (627, 1250), N=13; Mean(/oh/) = (590, 1053), N=11; Dist(/o/, /oh/) = 200

other speakers from the region. Based on the one interview I conducted in Fredonia (with a woman born in 1921), it does not appear that Wallace L'.s pattern indicates a community-wide transition to the merger in Fredonia. However, data from younger speakers from the town is necessary to confirm this.

Finally, Table 6.10 displays the results for the three speakers from Ripley, NY. One speaker, Jill C., maintains a clear distinction betwen /o/ and /oh/. Her plot for these two

Name	Year of Birth	Mean(/o/)	Mean(/oh/)	Dist(/o/, /oh/)
Clarence T.	1886	(714, 1417), N=19	(700, 1253), N=14	165
Jill C.	1889	(733, 1542), N=12	(659, 1114), N=13	434
Jonas H.	1898	(746, 1390), N=46	(674, 1249), N=16	158

Table 6.10: /o/ and /oh/ from three DARE speakers from Ripley, NY

vowels is displayed in Figure 6.12. It shows only a slight amount of overlap at the boundary between the two distributions. Furthermore, the distance between the two mean values is quite high, at 434 Hz.

The other two speakers from Ripley, however, do not appear to have a complete distinction between /o/ and /oh/. Clarence T.'s vowel plot in Figure 6.13 does show that the two distributions are mostly separated, and that several of the overlapping tokens are at the edges of the distributions. However, there is a token of *got* that is clearly within the /oh/ distribution,[15] and two tokens of *small* that are clearly within the /o/ distribution. This vowel plot indicates that the distinction between /o/ and /oh/ is not as great for Clarence T. as it originally was in Ripley, based on the evidence from Jill C.

The other DARE speaker from Ripley, Jonas H., indicates even more clearly that /o/ and /oh/ were in transition for him. A large number of tokens from each class overlap with each other, and only the F2 extremes of each distribution remain homogenous. Jonas H. and Clarence T. thus provide an early sign of the transitional nature of Ripley which laid the groundwork for the merger to spread completely through the town in the 20th century (see Section 6.9).

---

[15] The very low F2 value in his pronunciation of *Holland* can be explained by the fact that the /o/ vowel is followed by an /l/. Simiarly, the token of *washing* in the /oh/ distribution can be explained by the preceding /w/ (additionally, it is possible that *washing* has the phoneme /oh/ for Clarence T).

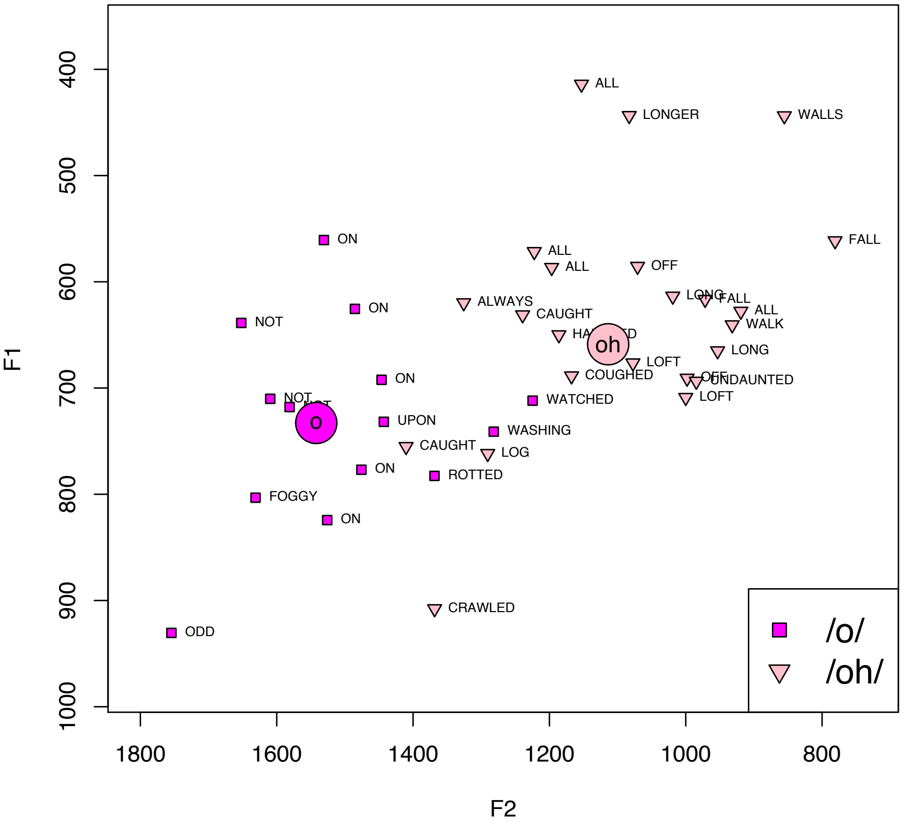

Figure 6.12: /o/ and /oh/ from Jill C., born 1889 in Ripley, from the DARE corpus, Mean(/o/) = (733, 1542), N=12; Mean(/oh/) = (659, 1114), N=13; Dist(/o/, /oh/) = 434

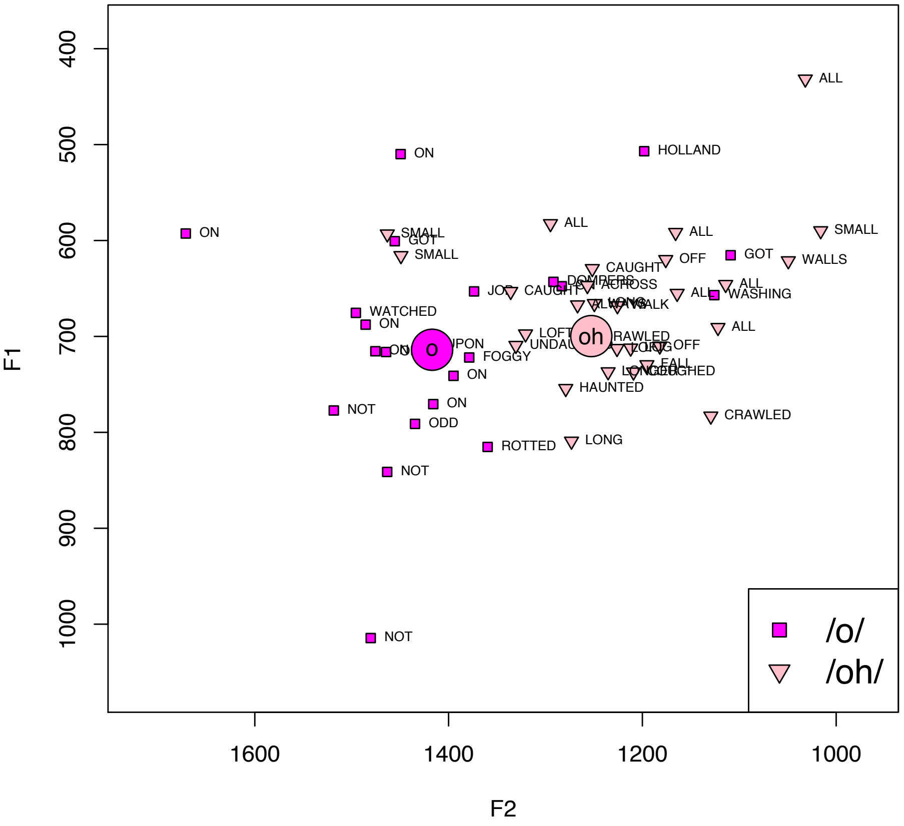

Figure 6.13: /o/ and /oh/ from Clarence T., born 1886 in Ripley, from the DARE corpus, Mean(/o/) = (714, 1417), N=19; Mean(/oh/) = (700, 1253), N=14; Dist(/o/, /oh/) = 165

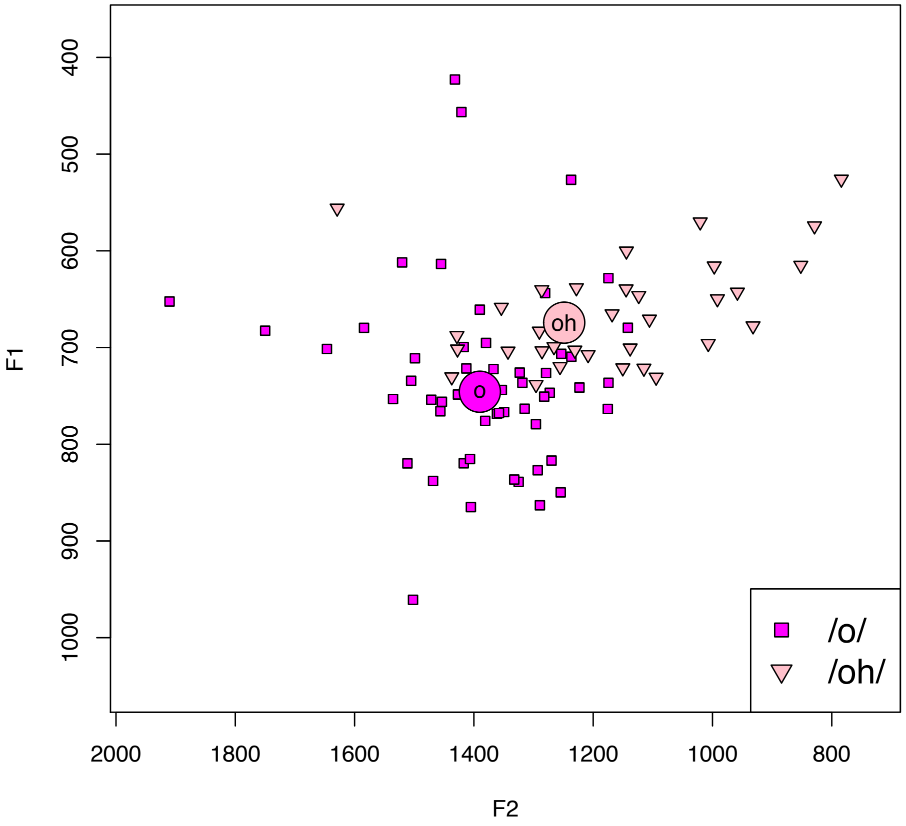

Figure 6.14: /o/ and /oh/ from Jonas H., born 1898 in Ripley, from the DARE corpus, Mean(/o/) = (746, 1390), N=46; Mean(/oh/) = (674, 1249), N=16; Dist(/o/, /oh/) = 158

6.8 The current geographic extent of the merger around Erie

Figures 6.15 and 6.16 show the geographic extent of the merger of /o/ and /oh/ for 72 speakers from Erie and the surrounding areas.[16] These two maps display the production data for the pairs *cot* / *caught* and *Don* / *dawn*: the blue points show speakers who pronounced the two words in a pair identically (based on my perception of their pronunciation), the red points show speakers who pronounced them as clearly distinct, and the green points show speakers who pronounced them similarly, but not identically.

The two maps show a clear boundary between the entire area of western Pennsylvannia stretching from Erie to Pittsburgh, on the one hand, and Chautauqua County, NY, on the other. There is no variation in western Pennsylvania: all speakers are categorically merged. The converse is true for most towns in Chautauqua Co., NY: all speakers maintain a clear distinction between /o/ and /oh/, except for speakers in the town of Ripley, NY. In Ripley, a clear apparent-time distribution of the merger is visible in Figures 6.15 and 6.16. The 19 speakers from Ripley are ordered on the maps by their ages (as are speakers in all towns): the speaker represented by the point in the upper-left corner is the oldest speaker in Ripley, and the speaker in the lower-right corner is the youngest (speakers are arranged in decreasing age order by row). The only speakers from Ripley who pronounced the minimal pairs as either close or distinct are middle-aged and older. No trace of the distinction was found in any of the younger speakers from Ripley. Section 6.9 will analyze the data from the Ripley speakers in more detail, in an attempt to interpret this apparent time distribution.

---

[16] These 72 speakers include the ANAE speakers who were re-analyzed for this dissertation. Their minimal pairs data was obtained from the database file that is released with the ANAE.

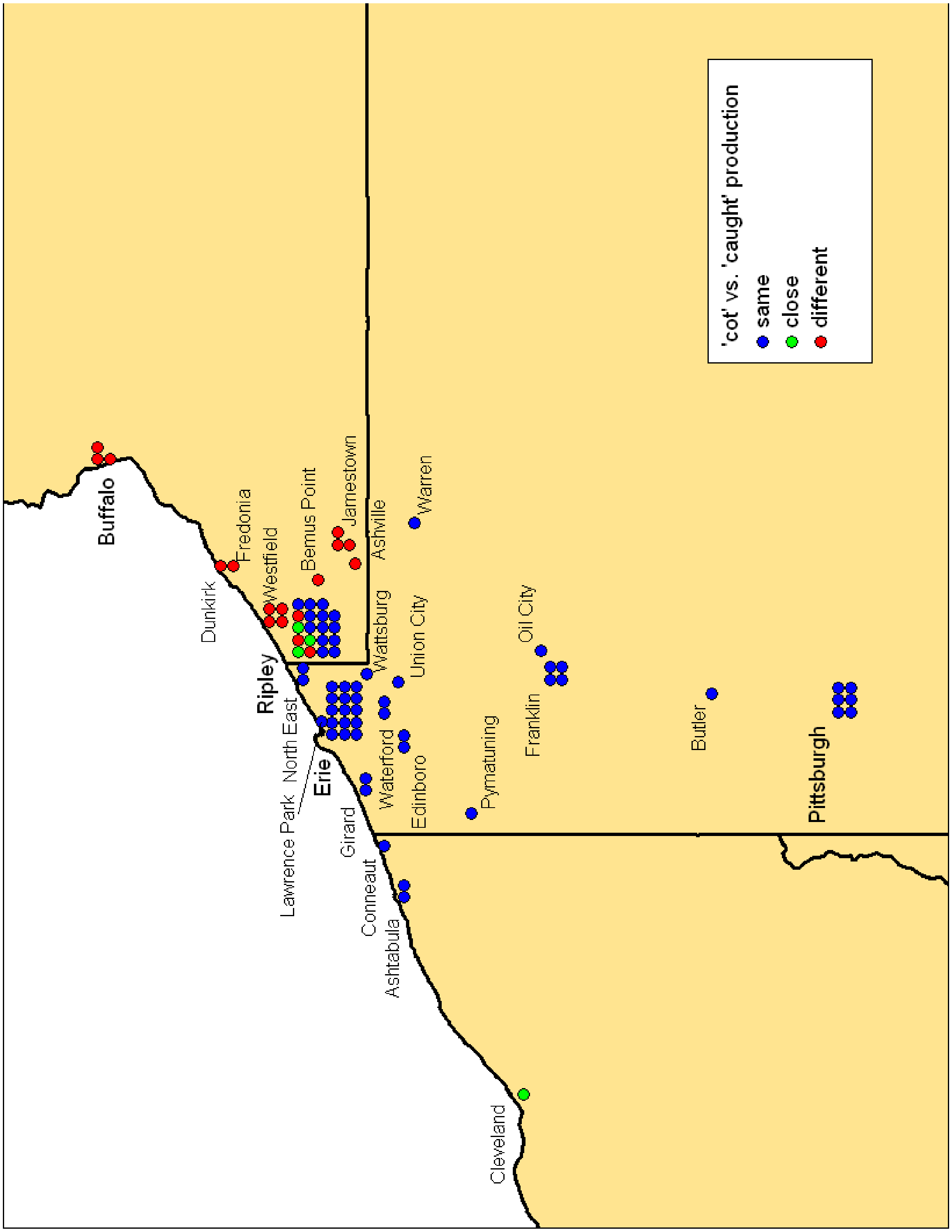

Figure 6.15: Geographic extent of the merger of *cot* and *caught* around Erie, minimal pair production data

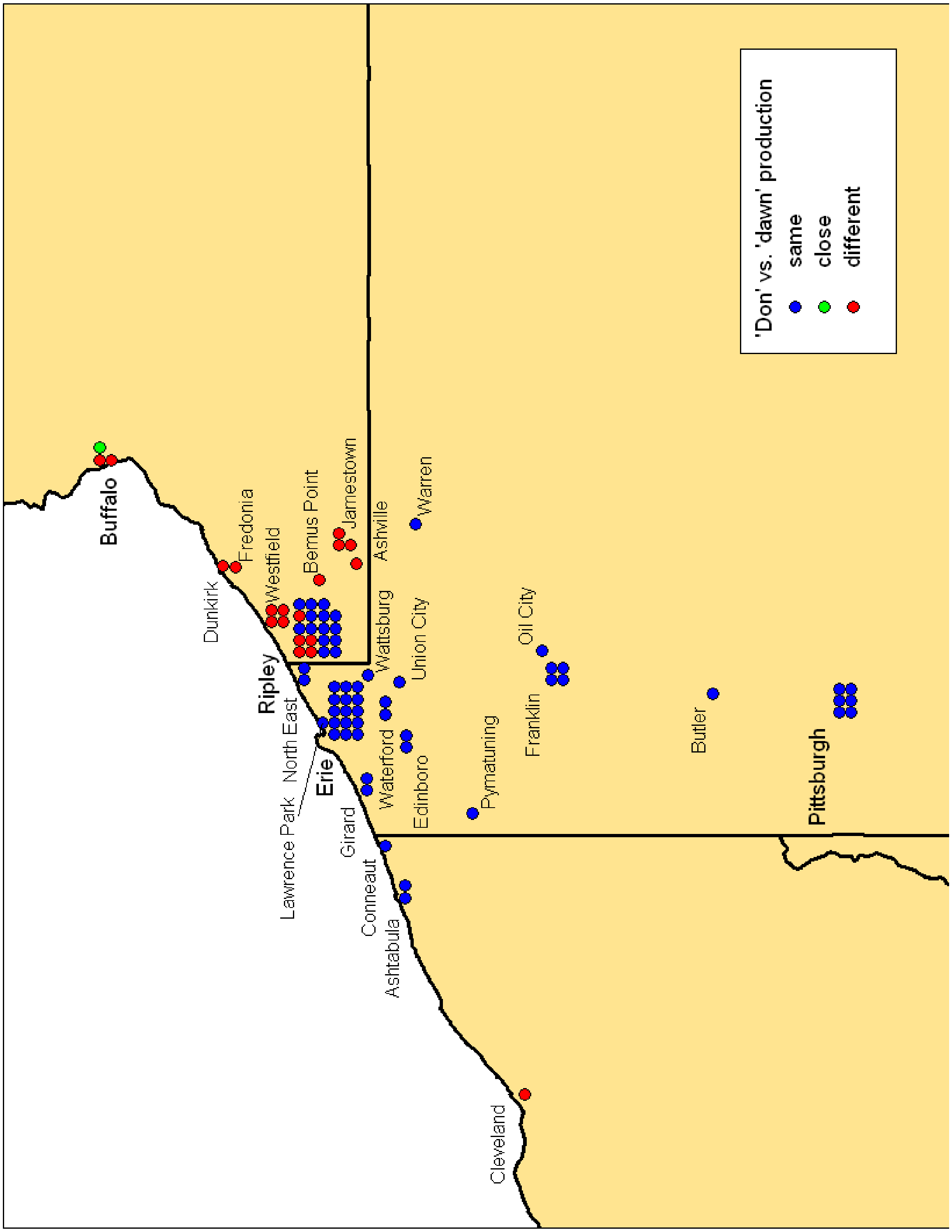

Figure 6.16: Geographic extent of the merger of *Don* and *Dawn* around Erie, minimal pair production data

Figures 6.17 and 6.18 display the perception data for the pairs *cot* / *caught* and *Don* / *dawn*. The general pattern is the same as what was observed for the production data: the merger is present in all towns in western Pennsylvania, and has spread to the younger population of Ripley, NY. Other towns in Chautauqua Co., NY still maintain the distinction.

There are a few cases where there is a mismatch between the speaker's perception of the merger and my evaluation of their production data. In nearly all of these cases, the merger is more advanced in production than in perception. For the *cot* / *caught* pair, there are six speakers for whom the merger is more advanced in production than in perception, and only one speaker for whom the reverse is true.[17] Similarly, there are three speakers for whom the merger of *Don* and *dawn* is more advanced in production, and none for whom the reverse is true. This result goes in the opposite direction to previous findings, where the tendency was for the merger to occur earlier in perception than in production (DiPaolo (1988), Herold (1990:97), Labov (1994:319), Labov et al. (2006:63)[18]). However, the number of speakers involved in the present study is too small to make any reliable generalizations. Furthermore, the perceptual judgments for at least some of the speakers likely represent the influence of orthography. Two of the speakers for whom the merger is more advanced in production than in perception come from areas of western Pennsylvania where the merger almost certainly occurred before the speakers were born. These speakers are a 53-year-old woman from Waterford and a 62-year-old man from Franklin. Since it is very unlikely that either of these two speakers have anything but a total merger of /o/ and /oh/, their perceptual judgments could simply reflect an intrusion of the orthographic difference into their ability to perceive the vowel sounds properly.

---

[17] "More advanced in production" is defined here to mean a production rating of "same" or "close" if the perception rating was "different", or a production rating of "same" if the perception rating was "close".

[18] It should be mentioned, however, that the there was also a large minority of ANAE speakers who showed the opposite pattern, namely the merger occurring earlier in production than in perception.

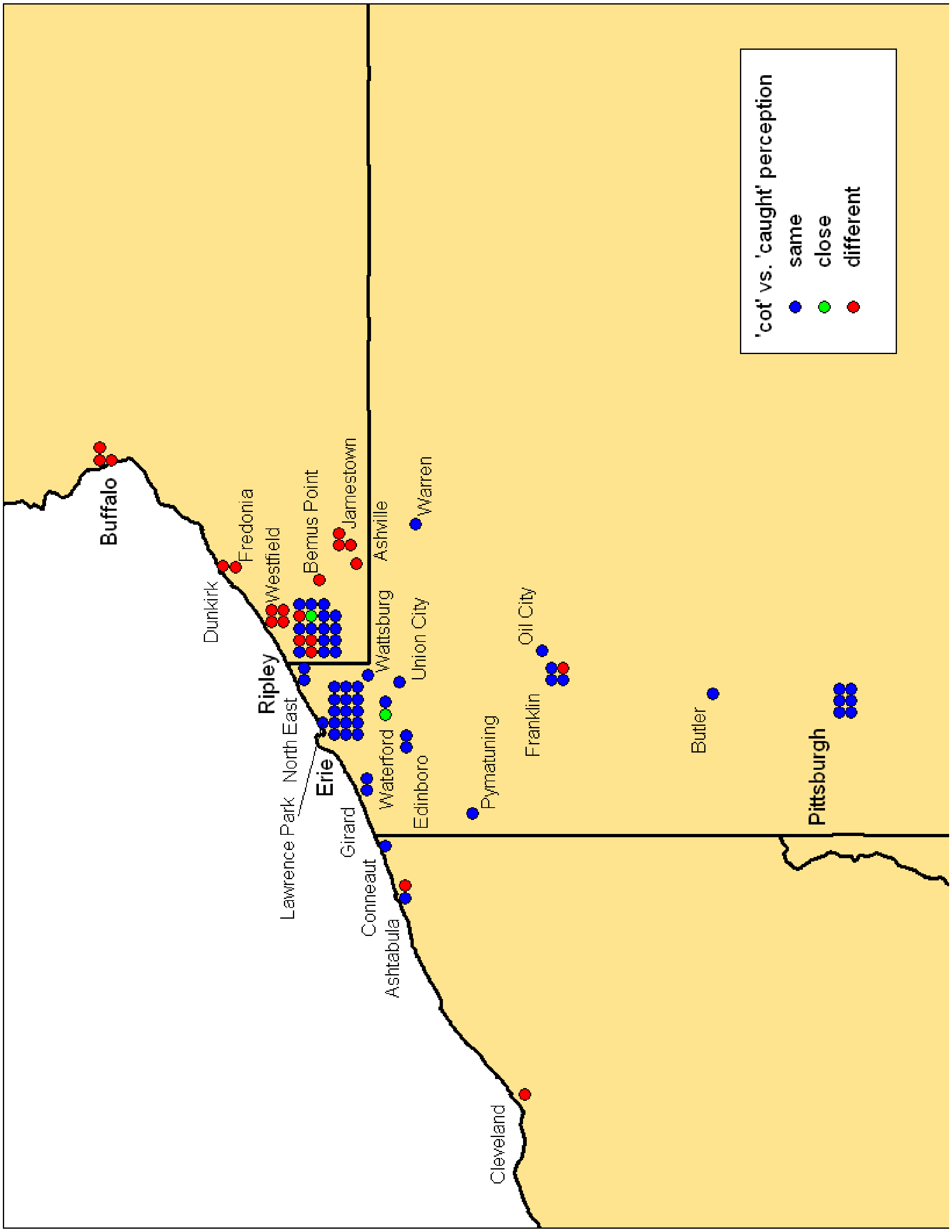

Figure 6.17: Geographic extent of the merger of *Don* and *Dawn* around Erie, minimal pair perception data

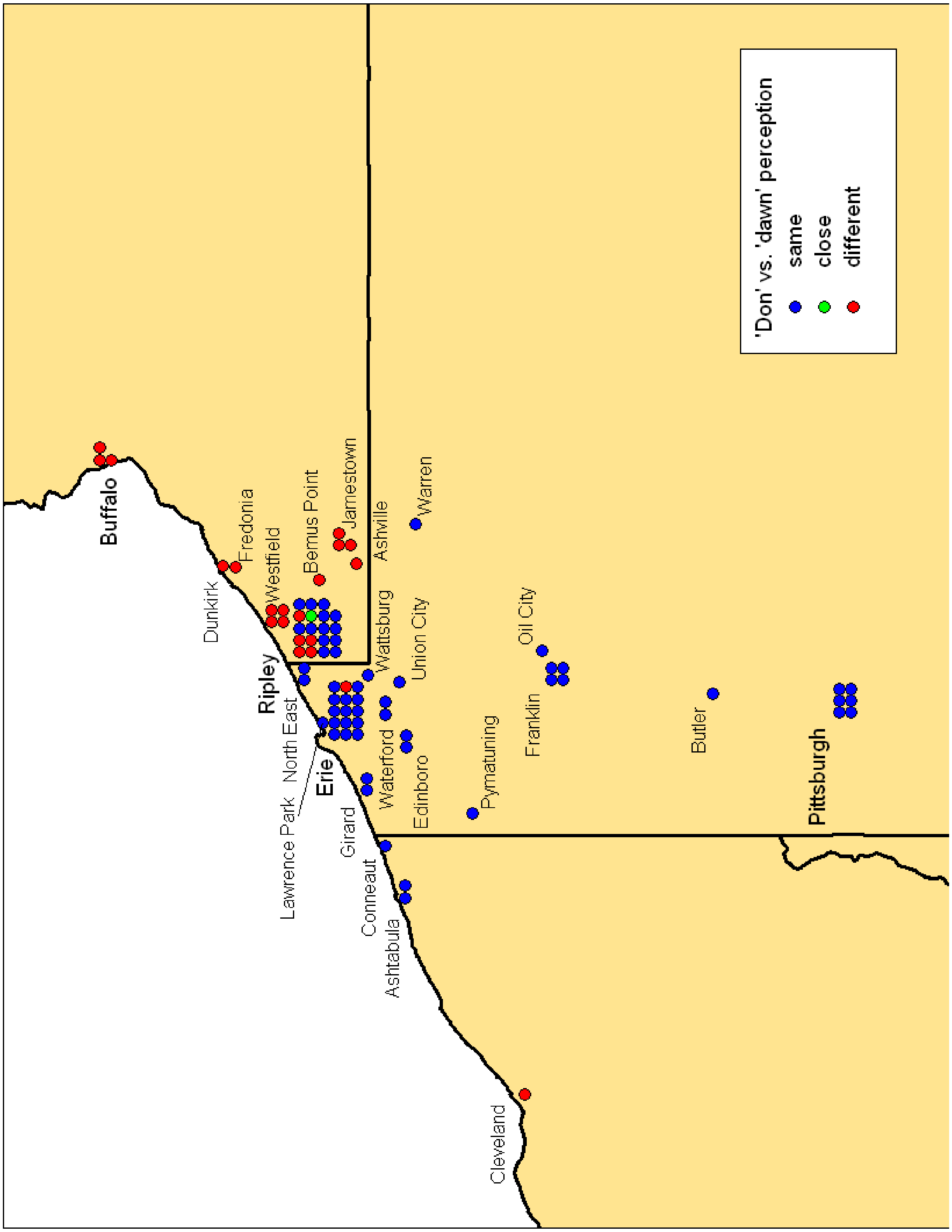

Figure 6.18: Geographic extent of the merger of *Don* and *Dawn* around Erie, minimal pair perception data

Finally, Figures 6.19 through 6.22 present the production and perception results for minimal pair tests for the pairs *collar* / *caller* and *stock* / *stalk*. These two pairs test whether the speaker maintains the /o/ ∼ /oh/ distinction before /l/ and /k/, respectively. The number of speakers in my corpus who did minimal pair tests for these pairs is smaller than the number who did tests for *cot* / *caught* and *Don* / *dawn*, since *collar* / *caller* and *stock* / *stalk* were not added to the list until mid-way through my field work.

However, even with this smaller set of responses the same general geographic pattern for the merger of /o/ and /oh/ is observable. Nearly all speakers in Western Pennsylvania are merged in both production and perception for the two pairs. The only speakers in Pennsylvania whose production tokens were not given the rating "same" are two speakers who produced the pair *collar* / *caller* as slightly distinct and the pair *stock* / *stalk* as different.[19]

---

[19] The speaker from North East pronounced *stalk* (presumably unnaturally) with an /l/, and the speaker from Erie produced *stock* and *stalk* differently during his first reading, but produced them identically when he repeated them.

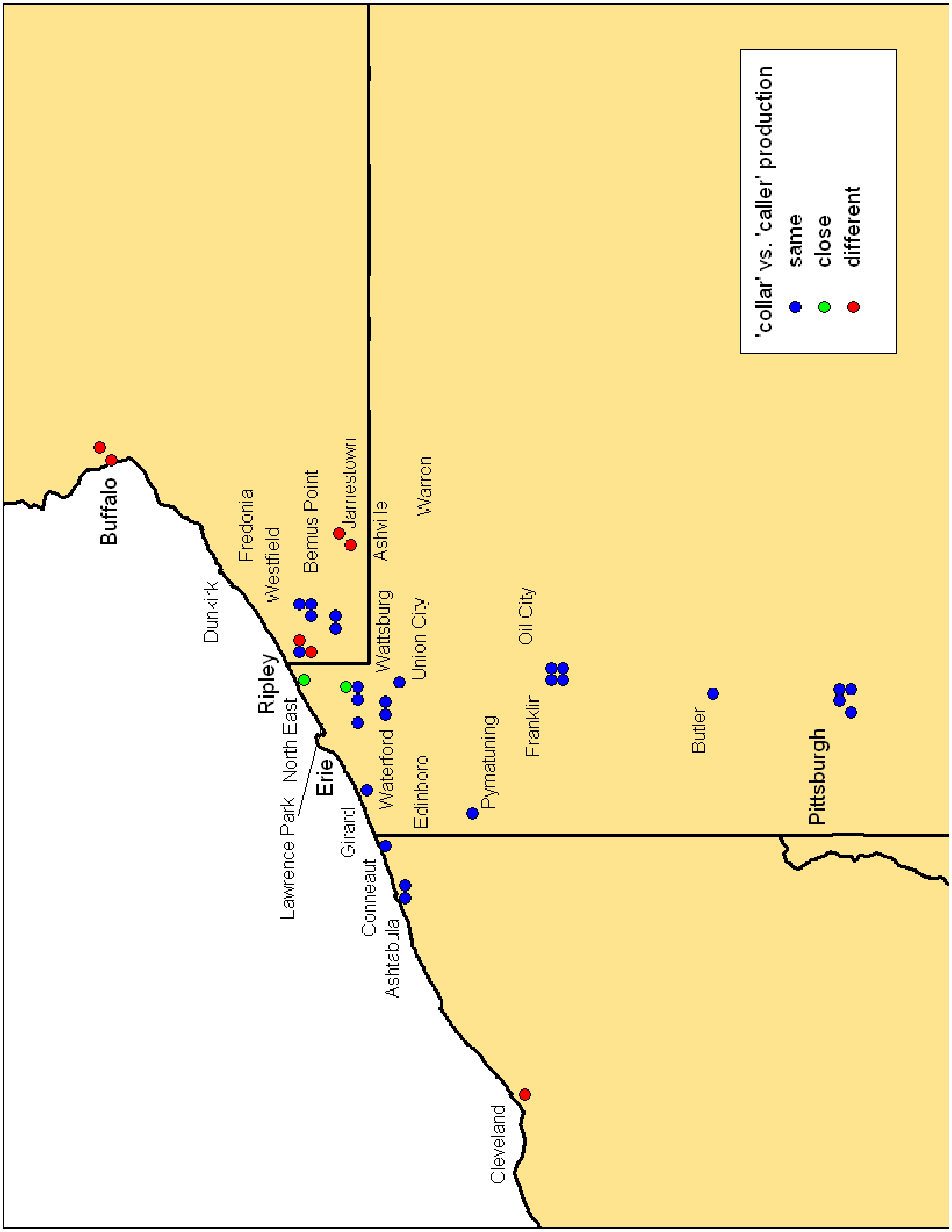

Figure 6.19: Geographic extent of the merger of *collar* and *caller* around Erie, minimal pair production data

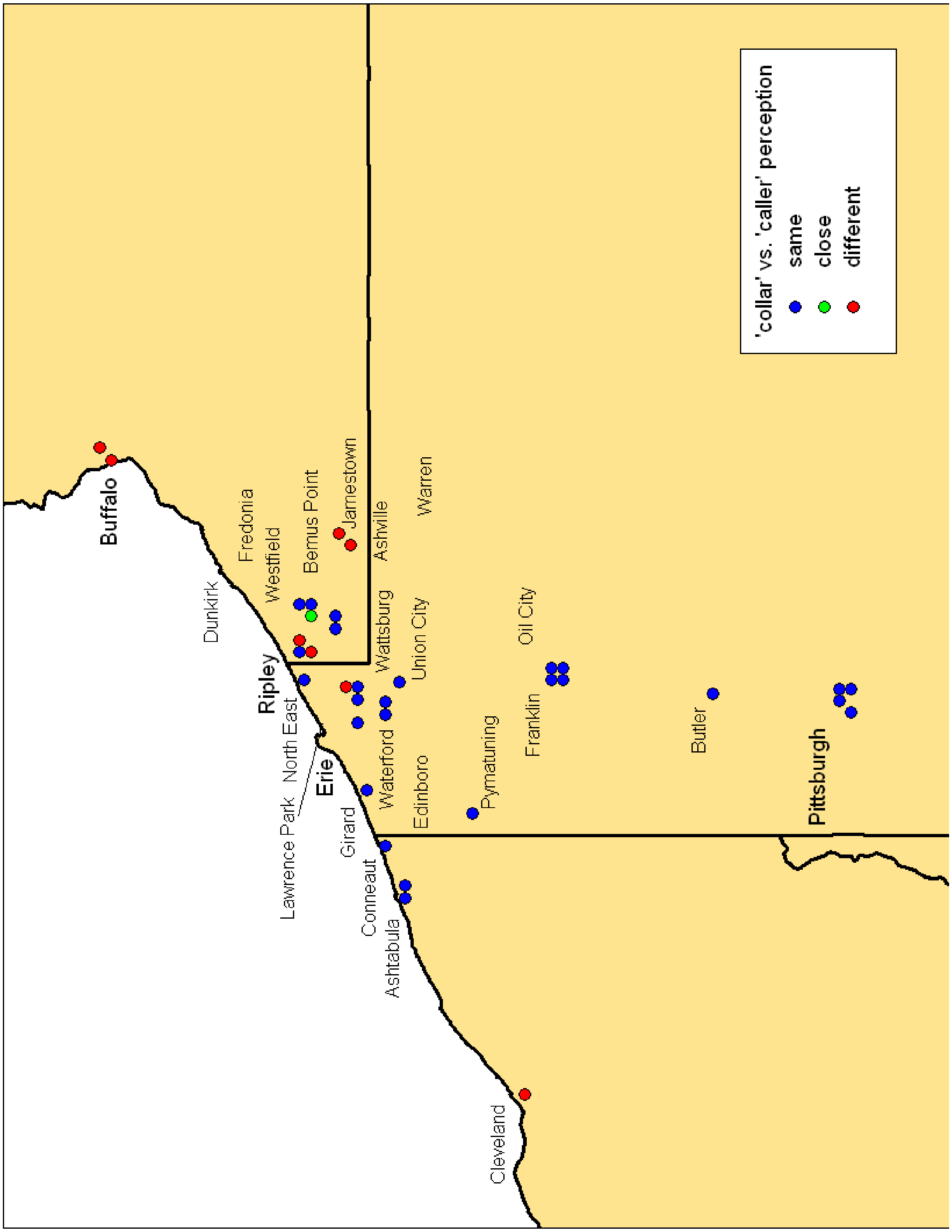

Figure 6.20: Geographic extent of the merger of *collar* and *caller* around Erie, minimal pair perception data

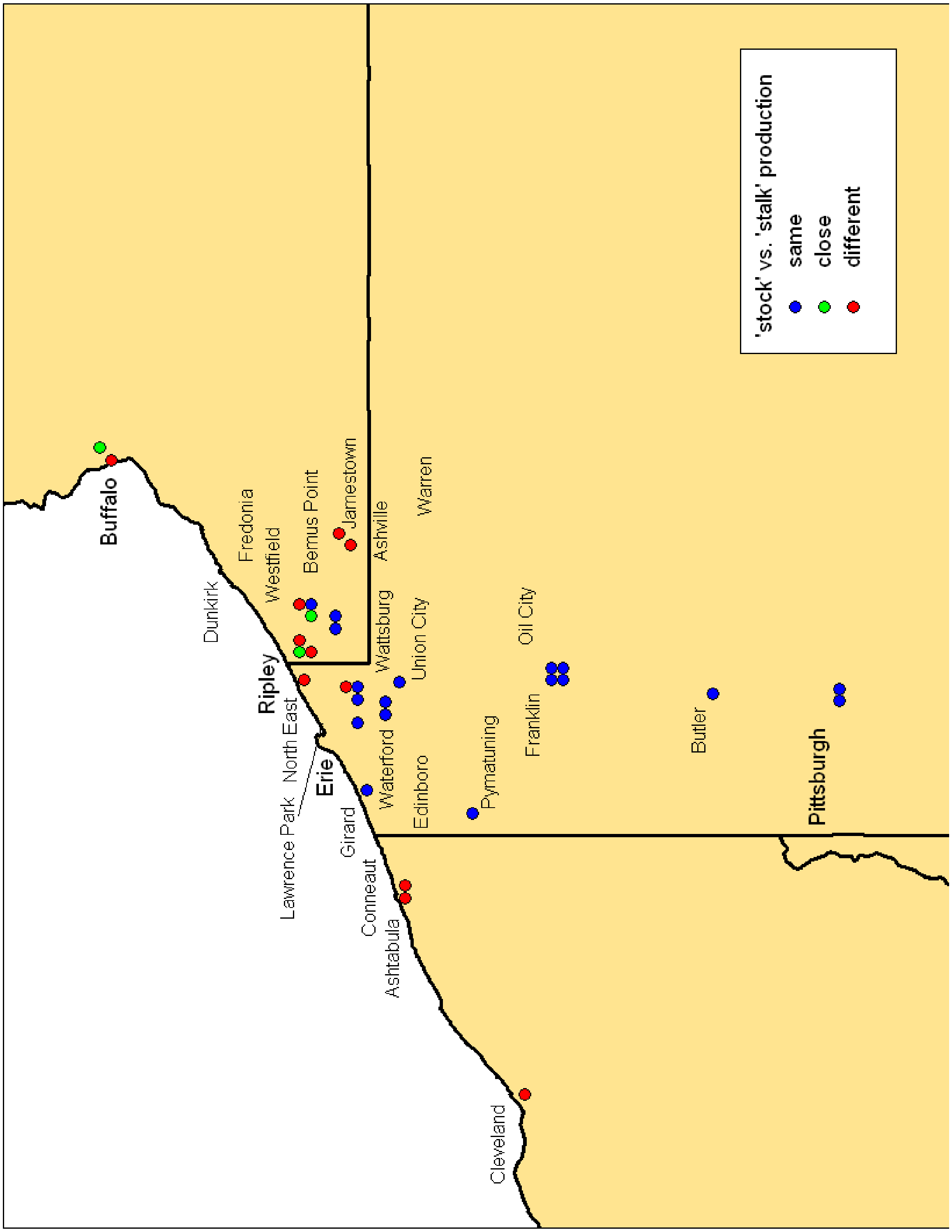

Figure 6.21: Geographic extent of the merger of *stock* and *stalk* around Erie, minimal pair production data

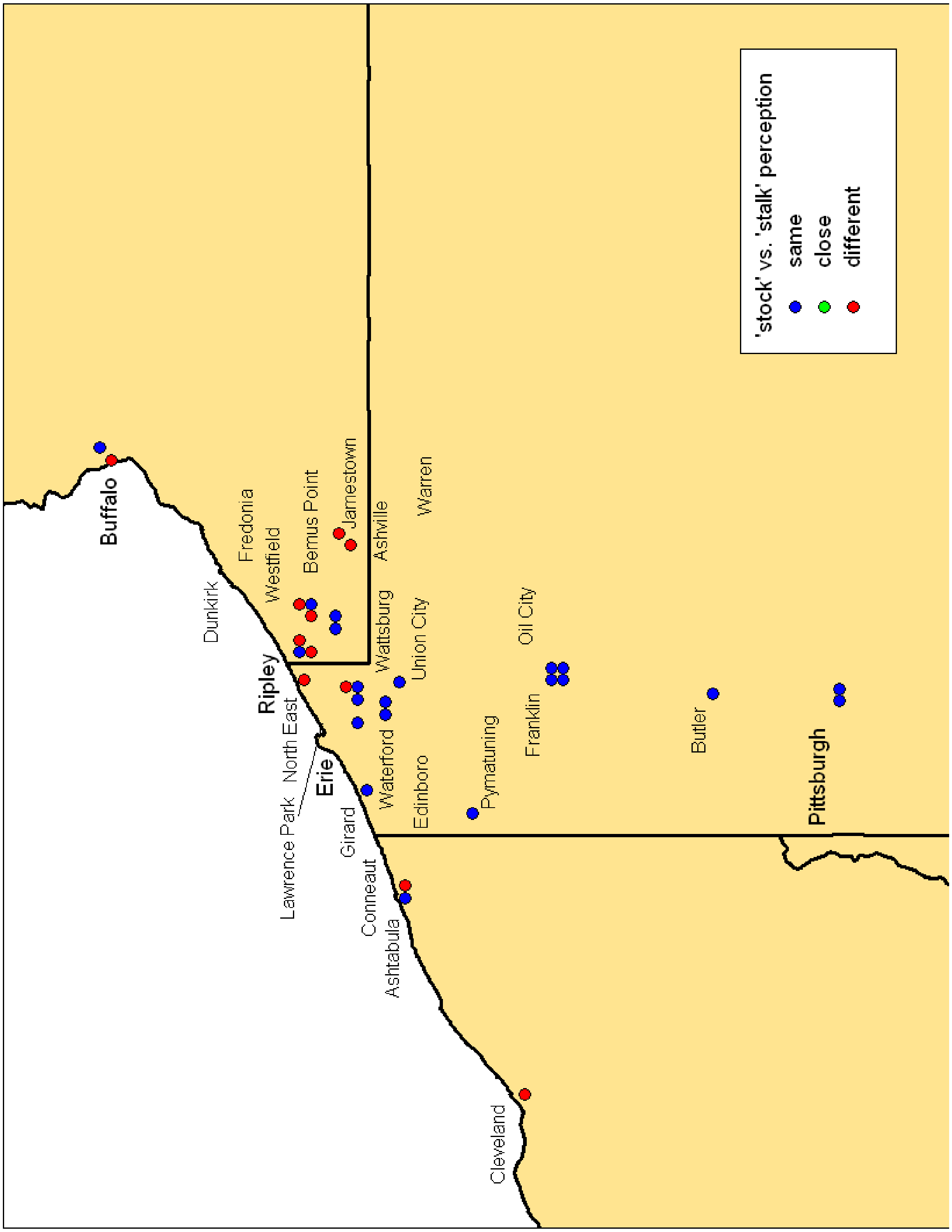

Figure 6.22: Geographic extent of the merger of *stock* and *stalk* around Erie, minimal pair perception data

Another intriguing result that is apparent in Figures 6.15 through 6.22 is that the merger of /o/ and /oh/ has spread to the towns of Conneaut and Ashtabula in northeastern Ohio.[20] Unfortunately, there are no prior studies of these towns to compare this evidence to, but it can be assumed that they were originally unmerged and Northern (as were Erie and Cleveland on either side of them). The mergerd speaker from Conneaut is a 52-year-old woman, and the two merged speakers from Ashtabula are women who were born in the 1930's. Thus, it appears that the merger spread to these two towns quite some time ago. Further research in the region between Ashtabula and Cleveland is necessary to determine the exact location of the current boundary for the merger of *cot* and *caught* in northeastern Ohio.

Finally, Figure 6.23 displays the Euclidean distance between /o/ and /oh/ for all speakers in the corpus. This map shows the same geographic pattern as the maps for the minimal pairs. The only speakers in Pennsylvania who have a distance of greater than 300 Hz are archival speakers from DARE (two from North East, one from Union City, and one from Warren). The town of Ripley displays considerable variation: some speakers maintain a clear distinction with a distances of greater than 300 Hz between /o/ and /oh/, while others appear to have merged the two classes. A more detailed analysis of these speakers from Ripley will be conducted in the next section.

6.9 A case study of the merger in progress: the town of Ripley

As the previous section showed, Ripley is the only town where inter-speaker variation was observed in the results of the minimal pairs tests for /o/ and /oh/. Additionally, Figures

---

[20] The only one of the four minimal pairs that was produced differently was *stock* / *stalk*. Again, the two speakers who produced a difference were apparently confused by the orthography of *stalk* and unnaturally inserted an /l/ into their pronunciations.

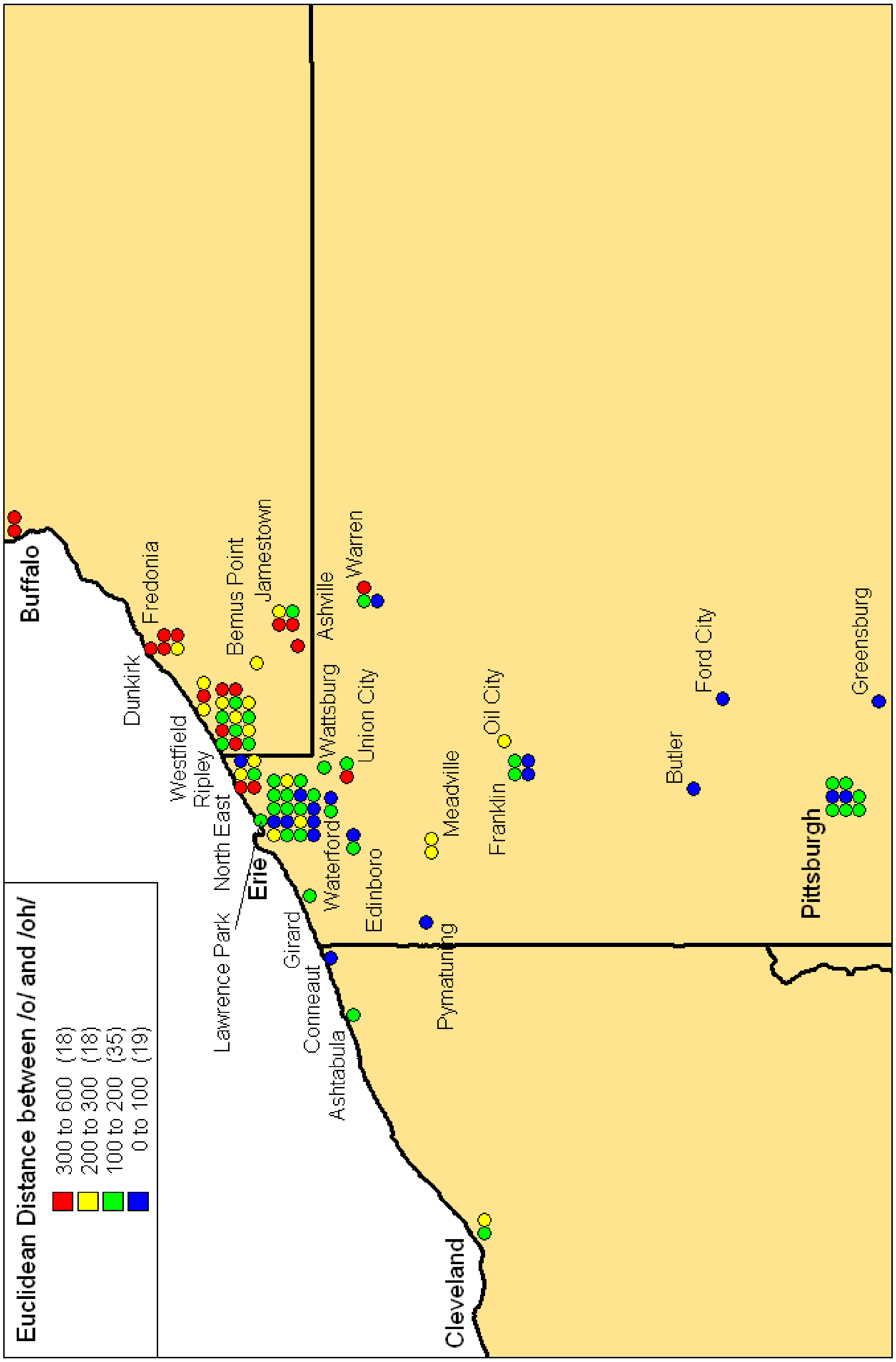

Figure 6.23: Euclidean distance between the mean values of /o/ and /oh/

6.15 and 6.16 clearly show an apparent time distribution in which the merger has become more prevalent in Ripley over time. This section will take a more in-depth look at the results for the individual speakers from Ripley in an attempt to understand how the merger spread to that town.

First of all, the evidence from the young speakers I interviewed in Ripley clearly shows that the merger has progressed to completion for both male and female speakers in Ripley. I conducted abbreviated interviews with five female students and two male students at the high school in Ripley, and their minimal pair tests for /o/ and /oh/ demonstrated a complete merger in production and perception. Furthermore, I conducted full interviews with three other adolescents from Ripley. Both the minimal pair tests and the acoustic data from these interviews demonstrate that these three speakers also have a complete merger of /o/ and /oh/. As an example, Figure 6.24 shows a plot of /o/ and /oh/ for Ryan N., a 15-year-old high school student. This plot shows that Ryan N. has an almost complete overlap between the two distributions, and a rather back /o/ with a mean F2 value of 1266 Hz.

While the adolescents I interviewed in Ripley were categorically merged, there is a large amount of inter-speaker variation among the adults in the town. Based on the minimal pair tests for /o/ and /oh/, the adults I interviewed fall into three categories, defined as follows:

- **Merged speakers**: both the production and perception values for the minimal pairs *cot* / *caught* and *Don* / *dawn* are “same”
- **Unmerged speakers**: both the production and perception values for the minimal pairs *cot* / *caught* and *Don* / *dawn* are “different”
- **Transitional speakers**: the production and perception values for the minimal pairs *cot* / *caught* and *Don* / *dawn* do not unambigiously characterize the speaker as merged or unmerged (there is either a mismatch in production and perception for one of the

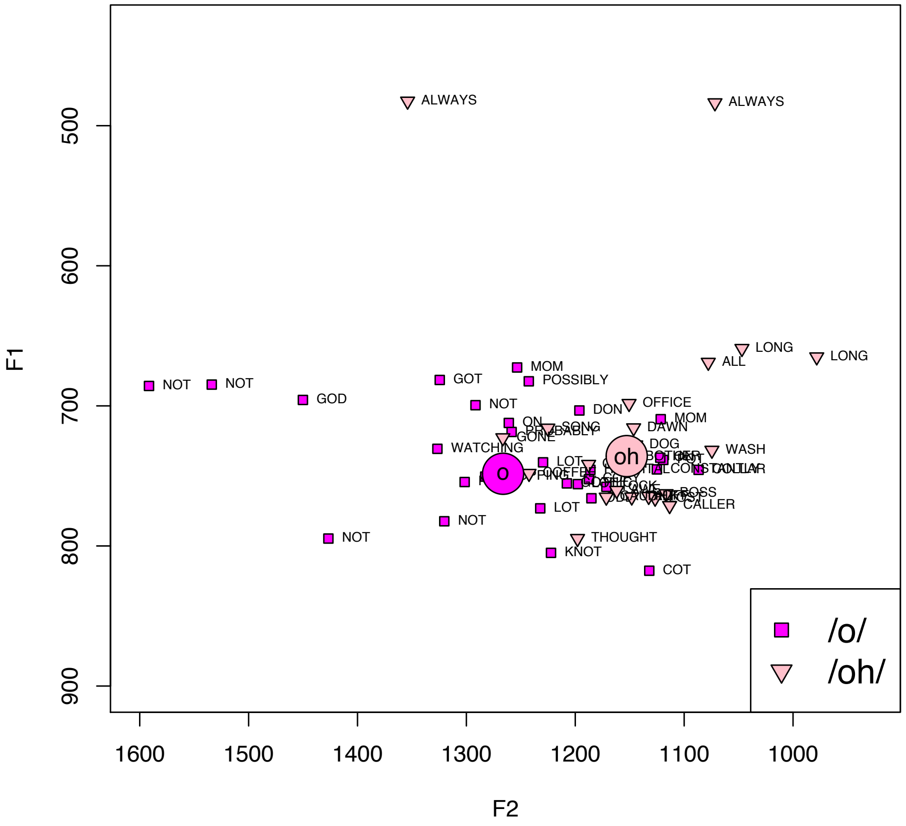

Figure 6.24: /o/ and /oh/ from Ryan N., born 1994 in Ripley
Mean(/o/) = (748, 1266), N=30; Mean(/oh/) = (736, 1153), N=16; Dist(/o/, /oh/) = 114

Name	Born	Occupation
Tracy N.	1972	waitress
Pam O.	1958	winery owner

Table 6.11: Demographic information for two adult speakers in Ripley who have the merger of /o/ and /oh/

Name	Born	Occupation
Sheila T.	1950	waitress
Rachel A.	1951	daycare provider
John M.	1953	town supervisor
Daphne R.	1958	grape farmer
Jane L.	1960	waitress
Rachel C.	1963	town clerk

Table 6.12: Demographic information for six unmerged adult speakers in Ripley

minimal pairs, or the two pairs produced different results)

Tables 6.11 through 6.13 display the demographic information for the speakers from Ripley that fall into each of these three categories.

The lists of speakers in Tables 6.11 – 6.13 suggest that the merger was in progress in Ripley at least 60 years ago. The ages of the speakers in each group suggest that the merger advanced more quickly among men, although the number of speakers is too small to say this with certainty. The data from the three DARE speakers from Ripley would fit well

Name	Born	Occupation
Stan R.	1948	grape farmer
Larry K.	1952	town supervisor
Rebecca R.	1980	baker

Table 6.13: Demographic information for three transitional adult speakers in Ripley

with this observation. The two male DARE speakers from Ripley (born in 1886 and 1898) appeared to already be in transition to the merger, while the female speaker (born in 1889) maintained a clear distinction. Thus, there appears to have been a period of nearly 100 years during which the merger was slowly spreading throughout the town of Ripley.

6.10 Explaining the chronology of the merger

The original starting point for the chronology of the merger was LAMSAS speaker PA67b from Amity township in Erie County. He was born in 1903, and maintained a clear distinction between /o/ and /oh/, according to Kurath and McDavid (1961) and Wetmore (1959). This suggested that the merger did not reach Erie until around the second decade of the 20th century, at the earliest. However, the apparent time date from the elderly speakers from Sun Valley demonstrated that a complete merger had spread through the city of Erie by the 1920's. Additionally, three archival speakers (the two from the SWV corpus and H.O. Hirt) suggest that the merger was in transition in the city of Erie and the neighboring town of North East by the turn of the 20th century.

The apparent contradiction between the LAMSAS data and the other evidence can be explained by considering the specific location in Erie County of the two unmerged LAMSAS speakers. Neither of them were from the city of Erie itself; rather they were born and raised in small farming communities in the southeastern part of Erie County. On the other hand, the speakers who provide evidence for an earlier date for the merger are much more connected to the city than the LAMSAS speakers: H.O. Hirt and the Sun Valley residents are all from the city of Erie itself, and the two SWV speakers are from North East. North East is only slightly closer to Erie than Amity in terms of distance, but is much more closely connected with Erie, since it is a larger community and a major highway passes between Erie and North East.

So, if all of the temporal and geographic evidence is taken at face value, then it indicates that the merger first occurred in the city of Erie, and then spread gradually to the nearby townships in Erie County. The spread of the merger proceeded in accordance with the Cascade Model of Labov (2003), first to the more populous ones, then, finally, to the smaller, more isolated ones. H.O. Hirt's data indicates that the merger probably took place in Erie already before the turn of the 20th century. The two SWV speakers indicate that it had spread to North East by around 1910. Finally LAMSAS speaker PA67b indicates that the merger had not yet reached Amity township by 1910.

This Cascade Model pattern of the merger spreading to the larger cities in a county first and from there to the smaller towns appears to be applicable to the other counties of northwestern Pennsylvania as well, although the data for Warren and Crawford Counties is not as clear. Table 6.7 showed that a female speaker born in 1907 in the city of Warren was transitional, and Figure 6.3 showed that the two LAMSAS speakers from rural areas of Warren County were unmerged. These two speakers, however, were born one and two generations earlier than the DARE speaker from the city of Warren. Thus, a Cascade Model spread of the merger to Warren County is not contradicted by the evidence from these three speakers, but a more geographically continuous model also cannot be ruled out by the dates. Finally, the oldest DARE speaker from Meadville, the largest city in Crawford County, also appeared to be transitional (see Table 6.4). She was born in 1899. The mergerd LAMSAS speaker in Crawford County was from the rural town of East Fallowfield, in the southern part of the county, and was born in 1859. Again, it is possible for the Cascade Model to explain this situation (assuming the merger was in transition in Meadville for two generations); however, a model in which the merger spread monotonically from the South to the North would also apply.

The only county which is a clear counter-example to the Cascade Model for the spread of the merger is Chautauqua County, NY. In that county, the merger has spread to Ripley, a small farming town, but none of the other larger towns. In this case, proximity to the large city of Erie seems to be the dominant factor. Ripley is just across the state line from North East, PA, and is only about 25 minutes away from Erie. Residents of Ripley are much more connected to Erie than to other cities in Chautauqua County, such as Jamestown. Furthermore, the other cities and towns in Chautauqua County have more structural resistence to the merger of /o/ and /oh/, because /o/ is more strongly fronted there than in Ripley. This aspect of the spread of the merger will be explored more in Section 7.3.2.

Chapter 7

Other Phonological Features

7.1 Introduction

The most important phonological feature of Erie from a dialectological standpoint is the merger of /o/ and /oh/. The previous chapter described the status of this merger in the region, and showed that Erie speakers are categorically Midland with respect to this feature. This chapter will consider the other phonological features that display regional and age-based differences for the speakers in my corpus.

First, Section 7.2 will consider the system of back upgliding vowels, which are shown in the ANAE to have clearly distinct distributions in the North and the Midland. Next, Section 7.3 will consider the structurally related changes of the Northern Cities Shift and how they relate to the spread of the merger of /o/ and /oh/ discussed in the previous chapter. Finally, Section 7.4 will present results for two other phonological variables that provide apparent time evidence for change among the speakers in my corpus.

7.2 Back upgliding vowels

This section examines the system of back upgliding vowels: /uw/, /ow/, and /aw/. These three vowels are undergoing fronting in many dialects of North American English; however, the progress of the change in F2 and the effect of segmental environment depends both on the region and the specific vowel (see Labov et al. (2006:152–168) for a complete description of how the behavior of the back upgliding vowels varies by region). In general, the fronting of /uw/ after coronals is most widespread,[1] and the fronting of /aw/ is most advanced.[2]

The division between the North and the Midland is quite strong with respect to the behavior of the back upgliding vowels. The difference is especially strong for /ow/, where the Midland has the strongest and most consistent fronting, and the North has very little. The fronting of /uw/ after non-coronals is also much more advanced in the Midland than in the North. This section will show that the speakers from Erie, in general, display Northern patterns in the back upgliding system. This systematic behavior is in marked contrast to the categorical Midland behavior of the Erie speakers with regard to the merger of /o/ and /oh/ shown in Chapter 6.

7.2.1 /uw/

The ANAE shows that the degree of /uw/ fronting depends heavily on the identity of the syllable onset. Specifically, /uw/ after coronal onsets is fronted (with an F2 mean value greater than the F2 midline of 1550 Hz) for nearly all speakers in North America, while fronting after non-coronal onsets is not as extreme, and limited to certain dialect regions.

---

[1] Only two dialects do not have strong fronting of /uw/ after coronals: in the North there is moderate fronting, and only in Eastern New England is /uw/ still a back vowel after coronals.

[2] "Advanced" is used here in the sense that the fronting of /aw/ has reached its maximum value and is receding. Only three regions show a significant age effect for /aw/ (Mid-Atlantic, South, and North), and these effects are all positive, i.e., the apparent time distribution shows that younger speakers have less fronting than older speakers in these regions.

When discussing this allophonic variation, tokens of /uw/ occurring after coronal onsets will be referred to with the notation /Tuw/, and tokens occurring after non-coronals with /Kuw/. In addition, a following /l/ has a uniform backing effect in the North and the Midland (only some Southern speakers also have fronting of /uw/ before /l/), so tokens before /l/ will not be discussed in this section.

Figure 12.2 in the ANAE shows that the mean F2 value of /Tuw/ for all speakers in North America is 1811 Hz, and Figure 12.3 shows that the mean F2 value of /Kuw/ is 1433 Hz. Additionally, Table 12.1 displays the results of a linear regression analysis for F2 of /uw/, and by far the strongest segmental effect is the presence of a coronal onset.

These general findings are reproduced for the speakers in my corpus as a whole. The mean F2 value of /Tuw/ is 1758 Hz (N = 2,032), and the mean F2 value of /Kuw/ is 1265 Hz (N = 639). The relative positions of the two allophones in my corpus are thus similar to the ANAE data. However, the absolute mean values are slightly lower, reflecting the fact that the region included in my corpus does not include speakers with the most extreme fronting of /uw/.

A linear regression analysis for F2 of /uw/ was conducted to determine the effects of the segmental environment on /uw/ fronting. This analysis takes into consideration the following features:[3]

- **Preceding segment**: oral labial, nasal labial, oral apical, nasal apical, palatal, velar, liquid
- **Following segment manner of articulation**: stop, affricate, fricative, nasal
- **Following segment place of articulation**: labial, labiodental, interdental, apical, palatal, velar

---

[3] These features represent the codes defined by the Plotnik program for the preceding and following segmental environments. Codes for the tokens in the environments that were excluded based on the criteria in Section 3.8 are not included in the regression analyses in this chapter.

Feature	Coefficient	Probability
Intercept	1228	< 0.0001
Preceding oral apical	539	< 0.0001
Preceding nasal apical	447	< 0.0001
Preceding palatal	497	< 0.0001
Preceding liquid	503	< 0.0001
Following interdental	-298	< 0.05
Following apical	-82	< 0.05

Table 7.1: Linear regression coefficients for F2 of /uw/ for all environmental features significant at $\alpha = 0.05$ (N = 2,671)

- **Following segment voicing**: voiceless, voiced
- **Following sequences**: one following syllable, two following syllables, complex coda, complex coda and one following syllable, complex coda and two following syllables

The results of the regression analysis are shown in Table 7.1. The four significant onset codes are all from coronal consonants, and they all show a strong fronting effect of about 500 Hz. This is similar to the 480 Hz coefficient shown for coronal onsets in Table 12.1 in the ANAE. In a departure from the ANAE results, no other onset consonants have a significant effect on the fronting of /uw/ (Table 12.1 shows significant effects for velar, obstruent+liquid, labial, and nasal onsets for the ANAE data). The effects of the following segment on /uw/ also differ somewhat from the observations in the ANAE. Table 7.1 shows a strong backing effect when /uw/ is followed by an interdental (in practice, this simply means the phoneme /θ/ since the phoneme sequence /uð/ does not occur in my corpus). This specific effect was not observed in the ANAE; however, they do note that following fricatives have a significant backing effect of -137 Hz. In addition, Table 7.1 shows a negative coefficient for following apicals, whereas the analysis for the ANAE data shows a positive coefficient for following coronals. However, the effects are not large in either study: -82 Hz and 70 Hz, respectively.

As an example of the main allophonic effect of coronal onsets on the fronting of /uw/, Figure 7.1 shows all tokens of /uw/ for Rachel A., from Ripley, NY, with separate symbols for tokens of /Kuw/ and /Tuw/.[4] Her mean F2 value of /Kuw/ is quite far back at 1108 Hz. Her fronting of /Tuw/ is not as extreme as it is for some speakers, but at 1685 Hz it is well in front of the mid line. There is a fair amount of overlap between the two classes, mostly caused by several fronted tokens of /Kuw/. There are very few non-fronted tokens of /Tuw/.

The ANAE shows that the Midland and the North clearly behave differently with respect to the fronting of /uw/. The mean values of /uw/ by region in Table 12.2 in the ANAE show the two regions occupying opposite ends of the continuum: the Midland has the largest amount of /uw/ fronting, with a mean F2 value of 1713 Hz, whereas the North has the backest /uw/, with a mean F2 of 1359 Hz. However, based on the more detailed regional results from the ANAE, no differentiation among the speakers in my corpus would be expected for the /Tuw/ allophone. While it is true that the Midland and the North are divided in their behavior with respect to /Tuw/, the division does not cut across the entire region (see ANAE Map 12.1 for the basis for the following discussion). Specifically, many Midland and Northern speakers behave similarly and follow the general pattern in North America of moderate fronting of /Tuw/ (defined as having an F2 mean value between 1550 Hz and 2000 Hz). Areas with extreme fronting are sometimes found in the Midland (for example, in the cities of Indianapolis and Kansas City); however, this isogloss excludes most of the Midland. Most importantly for this study, none of the Pittsburgh speakers are contained within it—the mean value for F2 of /uw/ in Table 12.2 in the ANAE shows a value of 1529 Hz for Western Pennsylvania. On the other hand, the North does contain

---

[4] The vowel means for /iyC/ and /o/ are also displayed as reference points for her vowel system

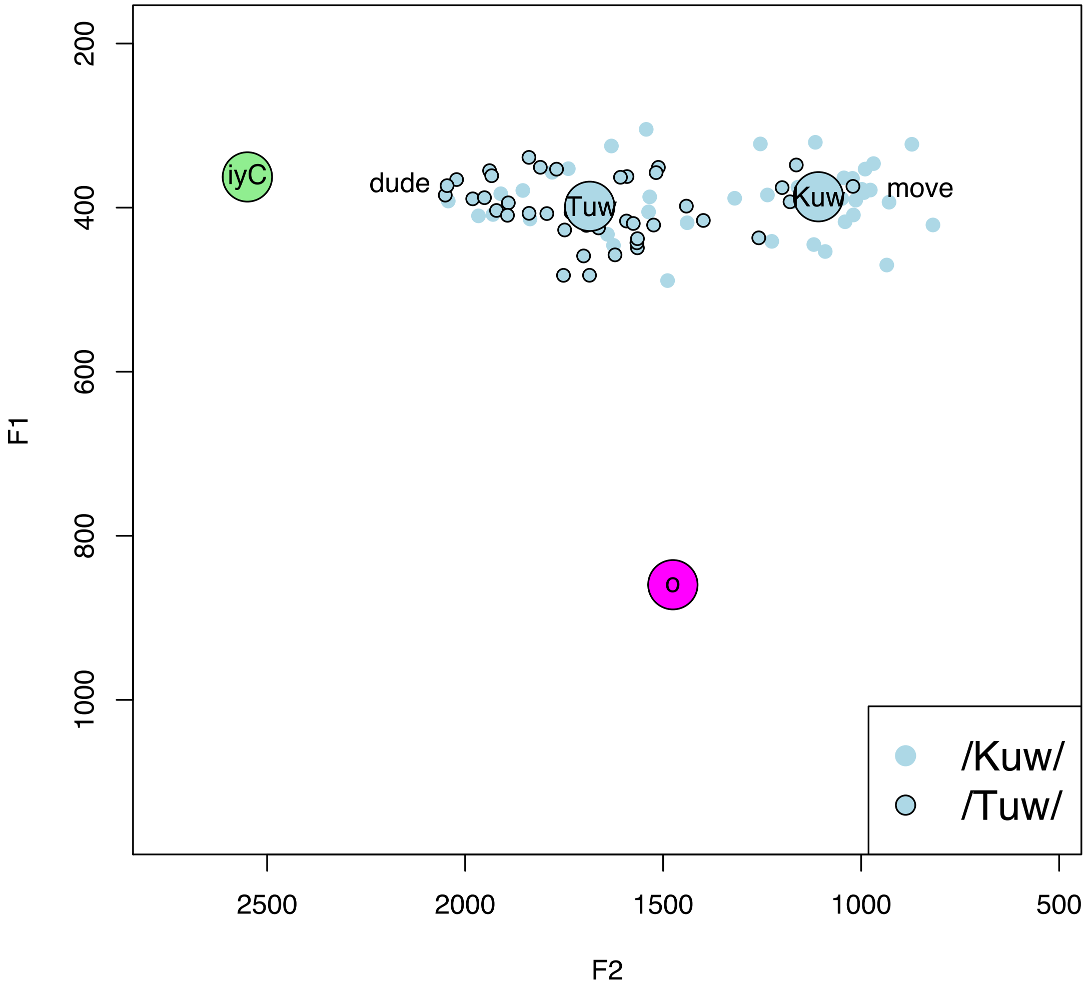

Figure 7.1: /Kuw/ and /Tuw/ for Rachel A., born 1951 in Ripley,
Mean(/Kuw/) = (387, 1108), N=20; Mean(/Tuw/) = (398, 1685), N=45

a large number of speakers with non-fronted /Tuw/; however, most of them are located in the western part of the region in Wisconsin and Minnesota. The results for the Northern speakers in the cities closest to Erie also mostly show the general pattern of moderate fronting of /Tuw/.

Figure 7.2 confirms this expectation for the speakers in my corpus. Nearly all of them have moderate fronting of /Tuw/ with an F2 value between 1550 and 2000 Hz. The speakers that have a mean value for /Tuw/ less than 1550 Hz are not mostly located in the North; neither are the speakers with a mean value for /Tuw/ greater than 2000 Hz mostly located in the Midland. Indeed, no clear regional pattern for these two groups of speakers with more extreme behavior of /Tuw/ is apparent.

The results for /Kuw/ from the ANAE, however, suggest a different regional pattern for the speakers in my corpus. ANAE Map 12.2 shows that the differentation between the Midland and the North with respect to /Kuw/ is much stronger than for /Tuw/. The entire Northern region is contained within the isogloss showing back values of /Kuw/, whereas most Midland speakers have either moderate or strong fronting of /Kuw/.[5]

Figure 7.3 shows the geographic distribution of speakers from my corpus with respect to the fronting of /Kuw/.[6] In this case, a clear regional pattern is observable: most of the speakers from Western Pennsylvania south of Erie County have moderate to strong fronting of /Kuw/, whereas most of the speakers from Erie County and New York show no fronting. The city of Erie clearly patterns together with the North with respect to this variable, and

---

[5] The isogloss for back values of /Kuw/, defined as speakers having a mean F2 of /Kuw/ less than 1200 Hz, cuts through the region between Pittsburgh and Erie. However, it is not clear why Erie was included inside this isogloss. Neither of the two Erie speakers match the selection criterion for this isogloss, and Erie is not entirely surrounded by communities inside the isogloss (due to the fact that Pittsburgh is also not included in it). So, according to the ANAE's isogloss construction procedure (Labov et al. 2006:42), they should be outside this isogloss.

[6] The displayed values of the feature are slightly different from those in the ANAE. The F2 threshold for the highest range of /Kuw/ fronting was set at 1450 Hz, 100 Hz lower than the 1550 Hz threshold displayed in Map 12.2 in the ANAE. This adjustment was necessary to produce a clear contrast between the Northern and Midland speakers, since few Midland speakers in my corpus have fronting of /Kuw/ greater than 1550 Hz.

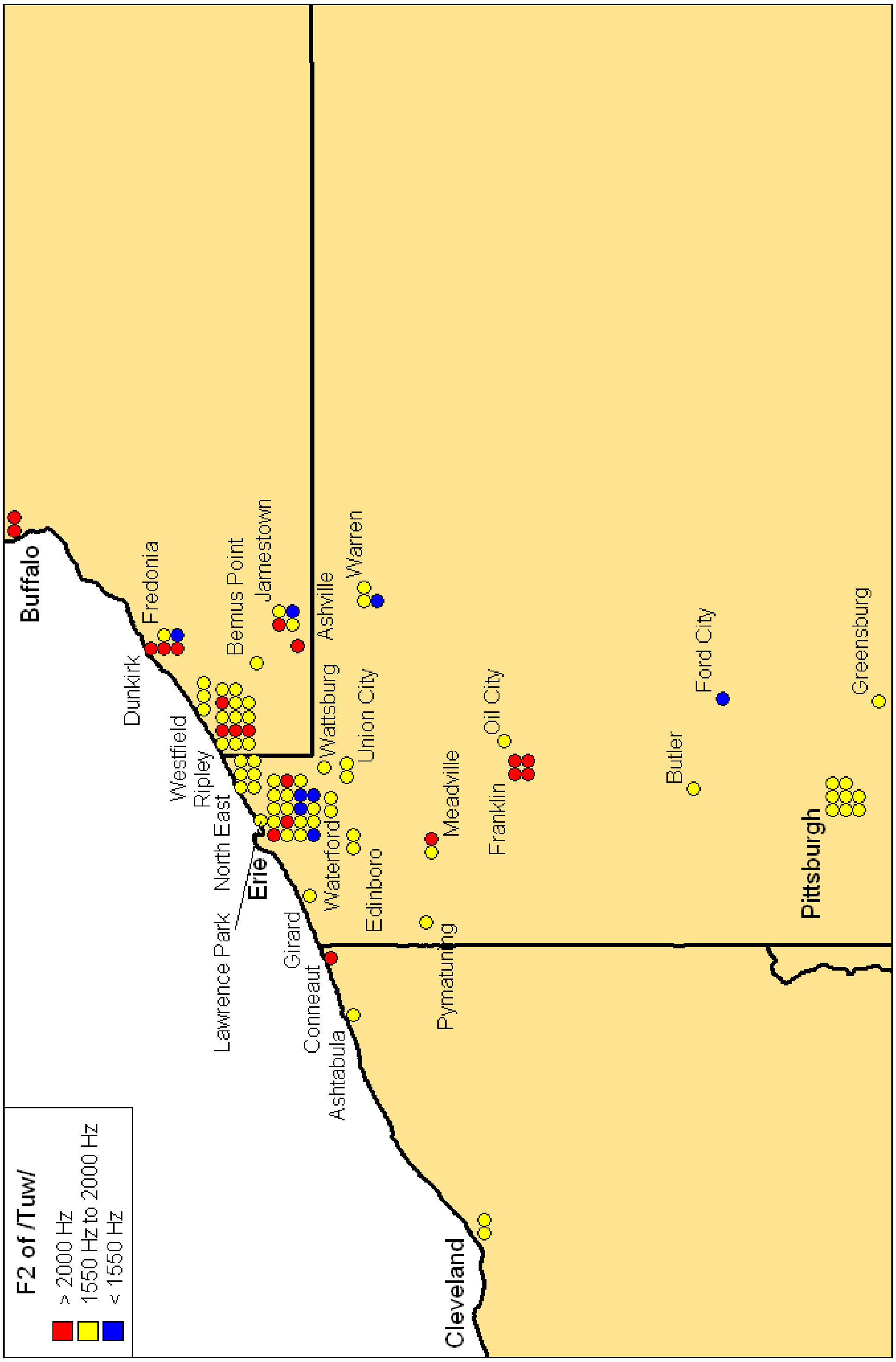

Figure 7.2: F2 of /Tuw/ (/uw/ after coronals)

Feature	Coefficient	Probability
Intercept	1148	< 0.0001
Preceding oral labial	-92	< 0.0001
Preceding nasal labial	-62	< 0.01
Preceding oral apical	123	< 0.0001
Preceding nasal apical	187	< 0.0001
Preceding palatal	124	< 0.0001
Preceding velar	41	< 0.01
Preceding liquid	-44	< 0.01
Following labial	-105	< 0.0001
Following interdental	-78	< 0.01
Following apical	22	< 0.05
Following velar	98	< 0.0001
Following fricative	-101	< 0.01
One following syllable	-25	< 0.05

Table 7.2: Linear regression coefficients for F2 of /ow/ for all environmental features significant at $\alpha = 0.05$ (N = 6,341)

not with Pittsburgh and the Midland. This is in contrast to the Northern behavior of /o/ and /oh/ in Erie described in Chapter 6.

7.2.2 /ow/

The mid vowel /ow/ is fronting in parallel with /uw/ for many speakers in North America; however, the fronting of /ow/ is, in general, not as extreme as /uw/, and most fronted tokens are not much further front than the mid line of 1550 Hz. As an overview for the allophonic conditioning on the fronting of /ow/, Table 7.2 shows the segmental environment features that have statistically significant coefficients.

The results in Table 7.2 are quite similar to a similar analysis for the ANAE data using a similar number of tokens (see Table 12.3 in the ANAE). Both regression analyses show that /ow/ fronting is increased with coronal and velar onsets but inhibited by labial onsets.

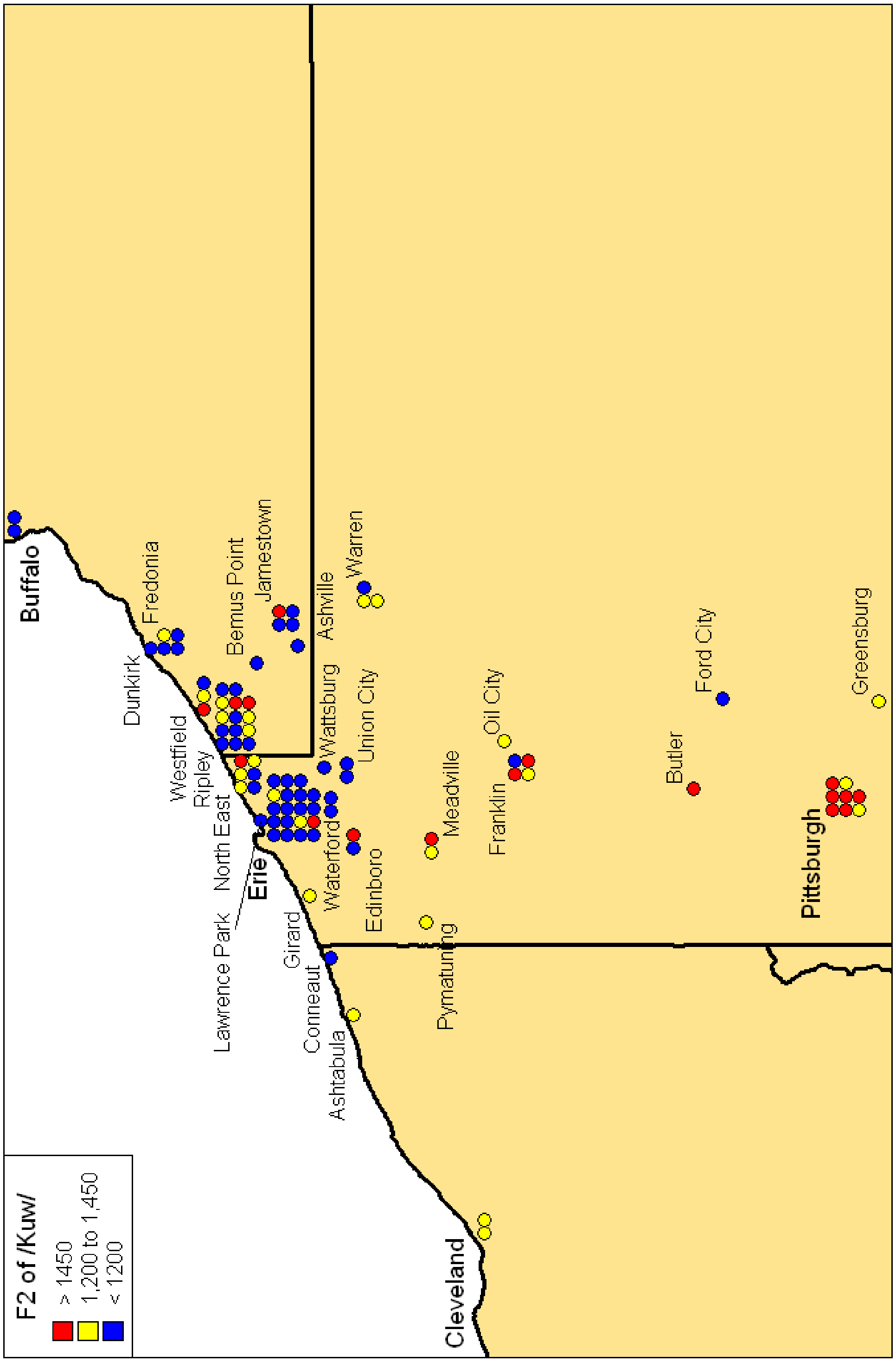

Figure 7.3: F2 of /Kuw/ (/uw/ after non-coronals)

Furthermore, both analyses show that fronting of /ow/ is inhibited by a following labial, a following fricative, or the presence of a following syllable. The analysis of my data set additionally shows a negative coefficient for preceding liquids and following interdentals as well as a positive coefficient for following velars and apicals where no statistically significant effect was found in the ANAE. On the other hand, the regression on my corpus did not find statistically significant effects for two environments that were significant in the ANAE analysis: word-final position and following nasal. Again, tokens before /l/ were excluded for this analysis (along with all other tokens matching the exclusion criteria listed in Section 3.8). A regression analysis including all tokens confirms the ANAE's finding that a following /l/ has by far the strongest effect on the position of /ow/, with a coefficient of -286 ($p < 0.0001$).

The division observed in the ANAE between the North and the Midland is even stronger with respect to the fronting of /ow/ than the fronting of /uw/. For this vowel, the Midland and the South behave quite similarly, and an isogloss based on the F2 value of /ow/ sharply divides the eastern part of North America into two regions. In the North, very few speakers have mean F2 values of /ow/ greater than 1300 Hz, whereas in the Midland and the South, very few speakers have values less than 1300 Hz. Again, Table 12.4 in the ANAE shows that the North has the lowest F2 mean value for /ow/ from among all the dialect regions, at 1127 Hz. In contrast to /uw/, however, where the speakers from Western Pennsylvania showed much less fronting than the rest of the Midland, speakers from these two regions have similar mean values for /ow/, and both are well over the 1300 Hz value separating the North and the Midland. In fact, the fronting of /ow/ is more advanced in Western Pennsylvania than in the rest of the Midland: the mean F2 of /ow/ in Western Pennsylvania is 1422 Hz, whereas it is 1367 Hz for speakers from the rest of the Midland.

As an example of a Pittsburgh speaker with strong fronting of /ow/, Figure 7.4 displays all tokens of /ow/ for Cecilia S., a speaker from the ANAE (TS 356). In this figure, all tokens of /ow/ before /l/ are marked with a black outline; these tokens are nearly all very back. On the other hand, most of the other tokens not before /l/ are fronted.[7]

As an example of a Northern speaker with a very back /ow/, Figure 7.5 shows all /ow/ tokens for Bill R. from Buffalo. In this plot, there is no separation between tokens before /l/ and tokens not before /l/: almost all tokens of /ow/ are very back. The non-fronted mean of /ow/ for Bill R. is aligned with his mean F2 value for /Kuw/, which is also quite far back (despite his strong fronting of /Tuw/).

Figure 7.6 shows the relative fronting of /ow/ for all of the speakers from my corpus. The red and orange symbols represent speakers with fronted /ow/ (defined as having a mean F2 value of /ow/ higher than 1300 Hz), whereas the symbols in three shades of blue represent speakers with non-fronted /ow/ (with a mean F2 value lower than 1300 Hz). The general pattern in Figure 7.6 is quite similar to the pattern observed for the fronting of /Kuw/ in Figure 7.3: speakers in Western Pennsylvania in the counties south of Erie mostly have fronted /ow/, whereas speakers from Erie County and Chautauqua County, NY, have non-fronted /ow/. Again, Erie's linguistic behavior with respect to this variable is clearly Northern.

The fact that the isoglosses for /ow/ and /Kuw/ are so similar suggests that the behaviors of /ow/ and /Kuw/ are structurally linked. That is, if a speaker has fronted /ow/, then he is very likely to also have fronted /Kuw/, and vice versa. A correlation test between the mean F2 values for /ow/ and /Kuw/ across all 90 speakers in the corpus provides further evidence for this: the correlation coefficient for the means of these two vowels is 0.64 ($p < 0.0001$). On the other hand, the behaviors of /ow/ and /Tuw/ are not linked at all: their correlation coefficient is 0.03 (*n.s.*).

---

[7] Cecilia S.'s mean F2 value of /ow/ is higher here at 1536 Hz than the value of 1457 calculated from the manual ANAE measurements. This discrepancy is due to the presence of several formant measurement errors (the tokens of /ow/ in Figure 7.4 with an F1 close to 800 Hz). However, the pattern wherein many tokens of /ow/ are fronted to around 1550 Hz, but most tokens remain non-fronted before /l/ is still apparent, despite the measurement errors.

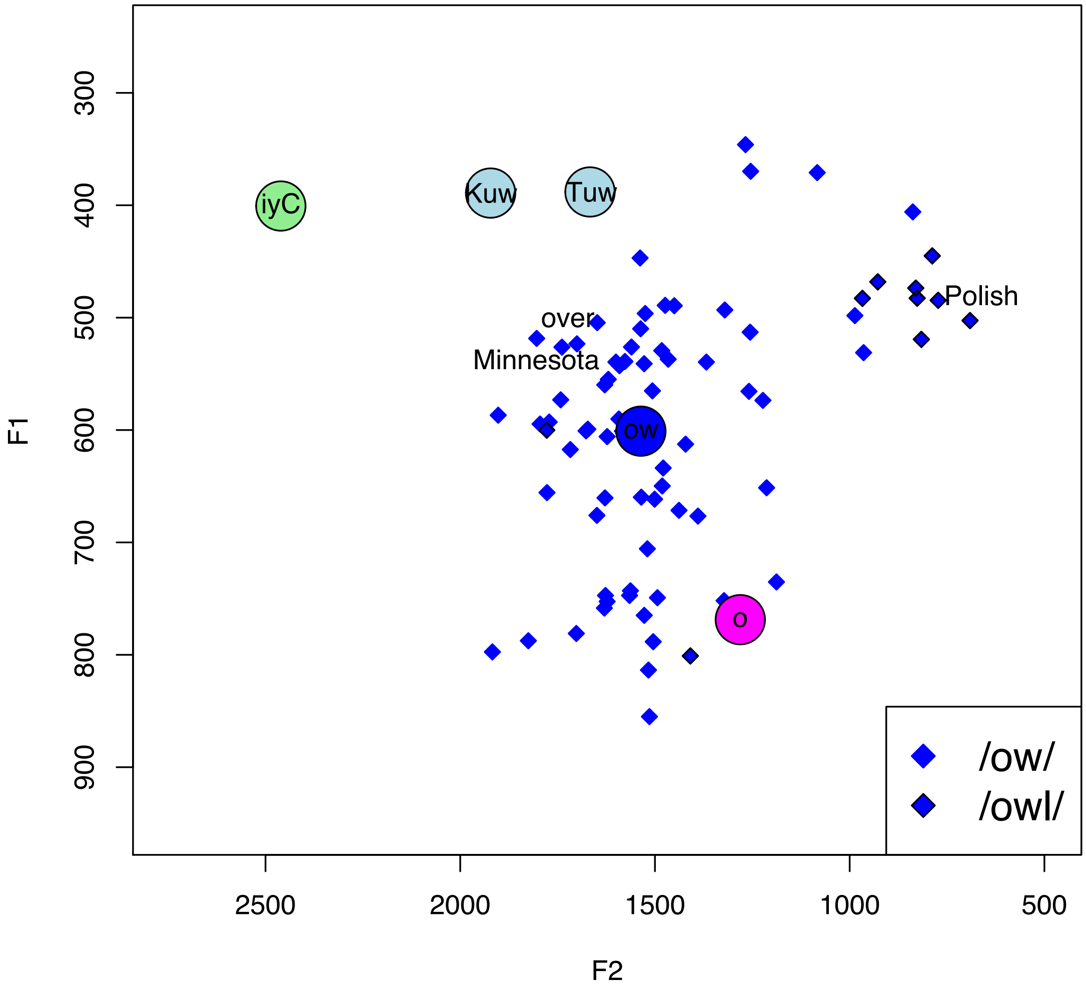

Figure 7.4: /ow/ for Cecilia S., born 1933 in Pittsburgh
Mean(/ow/) = (601, 1536), N=63

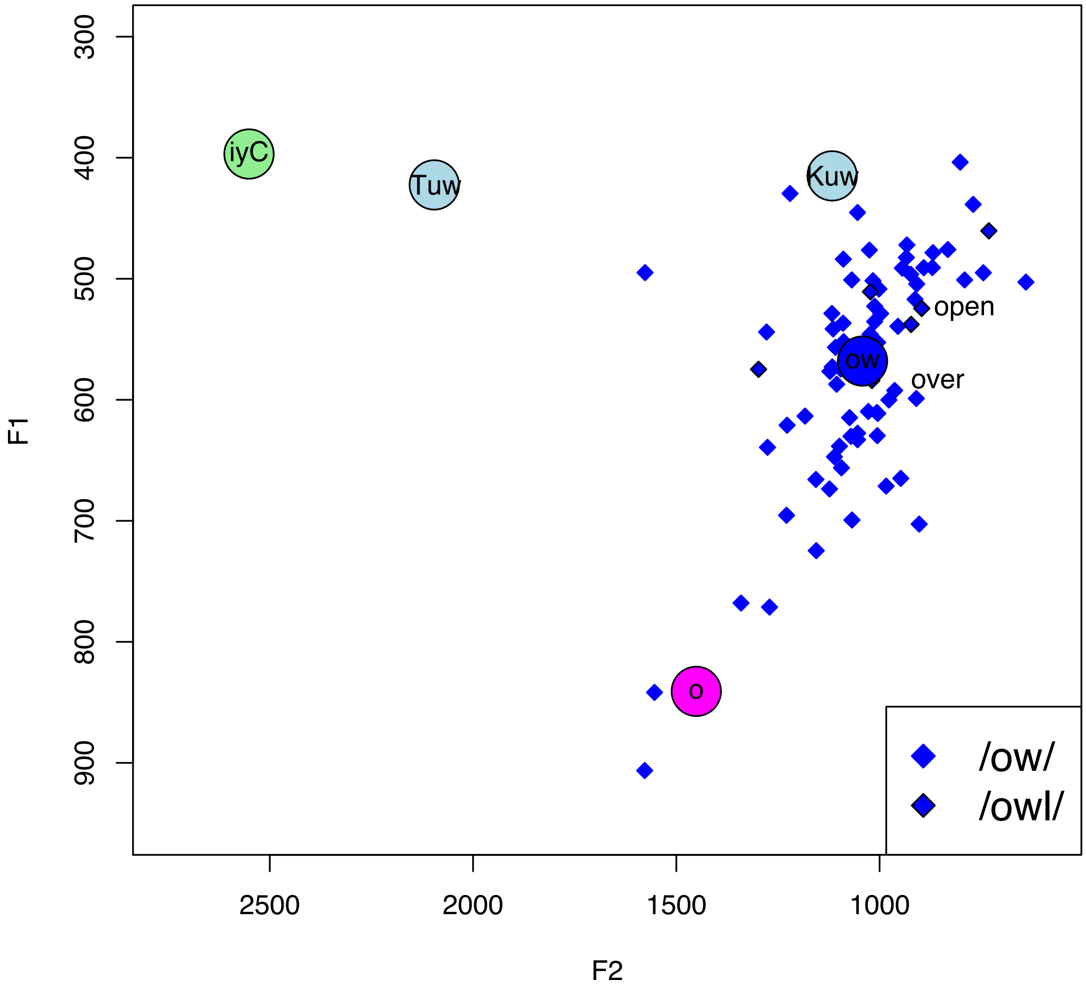

Figure 7.5: /ow/ for Bill R., born 1927 in Buffalo
Mean(/ow/) = (568, 1041), N=65

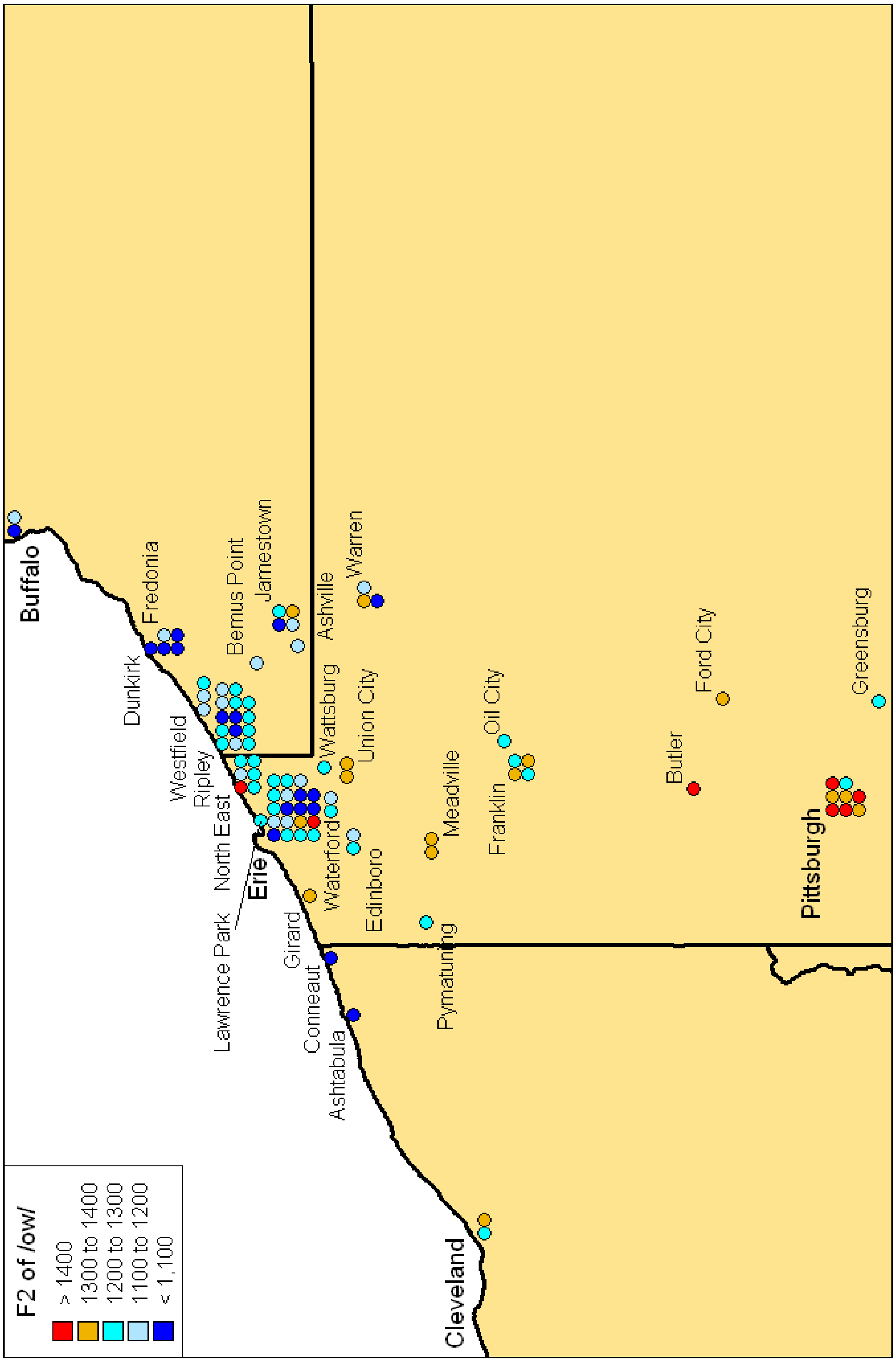

Figure 7.6: F2 of /ow/

7.2.3 /aw/

The lowest member of the system of back upgliding vowels, /aw/, is also fronted in much of the Midland and the South. For many of these speakers, the nucleus of /aw/ is consistently front of center, and is closer to [æw]. The general regional pattern is the same as it is for /Kuw/ and /ow/: Northern speakers mostly show conservative treatment of /aw/ with mean values less than 1550 Hz; Midland and Southern speakers show moderate to strong fronting. However, the strongest areas of fronting (where the mean F2 values of /aw/ are consistently over 1800 Hz) are found in the Mid-Atlantic region and the South. The Midland region around Western Pennsylvania is mixed with respect to this variable. The isogloss for the region with fronted /aw/ in Map 12.4 in the ANAE actually passes to the south of Pittsburgh. Thus, Pittsburgh and Erie are both included in the same isogloss as the North for /aw/. However, the speakers in both Pittsburgh and Erie are split 50-50 for this isogloss: three out of the six speakers from Pittsburgh have a mean F2 of /aw/ less than 1550 Hz, as does one out of the two speakers from Erie.

Figure 7.7 displays the values for this quantitative isogloss for /aw/ for the speakers in my corpus. As is the case for the ANAE data from the region, the speakers are mixed and no clear regional pattern is apparent. Roughly one-third (31 / 90) of the speakers from the corpus have a mean F2 value for /aw/ greater than 1550 Hz, and two-thirds (59 / 90) have a non-fronted /aw/.

The ANAE also defines a qualitative measure, the AWY criterion, for the fronting of /aw/ based on its position relative to the mean of /ayV/. Again, the Northern speakers, in general, show conservative behavior with a mean F2 value of /aw/ less than the mean of /ayV/. The Midland and Southern speakers, on the other hand, almost categorically show a mean value of /aw/ greater than the F2 mean of /ayV/. This structural isogloss shows higher consistency than the quantitative isogloss based on the value of 1550 Hz, and the ANAE authors use this fact to argue that the AWY line is better at dividing the fronted and

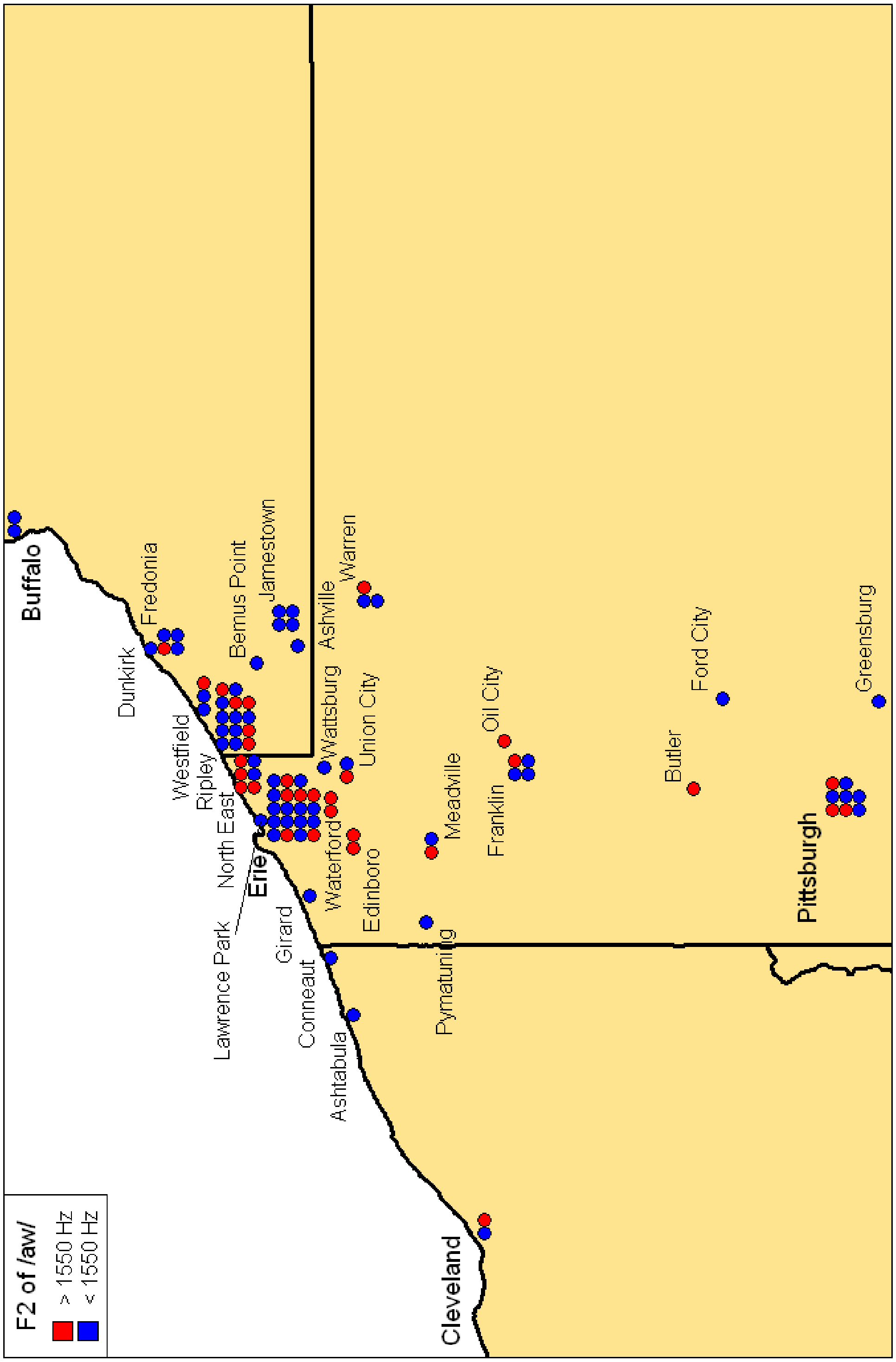

Figure 7.7: F2 of /aw/

non-fronted regions into two distinct groups Labov et al. (2006:160).

For the purposes of this study, a comparison of these two measures of the fronting of /aw/ is interesting, since they show different patterns in Pittsburgh and Erie for the ANAE speakers. As discussed above, the speakers in Erie and Pittsburgh were split evenly for the quantitative measure of /aw/ $>$ 1550 Hz, and were included inside the Northern isogloss in ANAE Map 12.4. Conversely, ANAE Map 12.5 shows that the six Pittsburgh speakers and the two Erie speakers categorically have a mean F2 value of /aw/ greater than the value of /ayV/. The structural isogloss for the non-fronted North thus does not include either of these two cities in western Pennsylvania. This area where /aw/ is more front than /ayV/ also stretches up from the Midland in northwestern Ohio to include Cleveland; thus, there is a discontinuity in the Northern isogloss for the AWY criterion stretching from Toledo to Buffalo.

Figure 7.8, however, shows that the speakers from my corpus are more mixed in this qualitative isogloss than the ANAE speakers from western Pennsylvania. Approximately one-fourth of them (23 / 85)[8] show a fronted value of /aw/ in comparison to /ayV/, whereas three-fourths (62 / 85) show non-fronted /aw/.

Thus, it appears that the fronting of /aw/ is not as closely tied to the fronting of /ow/ and /Kuw/ as the latter two are to each other. Correlation tests do show significant correlation coefficients between /aw/ and each of the other two back upgliding vowels. However, their values are much smaller than the value for the correlation between /ow/ and /Kuw/: the correlation between the F2 of /aw/ and the F2 of /ow/ is 0.26 ($p < 0.05$), and it is 0.27 ($p < 0.05$) for /aw/ and /Kuw/.

There is also no clear regional differentiation based on the fronting of /aw/ before /n/, an environment that generally favors stronger fronting, and might thus distinguish the

---

[8] Five speakers are not included in this count or in Figure 7.8 because they did not produce any tokens of /ayV/. These speakers were only recorded reading the word list, and they read an earlier version of the list that did not include the tokens such as *high*, *rider*, and *hide* that are listed in the word list in Appendix C.

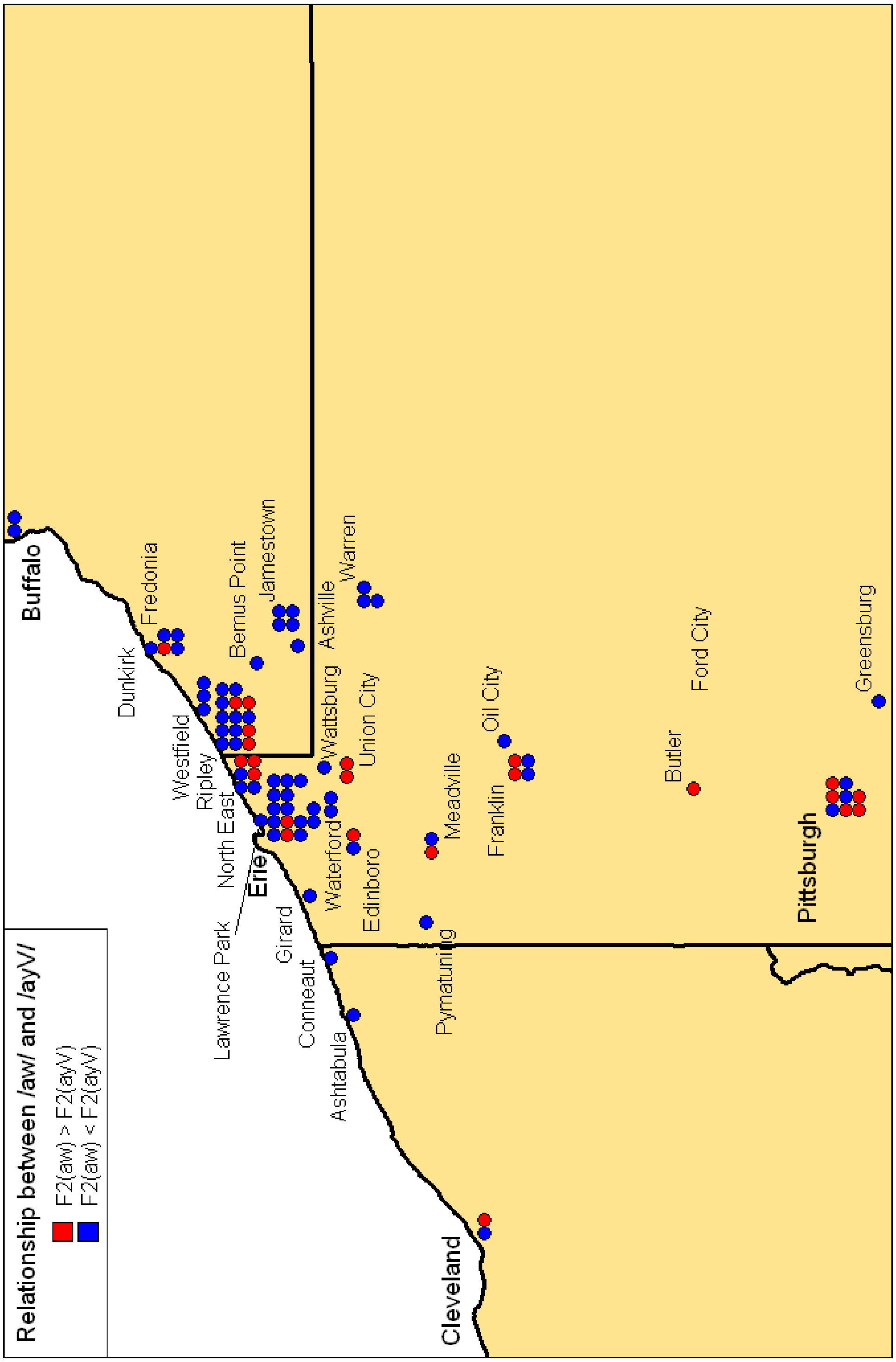

Figure 7.8: Relationship between /aw/ and /ayV/

Midland and Northern speakers better. Only six speakers in the corpus have a mean F2 value of /aw/ before nasals greater than 1800 Hz: they are scattered throughout the entire region, and are located in Oil City, Edinboro, Erie, North East, Ripley, and Cleveland. The remaining speakers do not show any regional pattern.

7.3 NCS vowels

Chapter 6 showed that the merger of /o/ and /oh/ spread to Erie around the turn of the 20th century. Subsequently, it continued to spread northward, and became a feature of the town of Ripley, NY in the latter half of the 20th century. On the surface, this slow spread of the merger to these areas is unsurprising. It coincides with the increasing spread of the merger throughout many areas of North America, and can be seen as an instantiation of Garde's Principle (by which mergers expand at the expense of distinctions). However, when considered in the specific context of where the merger has spread to, its advance through Erie County, PA and into Chautauqua County, NY is actually quite remarkable. This is because the North, with the fronting of /o/, represents one of the three regions of stable resistance to the spread of the merger of /o/ and /oh/ in North America.[9] The fact that /o/ is fronted in the North means that the distance between /o/ and /oh/ is generally quite large. This large margin of security brings about the Northern resistance to the spread of the merger.

This section will consider the the two main components of the Northern Cities Shift: the generalized raising of /æ/ and the fronting of /o/. Specifically, the progress of the NCS in Chautauqua County will be examined together with the spread of the merger of /o/ and /oh/ to Ripley in an attempt to understand why the merger was able to spread into this area of the North.

---

[9] The other two are the Mid-Atlantic region with raised /oh/ and the South with back upgliding /oh/.

7.3.1 /æ/

The fronting and raising of /æ/ in restricted phonological environments is quite common in North American English. For example, nearly all dialects have fronting and raising before nasal codas (Canada, especially Montreal, is the main exception to this trend (Boberg 2005)). Areas with a continuous system of fronting and raising before /æ/ often have tokens before /d/ and /g/ in an intermediate position between tokens before nasals and tokens before voiceless obstruents. An even more complex system is observed in Philadelphia and New York, where tensing occurs in a wide variety of phonological environments, in addition to being lexically specified for some words and sensitive to morphological constraints (see Labov (1994:429–431) for details of these systems). However, the generalized fronting and raising of /æ/ in all phonological environments is unique to the North. It was the triggering event for the Northern Cities shift, and is the structural precondition that enables the fronting of /o/ in the North. In other dialects where /æ/ is only raised in certain environments, some tokens of /æ/ still remain in low front position and inhibit the movement of /o/ in that direction.

A quantitative measure based on the degree of raising of /æ/ is one of the five criteria that define the North in the ANAE. This isogloss is called the AE1 line, and is defined to include speakers whose F1 mean value of /æ/ (excluding tokens before nasals) is less than 700 Hz. Map 14.4 in the ANAE shows that the largest concentration of speakers matching the AE1 criterion is in the North and the St. Louis corridor. Again, there is a discontinuity in the isogloss at Erie, since neither of the two ANAE speakers from Erie are selected by the AE1 criterion. The homogeneity and consistency values for this isogloss are less than they are for the other defining isoglosses of the North: there are many speakers outside the North with F1 mean values of /æ/ less than 700 Hz (especially in the South), as well as several speakers in the North who do not match this criterion.

Figure 7.9 shows the results for the AE1 criterion for the speakers in my corpus. The highest concentration of speakers with F1 of /æ/ less than 700 Hz is seen in Chautauqua County, NY, especially when the towns of Ripley and Westfield are excluded. There are also several speakers from Pennsylvania that match the AE1 criterion: 12 out of a total of 56. However, all of the speakers from Pennsylvania with an F1 of /æ/ less than 700 Hz are advanced in age. In fact, most of them come from archival sources: the red symbols in Figure 7.9 from Warren, Meadville, and Union City, as well as one of the three from North East are from the DARE recordings; the other two from North East are the speakers from the SWV archive. All of the other speakers from Pennsylvania who match the AE1 criterion (in Pittsburgh, Franklin, Wattsburg, Edinoboro, and Erie) are older than 60.

Figure 7.10 shows the relationship between the raising of /æ/ among the speakers from Pennsylvania and age. The dashed line represents the AE1 criterion, and the figure shows that all speakers with values below the line are older than 60.[10] In addition, many of the other elderly speakers have F1 values for /æ/ close to the 700 Hz line. A linear regression for the F1 value of /æ/ predicted by age shows a significant coefficient of -1.55 ($p < 0.0001$). This means that the regression model subtracts 39 Hz from the F1 value of /æ/ for every 25 years of age, i.e., the raising of /æ/ is decreasing in apparent time.

However, when the same analysis is conducted for the 30 speakers from New York (this group includes the 28 speakers from towns in Chautauqua County and two from the city of Buffalo), a different picture emerges. For these speakers, there is no significant effect based on age. In general, the speakers from New York show more raising of /æ/: there are only 6 speakers from New York with a mean F1 value greater than 750 Hz (compared to a total of 29 speakers from Pennsylvania).

The findings from the ANAE also show a recession in apparent time for the raising of /æ/, both in the North and elsewhere; however, the age coefficient is much larger for the

---

[10] Since several of the speakers shown in Figure 7.9 are drawn from archival sources, the x-axis represents their projected current age, calculated by subtracting their year of birth from 2009.

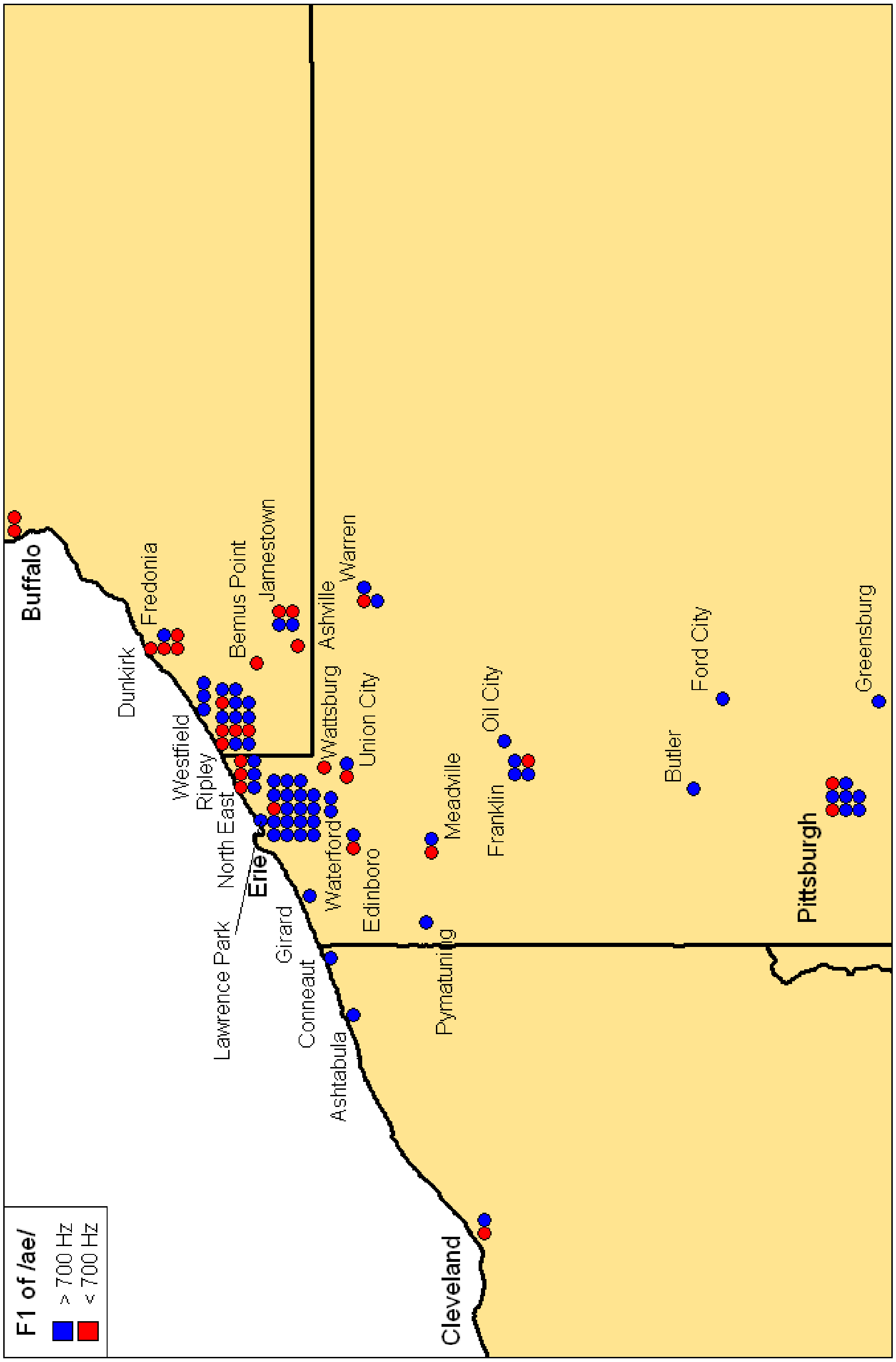

Figure 7.9: Raising of /æ/

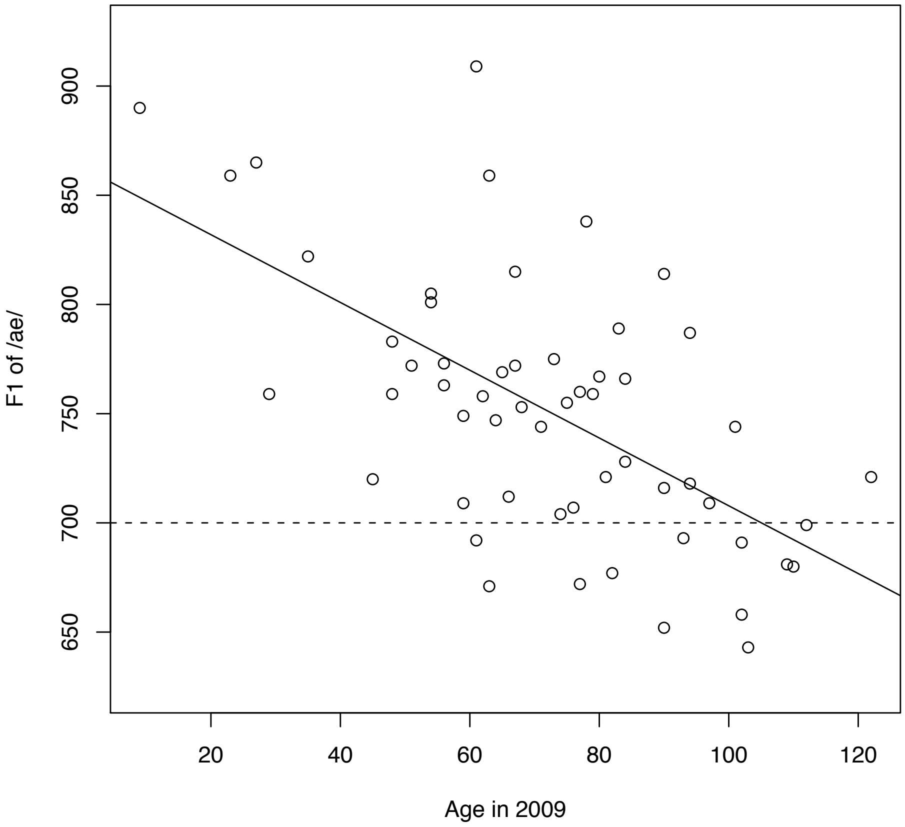

Figure 7.10: F1 of /æ/ by age for 56 speakers from Pennsylvania; Regression coefficient for age is -1.55 ($p < 0.0001$), $r^2 = 0.37$; dashed line shows AE1 criterion (F1 of /æ/ < 700 Hz)

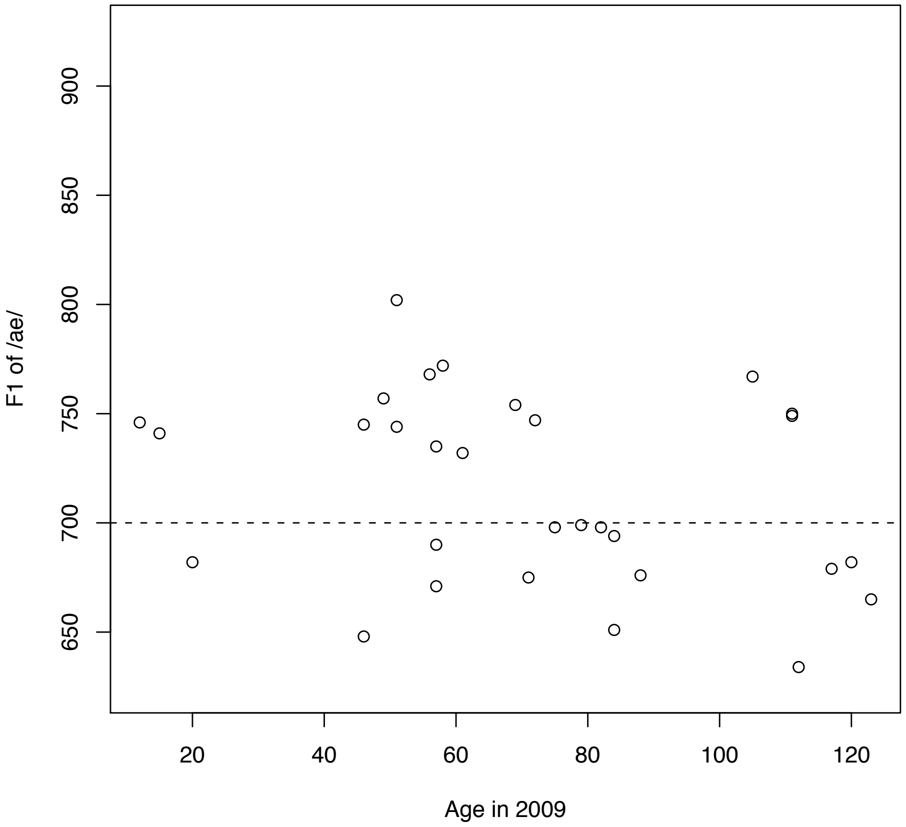

Figure 7.11: F1 of /æ/ by age for 30 speakers from New York; Regression coefficient for age is not significant at $\alpha = 0.05$; dashed line shows AE1 criterion (F1 of /æ/ < 700 Hz)

speakers from the North (Labov et al. 2006:195). Compared with this result, it is intriguing that the Northern speakers from Chautauqua County and Buffalo in my corpus do not also show a significant age effect on the F1 of /æ/. This is likely because the speakers from Chautauqua County do not have /æ/ raised nearly as much as the most advanced speakers in the Inland North. For the speakers from Chautauqua County who do match the AE1 criterion, Figure 7.11 shows that nearly all of them are clustered in the range between 650 and 700 Hz. This is in contrast to the ANAE, where the most advanced Northern speakers have a mean F1 of /æ/ less than 500 Hz (see Map 14.3). Chautauqua County can thus be considered to be on the fringes of the NCS area.

On the other hand, the age coefficient showing the lowering of /æ/ in apparent time for the speakers from Pennsylvania is much larger than the related coefficient calculated for all non-Northern speakers in the ANAE. In the case of the speakers from my corpus, this effect reflects a shift away from membership in the North. The fact that several elderly and archival speakers from the area of Pennsylvania around Erie have a raised /æ/ shows that they patterned with the other Northern speakers of their generation and had the structural precondition for the NCS. However, the fact that /æ/ is not raised for all speakers from Pennsylvania under age of 60 in my corpus shows that this precondition has disappeared. Stage 2, the fronting of /o/ is now blocked for these speakers, as will be demonstrated in the following section.

7.3.2 /o/

The second defining criterion for the North provided by the ANAE is based on the fronting of /o/, Stage Two of the NCS. It is called the O2 line, and is defined to select speakers whose mean F2 value of /o/ is greater than 1450 Hz. The isogloss based on the O2 criterion has a very high consistency in the ANAE: almost no speakers outside of the North are selected by it (see Map 14.5).

Figure 7.12 displays the results for the O2 criterion for the speakers in my corpus. The distribution of speakers is similar to the one shown in Figure 7.9 for the AE1 line, with even fewer speakers from Pennsylvania matching the O2 criterion. As was the case for the AE1 criterion, all of the Pennsylvania speakers who are aligned with the North in the fronting of /o/ are either elderly or drawn from archival sources.[11] This evidence suggests that the fronting of /o/ used to be more widespread in Erie County, but was in the process of receding around the turn of the 20th century.

The status of the speakers from Ripley, NY with regard to the O2 criterion is informative in understanding the spread of the merger of /o/ and /oh/ to that town. Figure 7.12 shows that four speakers from Ripley have an F2 mean value of /o/ greater than 1450 Hz. One of these is an archival speaker from DARE, Jill C. (born in 1889). The other three are Stan R. (born 1948), Rachel A. (born 1951), and Daphne R. (born 1958). These three speakers make up half of the list of six speakers in my corpus who maintain a clear distinction between /o/ and /oh/, judged on the basis of the minimal pair test results (see Table 6.12 in Chapter 6). Thus, based on this evidence alone, it would appear that satisfying the O2 criterion is a sufficient condition (but not a necessary) one for a speaker to maintain the distinction between /o/ and /oh/. However, the evidence from the two speakers in Jamestown who do not match the O2 criterion, but nevertheless maintain a clear distinction between /o/ and /oh/, shows that the actual relationship between a speaker's mean F2 value of /o/ and their status with regard to the merger is more complex than this.

Figure 7.13 plots the relationship between a speaker's mean F2 value of /o/ and their status for the minimal pair test of *Don* vs. *dawn*. In this figure, the production and perception scores are summed together to provide a single score on a scale of 0 to 4 for each speaker. On this scale, 0 represents a complete merger and 4 represents a complete distinc-

---

[11] The Pennsylvania speakers with a mean F2 greater than 1450 Hz from Warren, Union City, and North East are from archival sources and are no longer alive. The two other speakers from Erie and Wattsburg were born in 1916 and 1927, respectively.

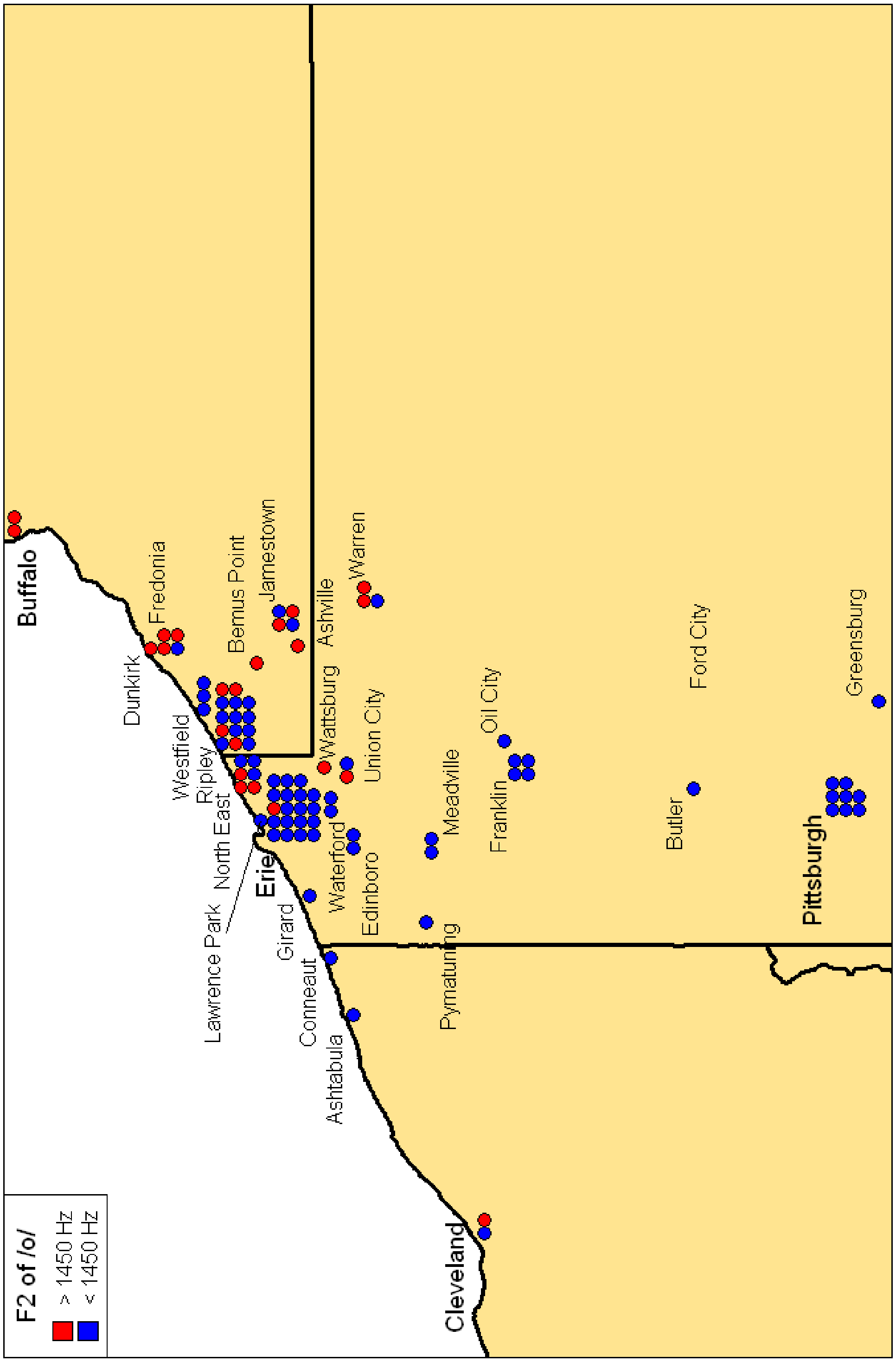

Figure 7.12: The fronting of /o/

tion.[12] The general trend shown in the figure is that speakers with higher F2 mean values for /o/ are more likely to maintain the distinction between /o/ and /oh/, and, conversely, speakers with lower values are more likely to have the merger. A linear regression model predicting the status of the minimal pair test by a speaker's mean F2 value of /o/ shows an increase of 0.96 points on the 5-point scale for every 100 Hz ($p < 0.0001$, $r^2 = 0.41$).[13]

A more precise relationship can be seen by examining the O2 criterion, drawn as a dotted vertical line in Figure 7.13. This line shows that only a single merged speaker has an F2 mean value of /o/ greater than 1450 Hz (Robert E. from Erie, born 1916); however, his mean F2 value is only slightly over the O2 line at 1456 Hz. On the other hand, the figure shows that there are several speakers with a complete distinction between *Don* and *dawn* who have a mean F2 value less than 1450 Hz. Thus, the correct characterization of the relationship between these two features appears to be that if a speaker has the merger of /o/ and /oh/, then their mean F2 value of /o/ will be less than 1450 Hz. That is, the merger of /o/ and /oh/ is a sufficient condition for a speaker not satisfying the O2 criterion.

In addition, Figure 7.14 shows the results for a third measure of the NCS from the ANAE, the ED criterion. This is a quantitative measure of the relationship between the mean values of /e/ and /o/, and is satisfied when their difference is less than 375 Hz. Again, the map based on the ED criterion (Map 14.7 in the ANAE) shows a very high consistency for this feature in the North—very few speakers outside of the North satisfy it. Figure 7.14 shows that only eight speakers in my corpus satisfy the ED criterion. One of them is an archival speaker from Warren, PA, another is one of the three unmerged Ripley speakers who satisfy the O2 criterion, and the remaining six are from other towns in Chautauqua County. The fact that only one out of 14 speakers from Ripley satsified the

---

[12] For the production and perception halves of each minimal pair test, a response of "same" is given the value 0, "close" is given the value 1, and "different" is given the value 2.

[13] Note, however, that almost all speakers in the corpus are either categorically merged or unmerged for this minimal pair. Only two speakers have intermediate values between 0 and 4. This is in contrast to the Midland speakers from the ANAE, who mostly show intermediate values.

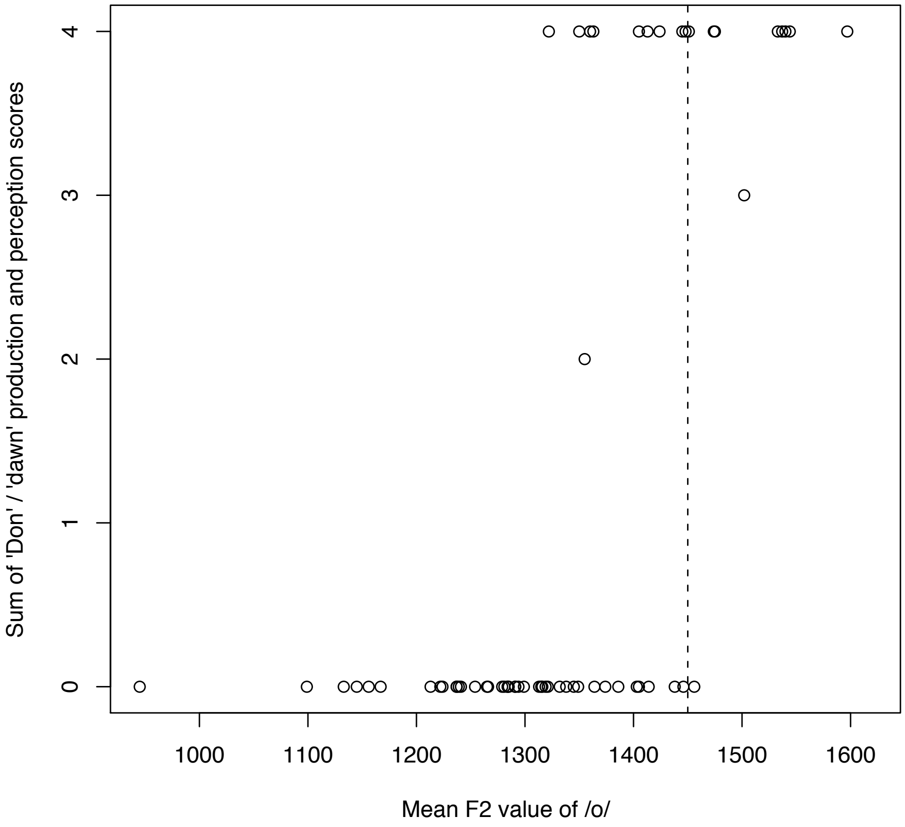

Figure 7.13: The relationship between a speaker's mean F2 value of /o/ and the merger of /o/ and /oh/ in *Don* vs. *dawn*;
dotted line shows the O2 criterion (F2 of /o/ > 1450 Hz)

ED criterion shows that the NCS is not active at all in that town.

Finally, the anomalous status of the three speakers from the town of Westfield must be discussed. All three of these speakers maintain a clear distinction between /o/ and /oh/ in the minimal pairs tests (see Figures 6.15 – 6.18 in Chapter 6). However, none of them are selected for any of the three criteria of the NCS discussed in this chapter: they all have a mean F1 value of /æ/ greater than 700 Hz, a mean F2 value of /o/ less than 1450, and a difference between the F2 means of /e/ and /o/ greater than 375 Hz. Figures 7.9, 7.12, and 7.14 show that these three speakers from Westfield are grouped together with most of the speakers from Ripley and Pennsylvania for these three criteria, and not with the speakers from the other towns in Chautauqua County, NY.

Thus, it appears that the NCS is also currently not a feature of the speech of Westfield. Figure 7.15 displays a plot of all tokens of /æ/, /o/, and /oh/ for Amy C., one of the three speakers from Westfield. There is little overlap between the distributions of /o/ and /oh/—her acoustic data thus matches her experimental results from the minimal pair tests. However, her mean F2 value of /o/, 1350 Hz, is 100 Hz below the O2 criterion.

The other two unmerged speakers from Westfield who do not match any of the NCS criteria were born in 1940 and 1960. All three of them thus represent the speech patterns of two to three generations ago. It is possible that these three speakers indicate that the Northern structural factors preventing the spread of the low-back merger were absent from Westfield approximately 60 years ago (as they were in Ripley), and that the merger of /o/ and /oh/ could be in the process of speading to Westfield as well. Unfortunately, my corpus does not contain any younger speakers from Westfield, so this hypothesis can not be tested with the data at hand. However, Westfield is the next town over on the shore of Lake Erie past Ripley, and is only 8 miles away from Ripley. Thus, its proximity to the merged area and its lack of NCS features would make it quite likely that the merger will spread there, too.

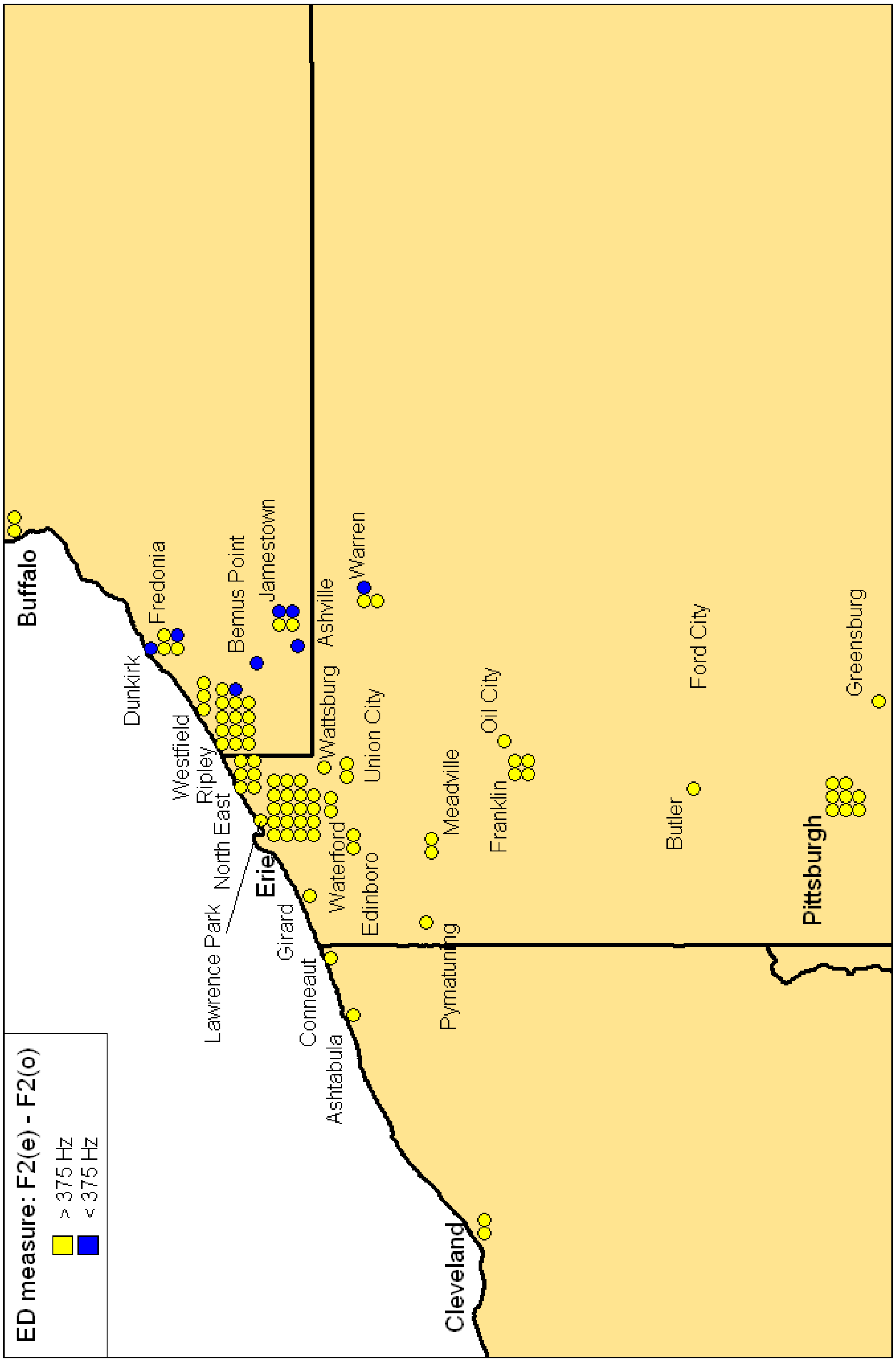

Figure 7.14: The relative position of /e/ and /o/ in the F2 domain

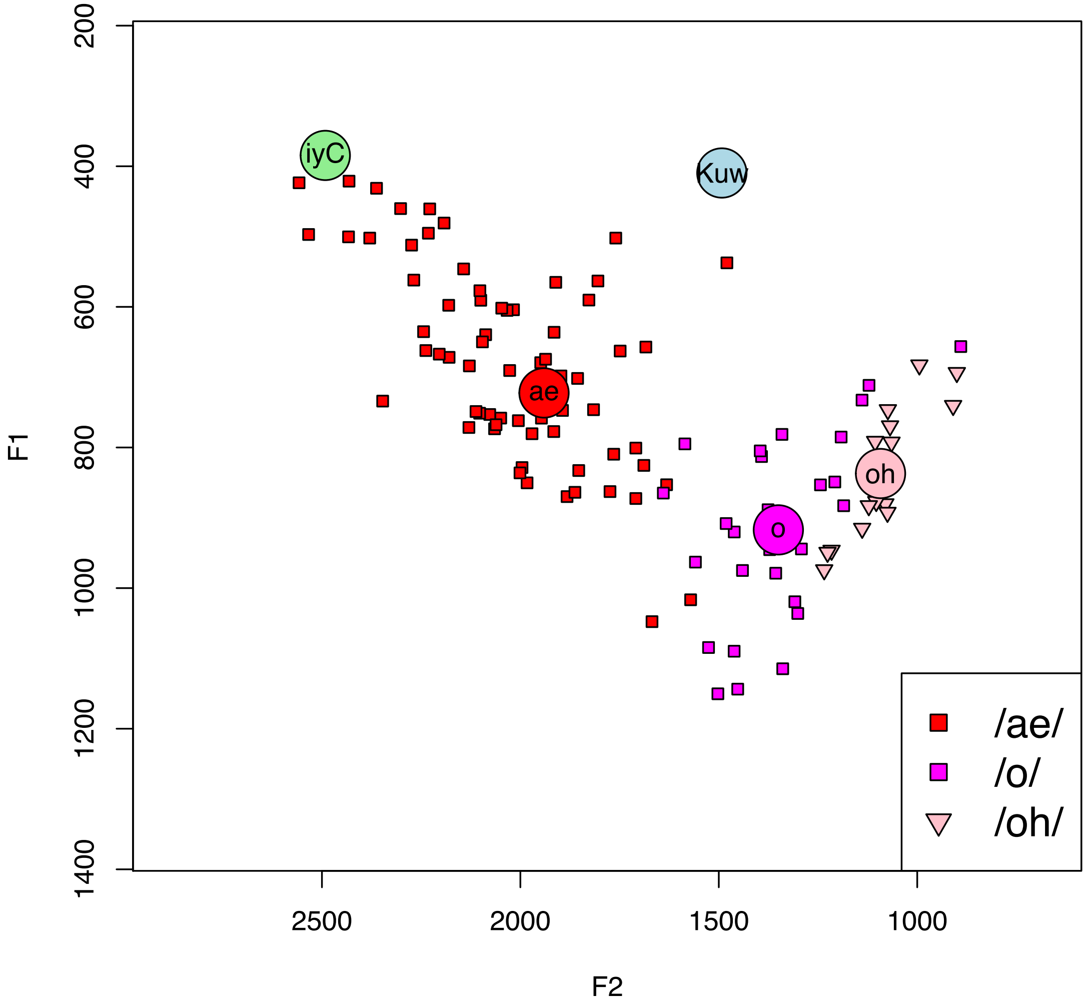

Figure 7.15: /æ/, /o/, and /oh/ from Amy C., born 1937 in Westfield; Mean(/o/) = (917, 1350), N=26; Mean(/oh/) = (837, 1092), N=14; Dist(/o/, /oh/) = 270

7.4 Other changes in progress

Finally, this section will consider two other phonological changes that can be observed in my corpus. These changes, however, do not show regional differntiation, but rather a distribution based on age. Both involve the disappearance of phonemic distinctions that were originally more widespread, specifically the distinction between /hw/ 'which' and /w/ 'witch' and the distinction between /uhr/ 'poor' and /ohr/ 'pour'.[14]

Both of these pairs display an apparent time distribution among the speakers in my corpus: in both cases, the older speakers are more likely to maintain a distinction, while the younger speakers are more likely to have them merged. These two pairs are the only minimal pairs that show an apparent time distribution among the 17 minimal pairs studied. They are the only two for which correlation tests between the minimal pair scores and the speaker's age only produce significant results. The correlation coefficients are 0.62 ($p < 0.0001$) for *poor* vs. *pour* and 0.43 ($p < 0.05$) for *which* vs. *witch*. The positive correlation here means that as age increases, the production value for the minimal pair test also goes up (as described above, a value of 2 represents a distinction in production; 1 means that the two pairs were close, but slightly different; 0 means represents a merger).

Map 8.1 in the ANAE shows the distribution for the merger of /hw/ and /w/ across all of North America. The only region where the distinction is consistently maintained is the South. Among the speakers in my corpus, seven maintain the distinction in production for the pair *which* vs. *witch*, and most of them are advanced in age. The range of ages for these seven speakers spans from 57 to 78, with an average age for the group of 68. It is quite likely that all traces of this distinction will have disappeared within the next one or two generations.

The data for the distinction between /uhr/ and /ohr/ is quite sparse in the ANAE, and

---

[14] All speakers in my corpus have the merger of /ohr/ as in 'hoarse' and /ɔhr/ as in 'horse'. The vowel resulting from this merger will be represented by the symbol /ohr/.

there is thus no map for it. However, the merger of these two vowels is discussed in the ANAE in connection with the Back Chain Shift before /r/. This chain shift is initiated by the merger of /ohr/ ‘four’ and /ɔhr/ ‘for’. This then enables the raising of /ahr/ to a mid-back position, and causes the subsequent raising of /ohr/ $\sim$ /ɔhr/. The final stage of this chain shift is then the merger of /ohr/ $\sim$ /ɔhr/ with /uhr/.

This chain shift is widespread in the Mid-Atlantic region, and is also found in other areas. However, it is not the cause of the merger of *poor* and *pour* among the speakers in the region around Erie. For these speakers, the merged phoneme is a mid-back vowel, similar to the vowel from the original /ohr/ class, not the high back vowel that would be the outcome of the Back Chain Shift before /r/. Thus, it is more likely that the speakers in my corpus instead provide evidence for a merger by transfer in which individual lexical items from the /uhr/ class are moving into the /ohr/ class. This lexical variation is widespread in North American English, and it often affects the more frequent words such as *sure* and *poor* first (Labov et al. 2006:273). Unfortunately, there are only a few tokens of other, less common, words with /uhr/ in my data set, so it is not possible to test this hypothesis to see whether the less frequent words are less likely to merge with the /ohr/ class.

Figure 7.16 displays the results for *poor* vs. *pour* geographically. While no regional pattern emerges from this figure, the apparent time distribution is visible: the unmerged speakers cluster towards the upper-left portion of the symbols for each location. Since the symbols are arranged in decreasing order by age, this distribution matches the result from the correlation test showing that the distinction is receding in apparent time.

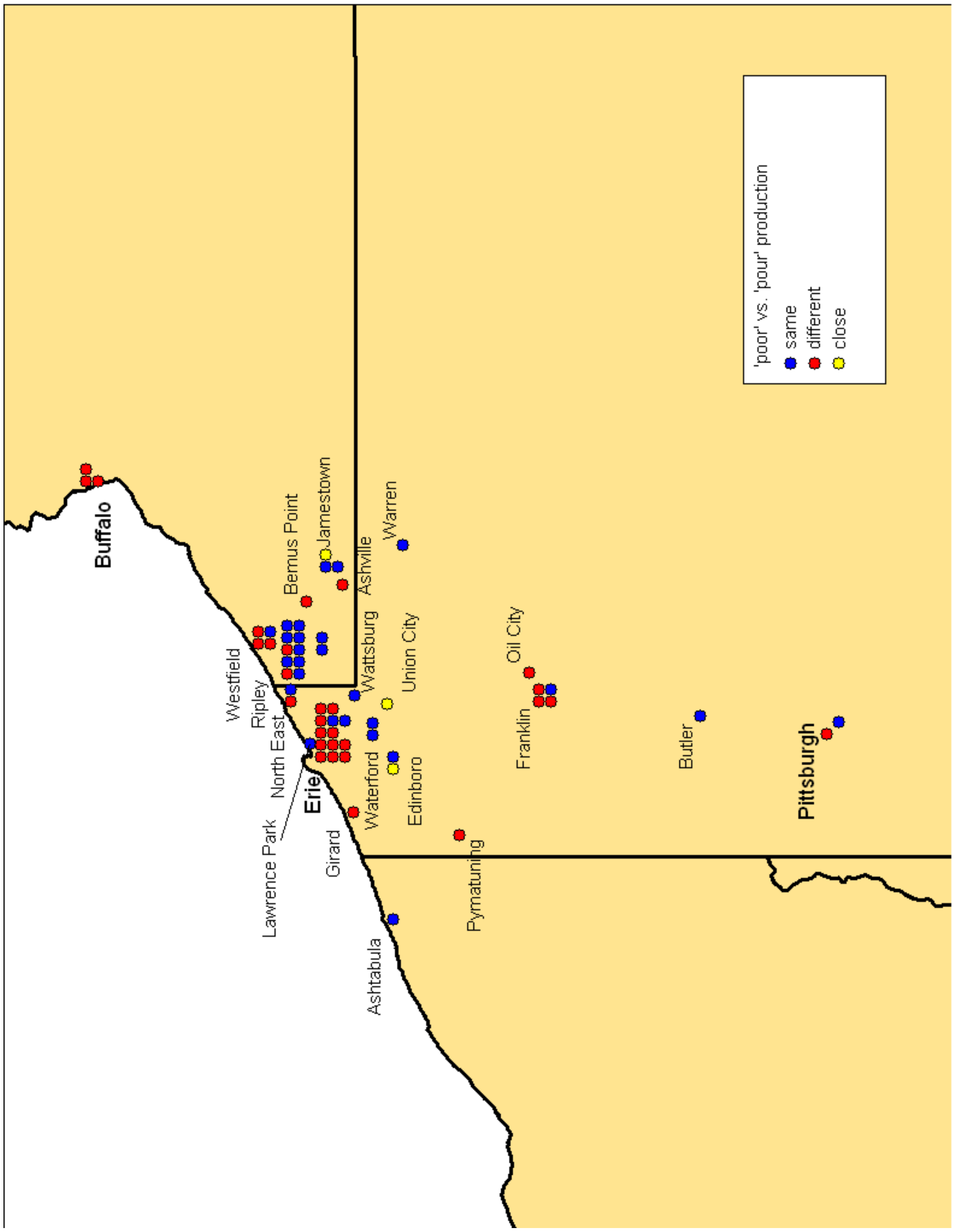

Figure 7.16: Minimal pair results for *poor* vs. *pour*: production data

Chapter 8

Lexical and Morphosyntactic Items

8.1 Introduction

Chapters 6 and 7 presented the present-day phonological evidence for considering Erie to be part of the Midland dialect region. This section will continue along those lines and consider the evidence from a few lexical and morphosyntactic variables. As was shown in Section 2.2, the lexical evidence from the earlier dialect atlases LAMSAS and DARE, representing the state of the language in the early 20th century, overwhelmingly supports grouping Erie with the North at that point in time. Unfortunately, most of the lexical items used in those earlier surveys have become obsolete, so it is not possible to conduct a controlled study based on the same words to determine how the boundaries have changed.

Despite the fact that lexical and morphosyntactic variables are not closely tied together as part of a larger structural system, as is the case for phonological variables, the evidence from LAMSAS and DARE shows that the regions defined by lexical isoglosses correspond well to the regions defined by phonological ones in ANAE. Figure 11.16 in ANAE (Labov et al. 2006:150) compares the regions based on phonological isoglosses to the regions defined in Map 8.1 in (Carver 1987:248). There is a close correspondance between Carver's

Upper North and ANAE's North, as well as Carver's Lower North and ANAE's Midland. Furthermore, Carver's boundary between the Lower North and the Upper South corresponds well with ANAE's boundary between the Midland and the South. Additionally, the regions defined in Figure 3 of Kurath (1949) correspond pretty well to the regions in ANAE, especially the division between the North and the Midland. Figure 14.11 in ANAE (Labov et al. 2006:207) emphasizes the close agreement between the lexical boundary between the North and the Midland from Carver (1987) and the phonological boundary. The only city which is to the north of the lexical boundary but is not within the phonological boundary of the North is Erie.

This section will present evidence from two lexical variables (stress assignment in the word *elementary* and the phrase *redd up*) and two morphosyntactic variables (positive *anymore* and *need* + Past Participle). This evidence will also show that Erie speakers fall on the Midland side of the boundaries; thus, the present-day lexical and morphosyntactic evidence from Erie is not as mis-aligned with the phonological evidence as is the case in ANAE Figure 14.11.

8.2 *Elementary*

An initial piece of non-phonological evidence comes from the lexical item *elementary*. Speakers in Upstate New York generally place secondary stress on the penultimate syllable of this word, whereas speakers in Erie, and the neighboring areas in western Pennsylvania pronounce *elementary* with an unstressed penultimate syllable (this is the normal pronunciation of this word throughout the rest of North America, too). The Upstate New York pattern appears to be unique to the region, at least in North America, and is also quite homogenous throughout most of the state (Dinkin 2009). The distribution of stress in *elementary* thus shows a clear boundary between Erie and the Midland, on the one side, and

Code	Stress pattern	Transcription
0	penultimate vowel deleted	[ˌɛləˈmɛntri]
1	penultimate vowel unstressed	[ˌɛləˈmɛntəri], [ˌɛləˈmɛntr̩i]
2	unsure or intermediate	e.g., [ˌɛləˈmɛnˌtəri]
3	penultimate vowel bears secondary stress	[ˌɛləˈmɛnˌtɛri]
4	penultimate vowel bears secondary stress; antepenultimate unstressed and reduced to schwa	[ˈɛləmənˌtɛri]

Table 8.2: Coding scheme used for lexical stress in *elementary*

Upstate New York, on the other.

The Upstate New York pronunciation of *elementary* has the following stress pattern: [ˌɛləˈmɛnˌtɛri]. This contrasts with the regular pronunciation in the rest of North America with an unstressed penultimate syllable, leading to pronunciations with either a schwa or a syllabic /r/ in the penultimate syllable, as in [ˌɛləˈmɛntəri] and [ˌɛləˈmɛntr̩i], or, most commonly, complete deletion of the syllable, as in [ˌɛləˈmɛntri]. Dinkin (2009) shows that the same stress pattern also affects other lexical items ending in *-mentary* in Upstate New York, thus leading to pronunciations such as [ˌsɛdəˈmɛnˌtɛri] for *sedimentary* and [ˌdakjəˈmɛnˌtɛri]. However, this study will only present results for *elementary*, since that was the only *-mentary* word included in the word list.[1]

Each speaker's word list pronunciation of *elementary* was provided with one of five codes representing the degree of stress on the penultimate syllable. The coding scheme, along with examples for each code, is presented in Table 8.2.

Figure 8.2 presents the results for the geographical distribution of the Upstate New York pattern of secondary stress in *elementary*. For this map, codes 0 and 1 were treated as qualitatively identical, since they both have an unstressed penultimate syllable. These

---

[1] The item *documentary* was added to a later version of the word list, but not enough speakers read this version for an analysis to be conducted.

codes were merged and are shown as red dots in Figure 8.2. Additionally, codes 3 and 4 were treated as identical, since both show secondary stress on the penultimate syllable. These codes, representing the Upstate New York pattern, were merged in Figure 8.2 as blue dots. (No pronunciations in my data set received a code of 2.)

Figure 8.2 shows that Erie clearly exhibits the normal widespread pronunciation lacking stress on the penultimate syllable—not a single speaker in Erie pronounced *elementary* with secondary stress on the penultimate syllable. On the other hand, the general Upstate New York pattern is robustly attested in Chautauqua County, NY. There, 18 out of 21 speakers produced tokens with secondary stress on the penultimate syllable.

The isogloss between the two regions appears to coincide closely with the state boundary. The only two speakers in Pennsylvania who exhibited the Upstate New York stress pattern live in Wattsburg and Union City, both located in Erie County just to the southwest of the border with New York. Interestingly, both of these speakers are older women: the speaker from Union City is 77, and the one from Wattsburg is 80. If younger speakers in Wattsburgh and Union City are found to have the normal unstressed pattern, then this could potentially indicate a change in progress away from the Upstate New York pattern.

According to the terminology in Chambers and Trudgill (1999:97), this boundary should not actually be referred to as a lexical isogloss, but rather a pronunciation isogloss. As they explain: "the former involves a difference in formatives from one dialect to the other while the latter involves a contrast in the phonemic representation of the same formative." An example of a lexical isogloss according to this terminology would be a boundary between the use of the terms *elementary school* and *grammar school*. However, I will continue to use the terminology *lexical variable* and *lexical isogloss* when referring to lexical stress assignment in *elementary* to distinguish this phenomenon from the phonological variables discussed in Chapter 6 and 7 that have deeper structural connections to other phenomena.

Chambers and Trudgill (1999:99) do say that lexical and pronunciation isoglosses are

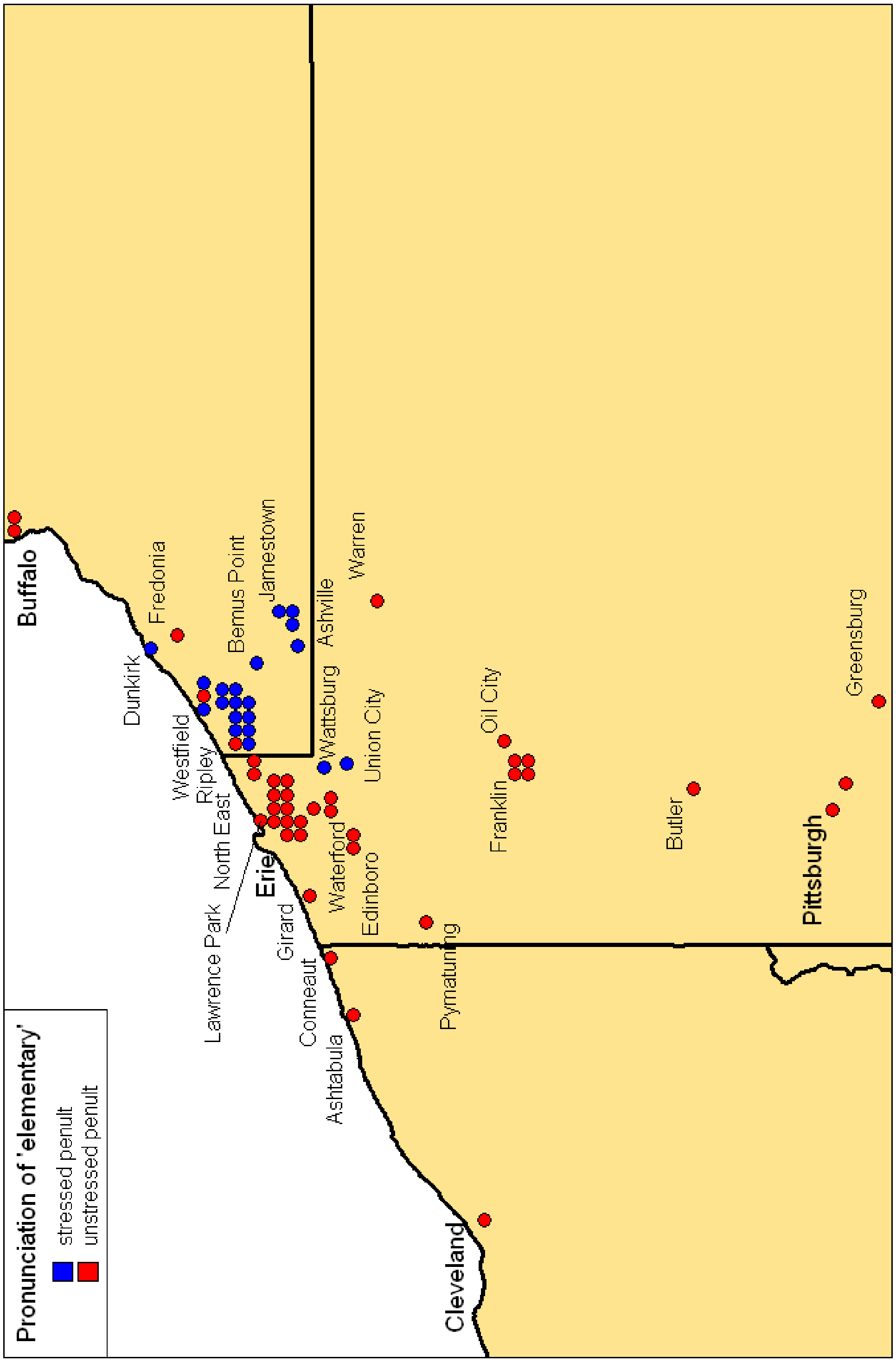

Figure 8.2: Lexical stress in *elementary*

quite similar from a structural standpoint, and group them together under the “lexical” heading when discussing the structural significance of different types of isoglosses. They also hypothesize that pronunciation variables are less likely than lexical variables to rise to the level of conscious awareness. Interestingly, evidence from one speaker that I interviewed shows that the stress variation in *elementary* is noticeable. This speaker grew up in Cleveland, OH, but currently resides in Findley Lake, NY (about halfway betwen Wattsburg and Ashville in Figure 8.2). This town is located in Chautauqua Co., about 15 miles east of the border with Erie Co., PA. As I was telling her about the purpose of my interviews and we began talking about regional pronunciation variation, the first thing she said was “Everyone around here says elementary (pronounced as [ˌɛləˈmɛnˌtɛri]), but I say elementary (pronounced as [ˌɛləˈmɛntri]).” At this point she had not seen the word list, and we had not previously discussed this variable at all. This shows that the Upstate New York stress pattern is clearly salient to speakers from other dialect regions.

8.3 *Redd up*

Another lexical item that shows a division between the Erie area and Chautauqua Co., NY is the verb *redd*, normally used in combination with the preposition *up* to form the phrasal verb *redd up*. Carver (1987:265) glosses *redd (up)* as “to clean or straighten (a room); to clean or clear off (a dinner table)”. My impression from eliciting judgments about *redd up* is that it is generally used to refer to a quick tidying up of a specific area, not a longer or more general cleaning. To illustrate this sense, one speaker from Erie said “Redd up is what you do to the room before guests come over.”

Published sources show that southwestern Pennsylvania is the area that exhibits the most concentrated use of *redd up*. McCool (1982:29) lists it as one of the features of Pittsburghese. Gooden and Eberhardt (2007:91–92) also cite *redd up* as a feature of Pittsburgh speech, but show that it use is restricted to White speech, as opposed to AAE. DARE describes its usage as "scattered, but chiefly North Midland, especially Pennsylvania" (Cassidy and Hall 2002:511): of the 97 DARE communities where "redd up" was attested in the questionnaire, 36 are located in Pennsylvania. Carver (1987:194) uses this distribution as the basis for including *redd up* as one of the 53 DARE isoglosses that comprise his Lower North (i.e., North Midland) layer.

Based on its geographical distribution in North America, it is quite likely that the use of *redd up* originated with Scots-Irish settlers (Montgomery 1997). However, it is difficult to demonstrate this direct path of transmission conclusively; since *redd up* also occurs in Scotland and northern dialects in England, its spread to North America could have been facilitated by other groups of immigrants using it as well (Crozier 1984:311). However, the fact that its isogloss seems to coincide well with the area of Scots-Irish settlement, and that speakers from other regions generally do not recognize it, strongly suggests that it was indeed brought over by Scots-Irish immigrants.

My own fieldwork in northwestern Pennsylvania and western New York show that the meaning of *redd up* is recognized by many speakers as far north as Erie. Also, several speakers from that region report that they themselves could use the phrase. On the other hand, only one speaker from Chautauqua Co., NY reported that she could use *redd up*, and only two others said they thought that they've heard other people use it. Most of them did not recognize the phrase and could only guess at what they thought it might mean. This geographic distribution is shown in Figure 8.3.

The distribution for *redd up* shown in Figure 8.3 corresponds well with the map for this item in DARE. In that map, Meadville (Crawford Co.) and Union City (Erie Co.) in Pennsylvania were both listed as communities where *redd up* was attested in the questionnaire (although it was not attested in North East, Erie Co.), whereas none of the communities in Chautauqua Co., NY were. This suggests that the current acceptability of *redd up* as far

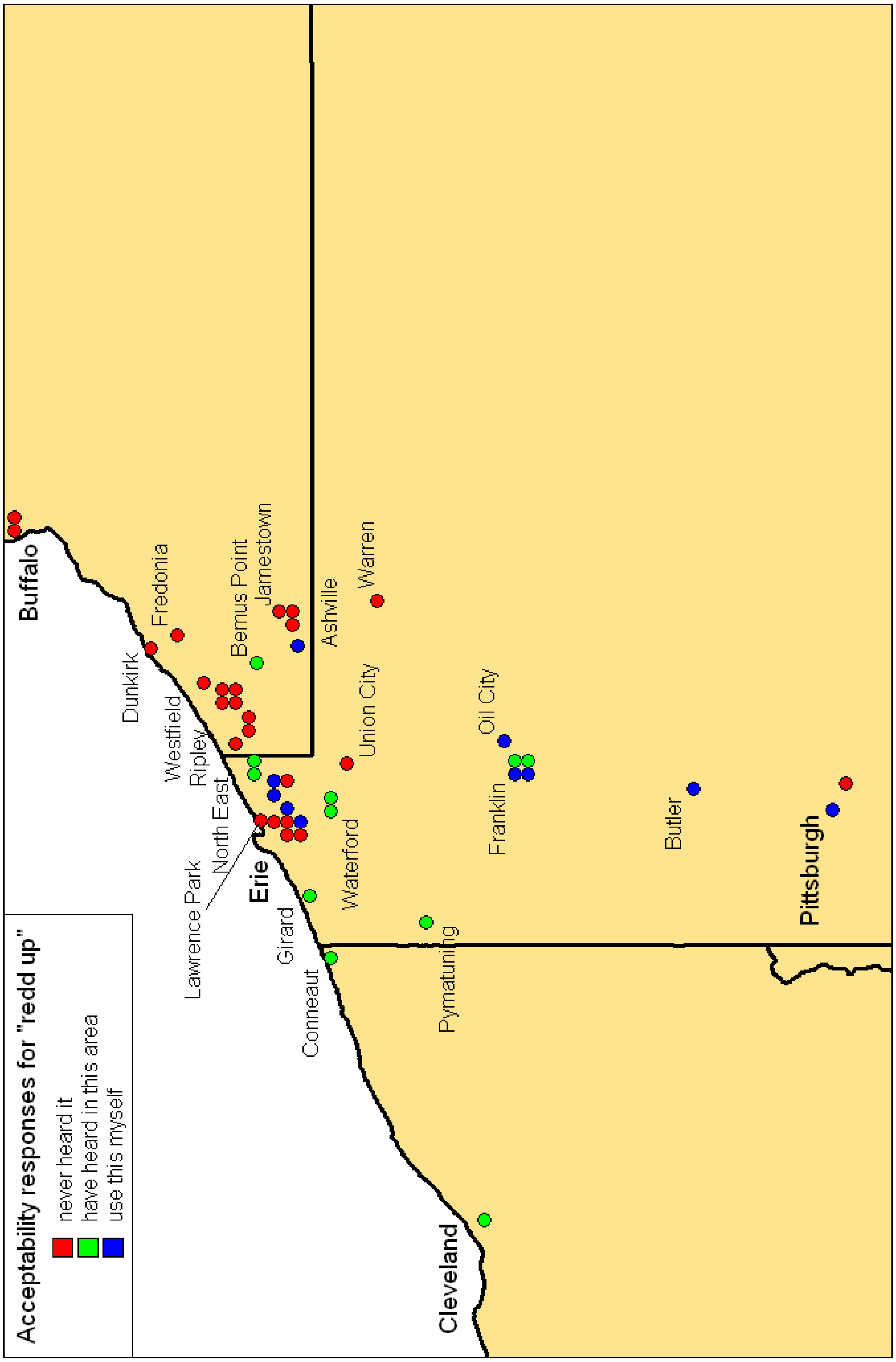

Figure 8.3: Acceptability responses for the sentence *I really should redd up the living room.*

north as Erie is a relatively old phenomenon, dating back to at least the early part of the 20th century, and is not necessarily due to recent influence from the Pittsburgh area. This could be one indication that the original linguistic situation in Erie was not identical to the neighboring Northern region, despite the evidence presented in Section 2.2. The presence of a substantial proportion of original Scots-Irish settlers in Erie (see Chapter 9) provides a possible explanation for this apparent contradiction.

8.4 Positive *anymore*

The use of *anymore* without an accompanying negative or question marker is widespread throughout the United States, and is a feature of most of the Midland region. Precise isoglosses for its geographic distribution throughout the country have not been determined, however, despite numerous studies. Two factors contribute to the lack of precision in our knowledge of the extent of its use. First of all, positive *anymore* is relatively infrequent in normal speech. Thus, targeted formal methods are required to elicit judgments about its use. However, introspective judgments about positive *anymore* have been consistently shown to disagree with actual usage (Labov 1973). Thus, any data obtained about positive *anymore* must be treated with caution.

Individual studies have shown positive *anymore* to be in common use in specific areas of the Midland, such as Missouri (Youmans 1986) and Southeastern Pennsylvania (Shields 1997), etc. Furthermore, three surveys with a larger geographic compass have investigated positive *anymore*, and their results show that its area of acceptance overlaps considerably with the Midland. Dunlap (1945) shows that positive *anymore* is most prevalent among his informants from Southeastern Pennsylvania and the neighboring areas of Delaware and Maryland, but is also widely attested in southern Illinois and Indiana. The isogloss in Labov et al. (2006:294) extends as far east as Philadelphia and as far west as Idaho, but stays south of the North / Midland boundary. The evidence from DARE (Cassidy and Hall 1985:73) also shows that positive *anymore* is concentrated most heavily in the Midland, showing especially high rates of use in Kentucky, Indiana, and West Virginia.

The fact that the use of positive *anymore* is strongest in the Midland corresponds well with the theory that it, too, originated in the speech of Scots-Irish immigrants (Dunlap 1945, Crozier 1984, Montgomery 2004). A Scots-Irish origin would also help to explain why positive *anymore* appears to be stronger in the South Midland, and does not extend as far north as the boundary between the North and the Midland in ANAE.

8.4.1 Examples from conversational speech

Since positive *anymore* is relatively infrequent in natural discourse, it is not possible to conduct a quantitative analysis of the occurences from the interviews I conducted. However, the small number of examples that did occur in the interviews provide clear evidence for the existence of positive *anymore* as a feature of Erie speech.

In total, I observed nine examples of positive *anymore* being used in natural discourse during my fieldwork. These nine examples came from six different speakers, all natives to the city of Erie. Examples 8.1 – 8.9 show these nine examples, demographic information about the speakers who produced them is presented in Table 8.3.

(8.1) That's the world's excuse to do anything anymore.

(8.2) A: It's amazing how much equipment kids need.
B: Oh, *anymore*.

(8.3) The way they strap these toys in anymore.

(8.4) I'm sure you've probably gone to a GNC or a health foods store or even a grocery store anymore and I mean it's crazy looking at all that stuff.

(8.5) It's so hard the way we build things anymore.

(8.6) I'm afraid to buy jewelry for her anymore.

Examples produced	Age	Gender	Town
8.1, 8.2	53	m	Erie
8.3	33	f	Erie
8.4	66	f	Erie
8.5	53	f	Erie
8.6	60	f	Erie
8.7, 8.8, 8.9	56	f	Erie

Table 8.3: Demographic information for the six speakers who produced positive *anymore*

(8.7) Any little town I go to anymore has a local espresso place.

(8.8) I don't know how long they take to score these things anymore. (about the SAT tests)

(8.9) Most of the planes I take seem fuller than that anymore.

The positive evidence exhibited by the nine naturally occurring examples of positive *anymore* produced by Erieites is a clear indication that positive *anymore* is a widely accepted feature of Erie speech. On the other hand, the negative evidence from other areas—i.e., the lack of examples of positive *anymore*—does not necessarily indiacte that positive *anymore* is not a feature of the speech of these regions, due to the construction's rarity in discourse.

8.4.2 Acceptability judgments for positive *anymore*

It is suggestive that the six speakers who spontaneously produced positive *anymore* all come from Erie, and that none of the speakers from Chautauqua Co., NY did so. However, in order to investigate the status of positive *anymore* in more geographic detail, it is necessary to obtain more data than just naturally occurring examples of the construction. In order to do this, speakers' judgments about the construction were elicited during the reading passage section of the interview. This survey used the methodology employed by

ANAE (Labov et al. 2006:293) for positive *anymore*, namely, a three-point forced choice scale. The speakers were presented with the following three sentences:

(8.10) Ticket prices are so high anymore, I never go to the movie theater.

(8.11) Anymore, there's too much crime in this neighborhood.

(8.12) John eats fast food so much anymore, it's no wonder that he's becoming overweight.

They were asked to rate the sentences as either 1) "I could say a sentence like this." 2) "I wouldn't say this, but I've heard people around here say something like it." or 3) "I've never heard anything like this before—it sounds like bad English." Examples 8.10 and 8.11 express complaints, which Labov et al. (2006:293) have argued to be the most natural pragmatic context for positive *anymore*. Example 8.11 has preposed positive *anymore*, which is generally judged to be less acceptable. As a control case, the survey also included the sentence in Example 8.13. This sentence contains *anymore* in a negative context, and should be judged to be perfectly natural by all speakers.[2]

(8.13) I was a pitcher when I was young, but now I don't play baseball anymore.

Table 8.4 shows the mean response values for all 49 speakers who took the survey. First of all, the mean response for the control sentence was 1.2, indicating that nearly all speakers judged this sentence to be perfectly acceptable, as was expected. The overall results for the three positive *anymore* sentences confirm that pre-posed *anymore* is less acceptable, and that framing the construction in a complaint speech act makes it more acceptable. The sentence with pre-posed *anymore*, Example 8.11, received the lowest overall response. Among the other two sentences with non pre-posed *anymore*, the one expressing a complaint, Example 8.10, received a higher overall response. However, the overall responses

---

[2] One informant did provide a rating of 3 for Example 8.13. This speaker, in fact, provided a rating of 3 for all of the eight sentences in the acceptability judgment portion of the survey. His responses to this portion of the survey were deemed unreliable and discarded.

	Ex. 8.10	Ex. 8.11	Ex. 8.12	Ex. 8.13
Mean acceptability response	1.9	2.2	2.1	1.2

Table 8.4: Mean response values (on a scale of 1 – 3) for acceptability judgments on three positive *anymore* sentences (Examples 8.10 – 8.12) and one control sentence (Example 8.13)

are all quite close, indicating that these effects are not very strong relative to the overall effect of the presence of positive *anymore*.

Figures 8.5 through 8.7 display the geographical locastions of the responses to the three sentences with positive *anymore*. Several Erieites responded that these sentences were something that they could say or that they have heard people around them saying. Taken along with the attested examples from Erieites of positive *anymore* in conversational speech presented in Section 8.4.1 this provides further evidence that Erie patterns with the Midland with regard to this feature.

It is not possible, however, to draw isoglosses in Figures 8.5 through 8.7 between an area where positive *anymore* is clearly acceptable and an area where it is not. The figures show that several speakers from Chautauqua County, NY also judged these three sentences as acceptable. Furthermore, the speaker from Butler, PA, judged two of them to be unacceptable. Butler is located just 40 miles north of Pittsburgh, and, based on previous studies, would thus be expected to have widespread acceptability of positive *anymore*. These results thus must be interpreted somewhat cautiously, and it must be remembered that speakers of positive *anymore* often do not have accurate introspective judgments on their own usage of the construction (Labov 1973).

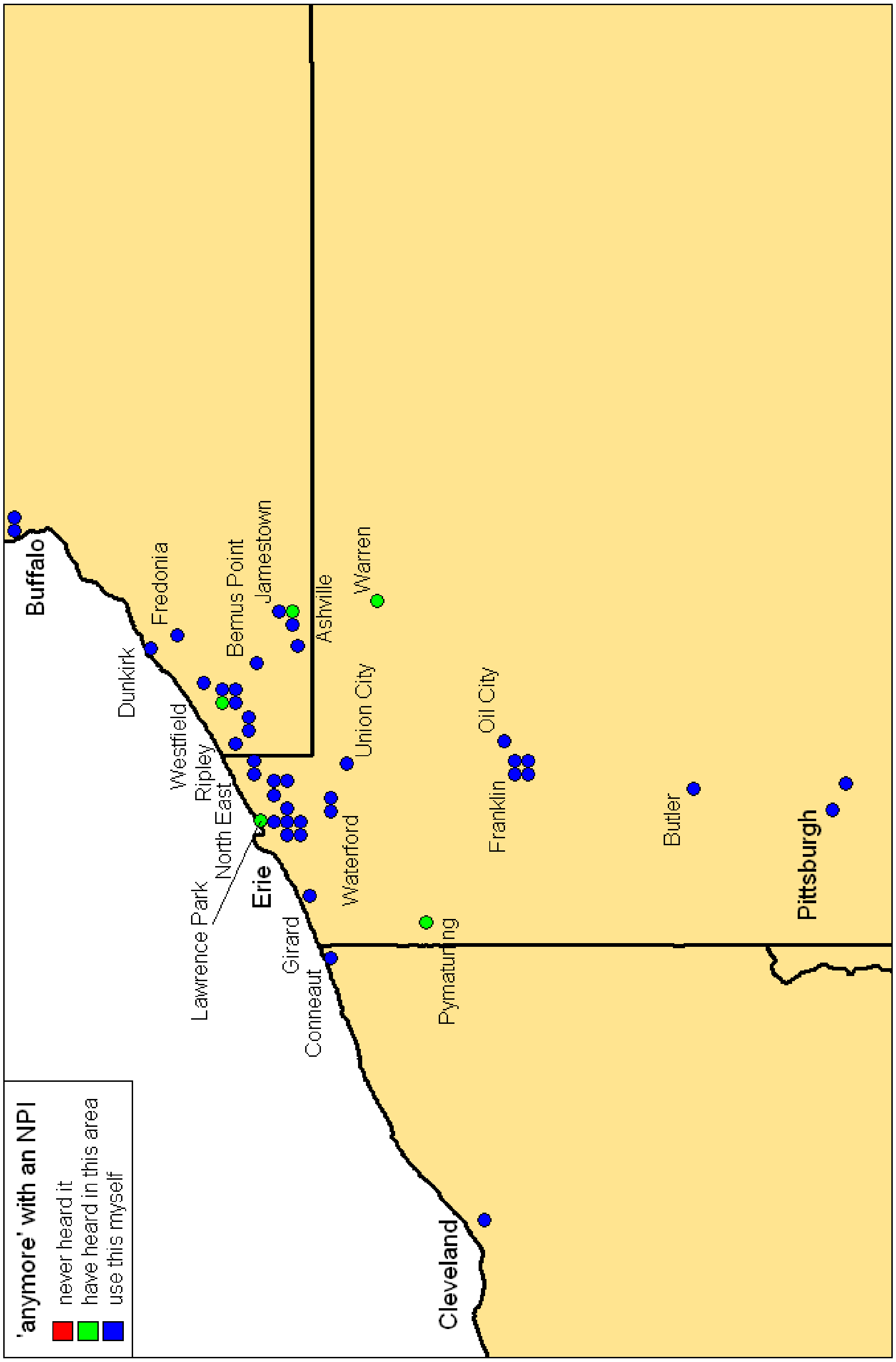

Figure 8.4: Acceptability responses for the sentence *I was a pitcher when I was young, but now I don't play baseball anymore.*

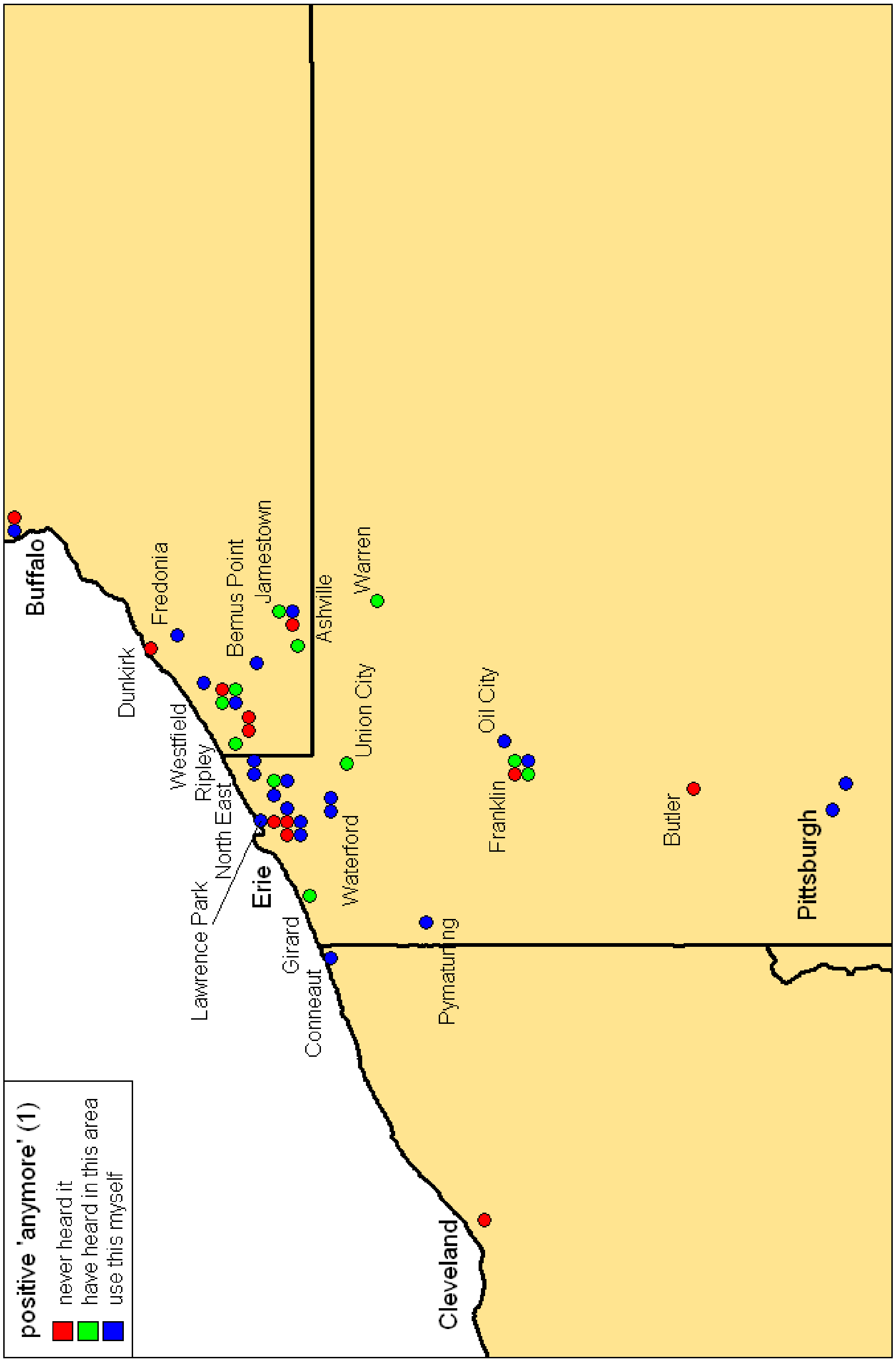

Figure 8.5: Acceptability responses for the sentence *Ticket prices are so high anymore, I never go to the movie theater.*

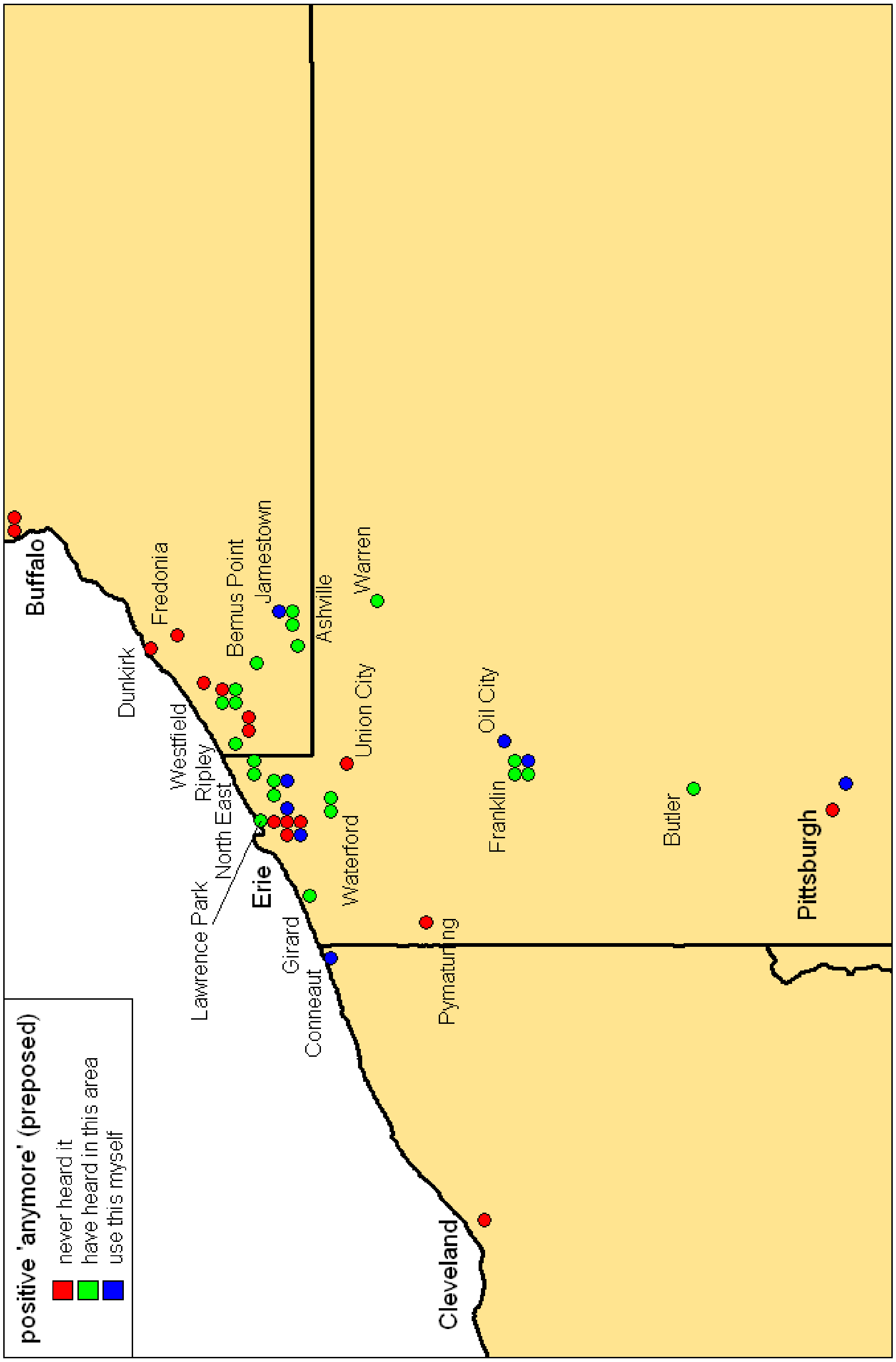

Figure 8.6: Acceptability responses for the sentence *Anymore, there's too much crime in this neighborhood.*

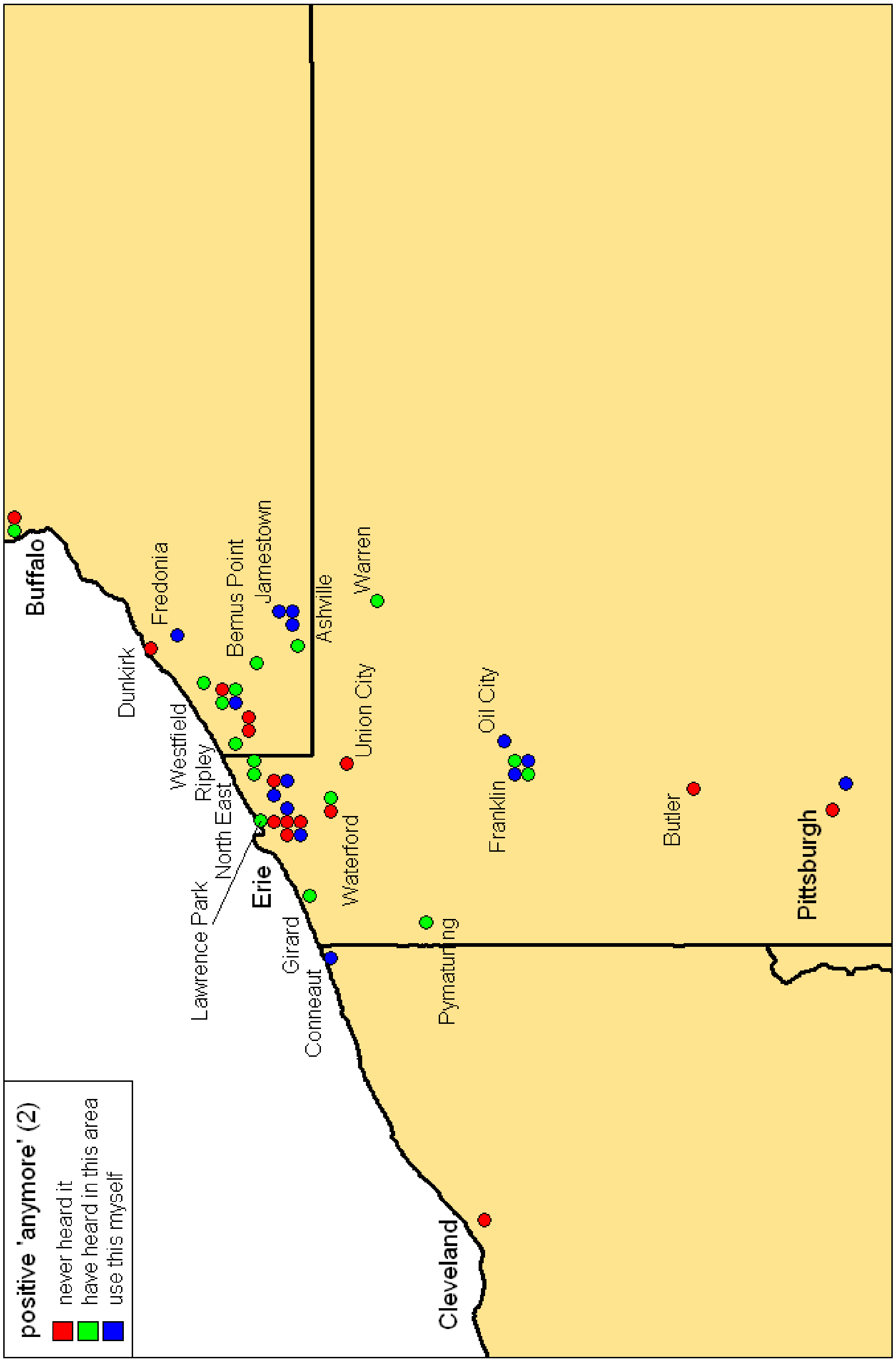

Figure 8.7: Acceptability reponses for the sentence *John eats fast food so much anymore, it's no wonder that he's becoming overweight.*

8.4.3 Discussion

There are two possible explanations for the current acceptability of positive *anymore* in and around the city of Erie. On the one hand, it could reflect a northward spread of the Midland system, and, thus, influence of Pittsburgh and the neighboring areas of Western Pennsylvania. On the other hand, it could be the case that positive *anymore* has always been a feature of Erie speech. The only potential negative evidence for the earlier existence of positive *anymore* in Erie is a single informant from the survey in Dunlap (1945) who marked the construction as “unfamiliar”. However, a DARE informant from Union City in Erie County did use *anymore* in response to the elicitation prompt: *People used to walk a lot, but everybody drives a car* ________ . This positive attestation near Erie suggests that positive *anymore* is probably not a recent addition to the area. As was the case with *redd up*, the acceptability of positive *anymore* in the Erie area may be attributable to the early presence of Scots-Irish settlers in the region.

8.5 *need* + Past Participle

A second Midland grammatical feature that is also attested in the Erie area is the use of *need* + Past Participle (V-*en*), as in *The car needs washed.* This use contrasts with the use of *need* + Present Participle (V-*ing*) in other dialect regions, as in *The car needs washing*. The full version containing *need to be* + Past Participle, as in *The car needs to be washed*, is acceptable in all areas.

The earliest source that mentions the geographic distribution of this feature is Stabley (1959) who reports attestations for *need* + V-*en* in several towns in western Pennsylvania: “Many western Pennsylvanians—educated as well as uneducated—often declare that the house *needs painted* or the television set *needs fixed* or the children *need spanked.* Certain radio and television announcers from Indiana, Johnstown, and Pittsburgh employ this construction, as do some newspaper writers and ministers...As an eastern Pennsylvanian, I have met this usage only in the Allegheny Mountain region of the state; wide inquiries yield no evidence of its currency elsewhere in the land.” This quote is instructive, since it mentions both the fact that the use of *need* + V-*en* is not restricted by the speaker’s social class, and the fact that it is not affected heavily by stylistic variation, since it is attested in printed sources. These two characteristics of *need* + V-*en* are reported as well in larger, more recent studies (Murray et al. 1996). However, the assertion that *need* + V-*en* is restricted solely to western Pennsylvania has been contradicted by more recent research. DARE’s entry for the construction says that it is “chiefly Midland, especially Pennsylvania”. Murray et al. (1996) demonstrate that its use is widespread throughout most of the Midland and limited in other regions. ANAE also shows that the geographical range of *need* + V-*en* coincides well with the Midland region, although its range is somewhat smaller than that of positive *anymore*—the isogloss of the former is almost wholly surrounded by the isogloss of the latter (Labov et al. 2006:295). Of the two ANAE speakers from Erie, one reported that she herself uses *need* + V-*en*, and the other reported hearing people in the area use it.

As is the case for positive *anymore* (Labov 1972:309), the alternation between *need* + V-*en* and *need* + V-*ing* operates below the level of consciousness for most speakers (Murray et al. 1996), and can thus be appropriately studied with a written questionnaire. In order to track the northern extent of the use of *need* + V-*en*, a forced-choice sentence completion task between *need* + V-*en* and *need* + V-*ing* was included in the written portion of my survey. The two sentences are reproduced as Examples 8.14 and 8.15:

washed.
(8.14) I drove through a big, muddy puddle yesterday. Now my car needs
washing.

mopped.
(8.15) I haven’t cleaned my kitchen in weeks. The floor really needs
mopping.

By using a forced-choice task instead of a sentence completion task, I was trying to maximize the percentage of relevant responses. Also, by not including an option for the construction *need to be* + V-*en*, the informant is forced to provide a response that unambiguously indicates whether or not they can use *need* + V-*en* This assumes, of course, that one, and only one, of the two constructions *need* + V-*en* and *need* + V-*ing* is grammatical for any given speaker. A few speakers did in fact respond that they did not like either of the two choices for completing Examples 8.14 and 8.15 and instead wrote in *needs to be washed* and *needs to be mopped*. The responses for these speakers were discarded from the analysis.

In addition to the forced-choice task between *need* + V-*en* and *need* + V-*ing*, the written portion of the survey elicited acceptability judgments for one sentence with *need* + V-*en*:

(8.16) I got into an accident last week, and now my front bumper needs repaired.

Figures 8.8 and 8.9 present the results for the forced-choice task in Examples 8.14 and 8.15, and the responses Example for 8.16 are presented in Figure 8.10.

In addition to the use of *need* + V-*en*, Murray and Simon (1999) and Murray and Simon (2002) have shown that a similar use of the Past Participle exists with the verbs *want* and *like*. These uses are also confined to the Midland region, and are both substantially more restricted than the use of *need* + V-*en*. Their research into the three constructions shows an implicational scale of acceptability such that if a speaker accepts *like* + V-*en* they will also accept *want* + V-*en*; similarly, if they accepts *want* + V-*en*, they will also accept *need* + V-*en*. Their maps for want + V-*en* and like + V-*en* show a heavy concentration of positive attestations in Western Pennsylvania around the Pittsburgh area. In order to determine whehter these constructions are also acceptable as far north as Erie, my survey also elicited acceptability judgments for the sentences in Examples 8.17 and 8.18:

(8.17) My cat looks really hungry. I think he wants fed.

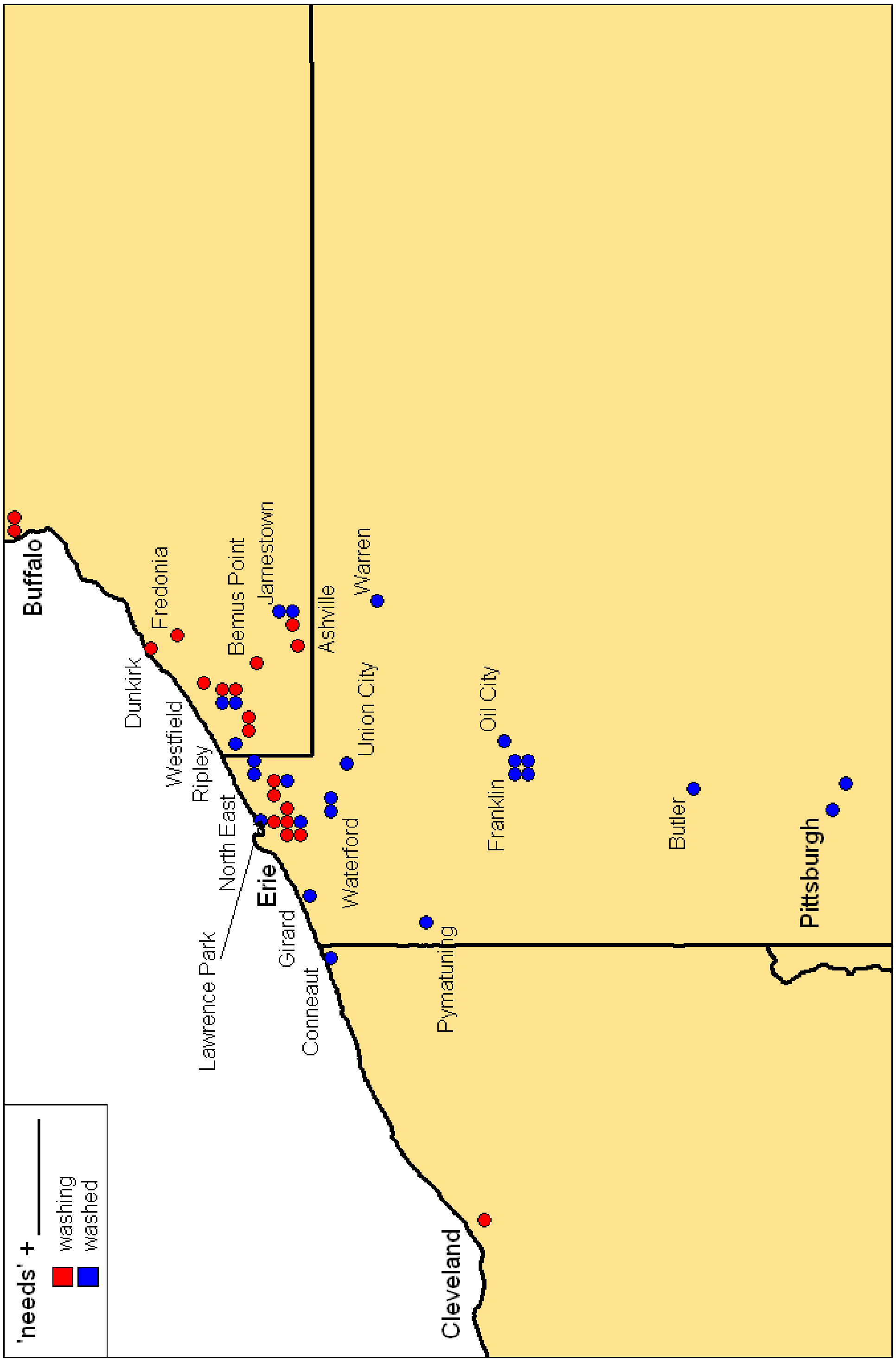

Figure 8.8: Responsed for the forced-choice completion task for the sentence *I drove through a big, muddy puddle yesterday. Now my car needs ______.*

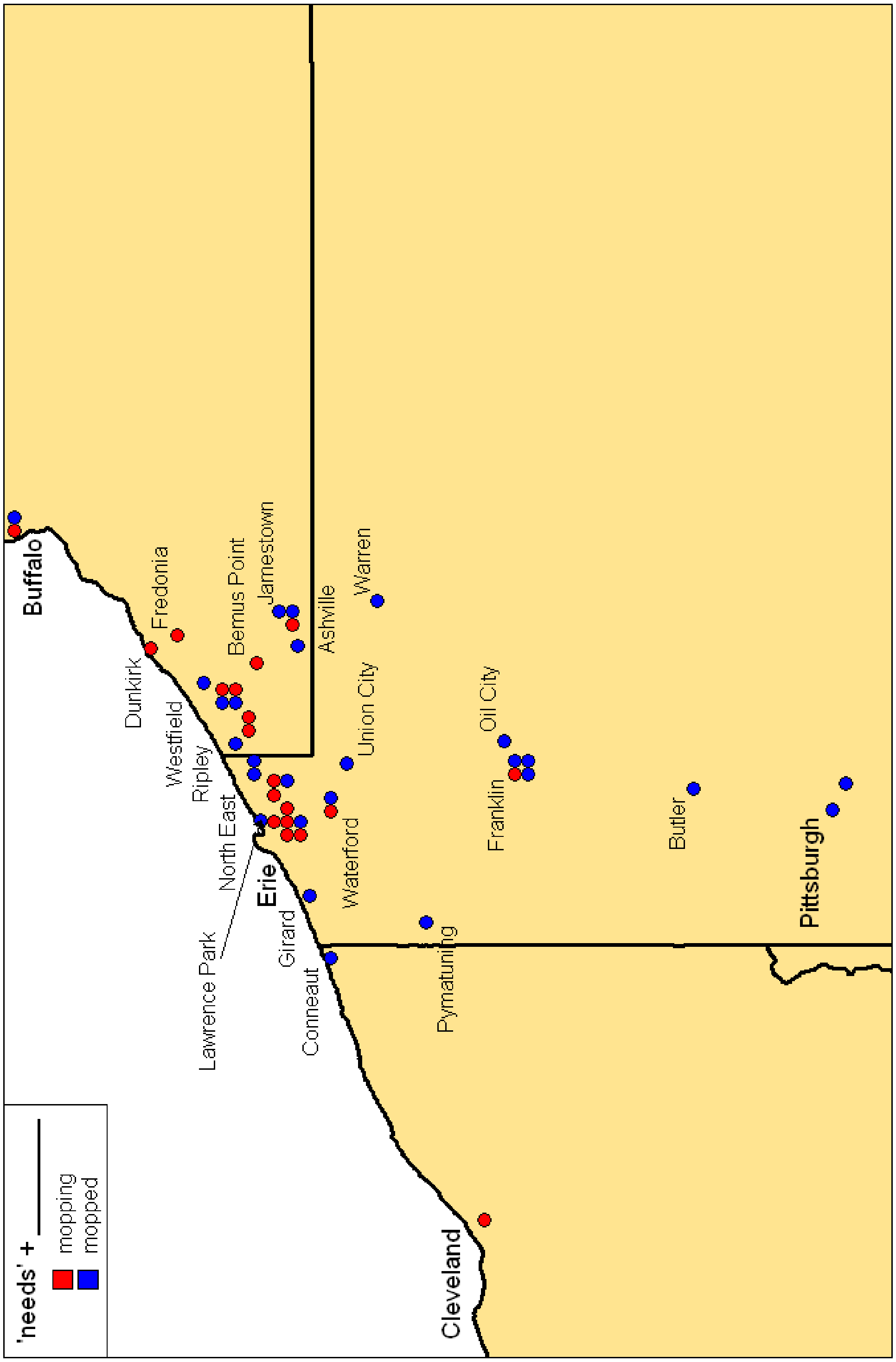

Figure 8.9: Responsed for the forced-choice completion task for the sentence *I haven't cleaned my kitchen in weeks. The floor really needs ______.*

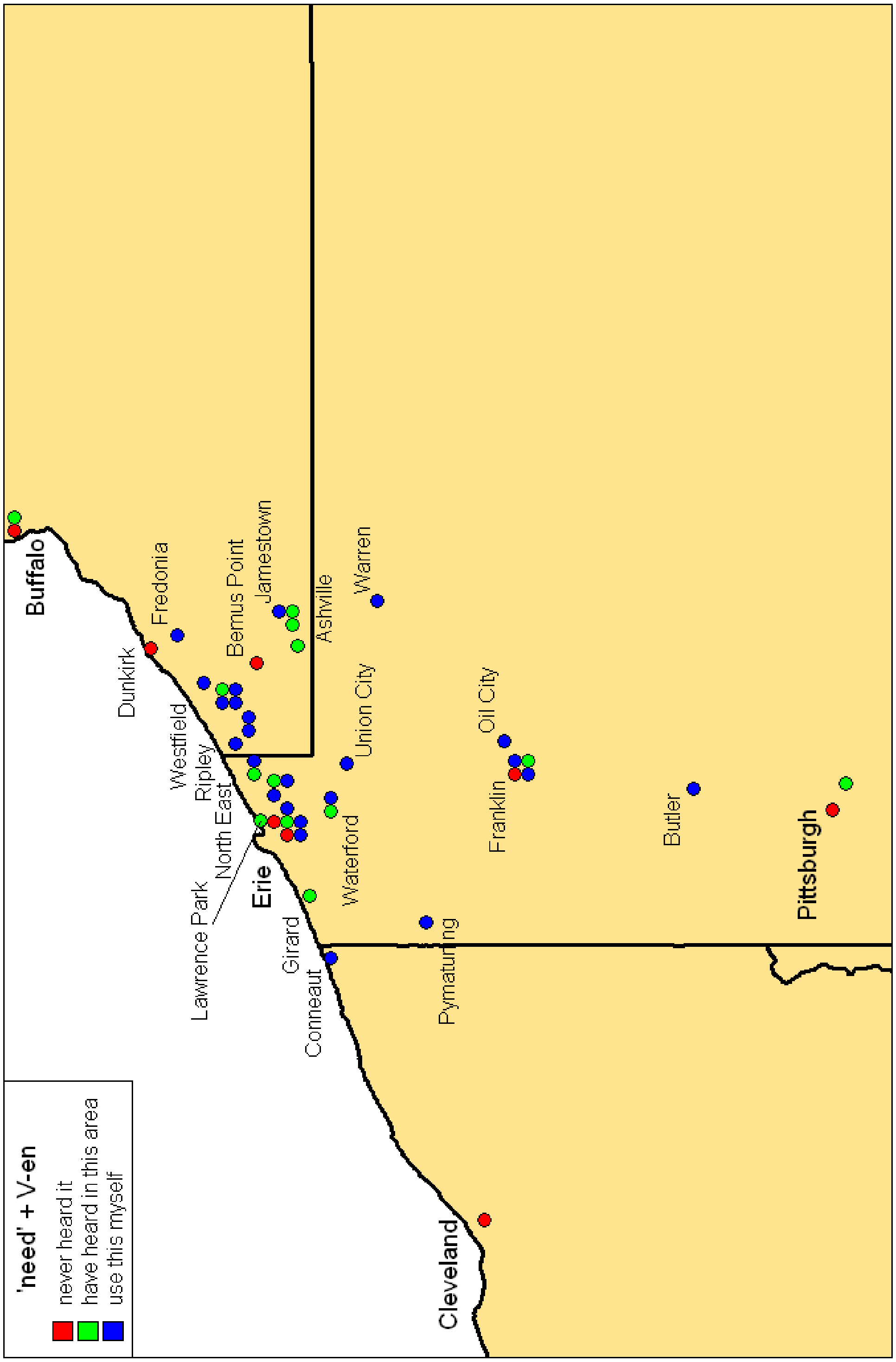

Figure 8.10: Acceptability responses for the sentence *I got into an accident last week, and now my front bumper needs repaired.*

(8.18) Every newborn baby likes cuddled.

Figures 8.11 and 8.12 present the results for the sentences in 8.17 and 8.18.

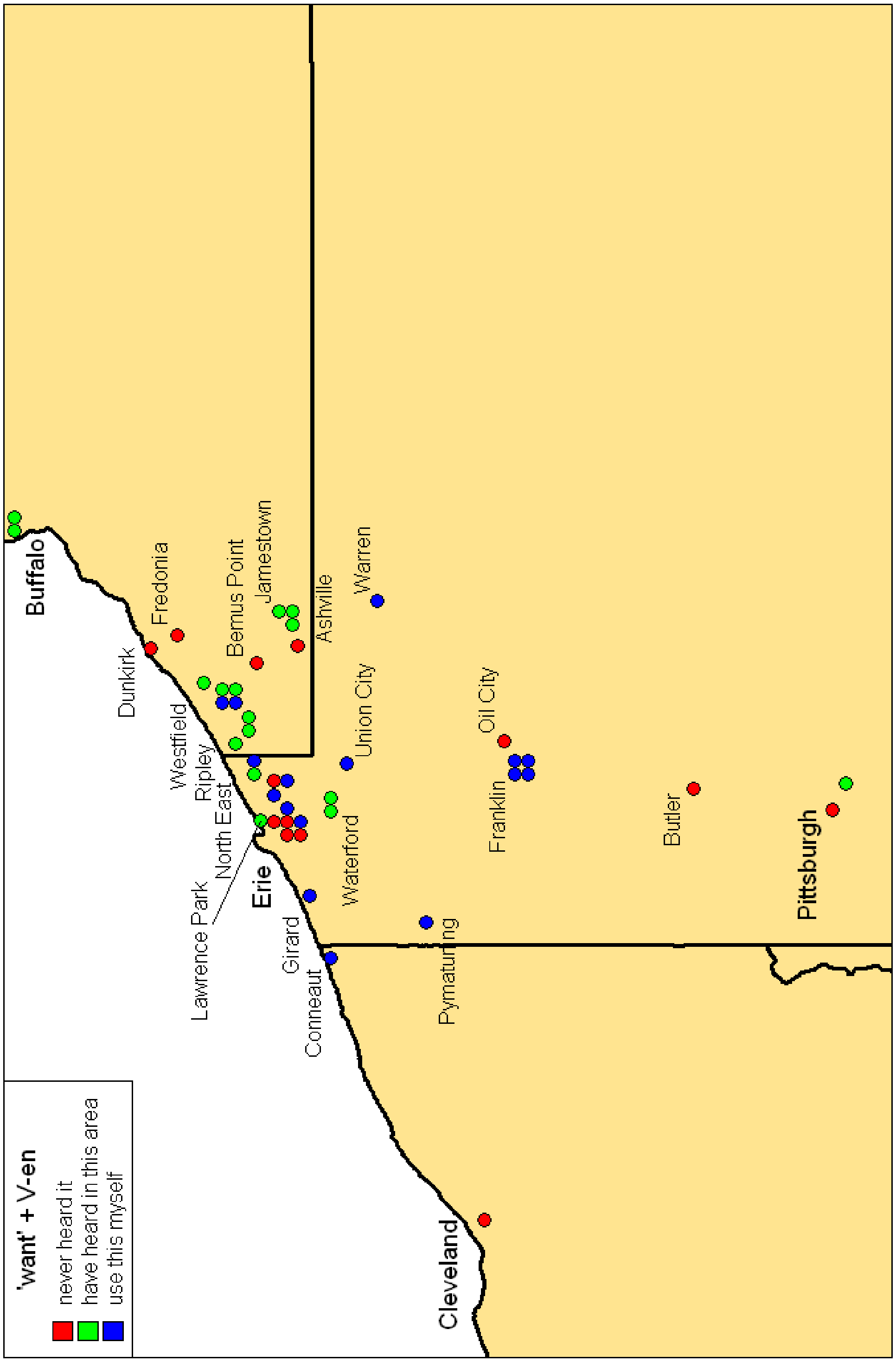

Figure 8.11: Acceptability responses for the sentences *My cat looks really hungry. I think he wants fed.*

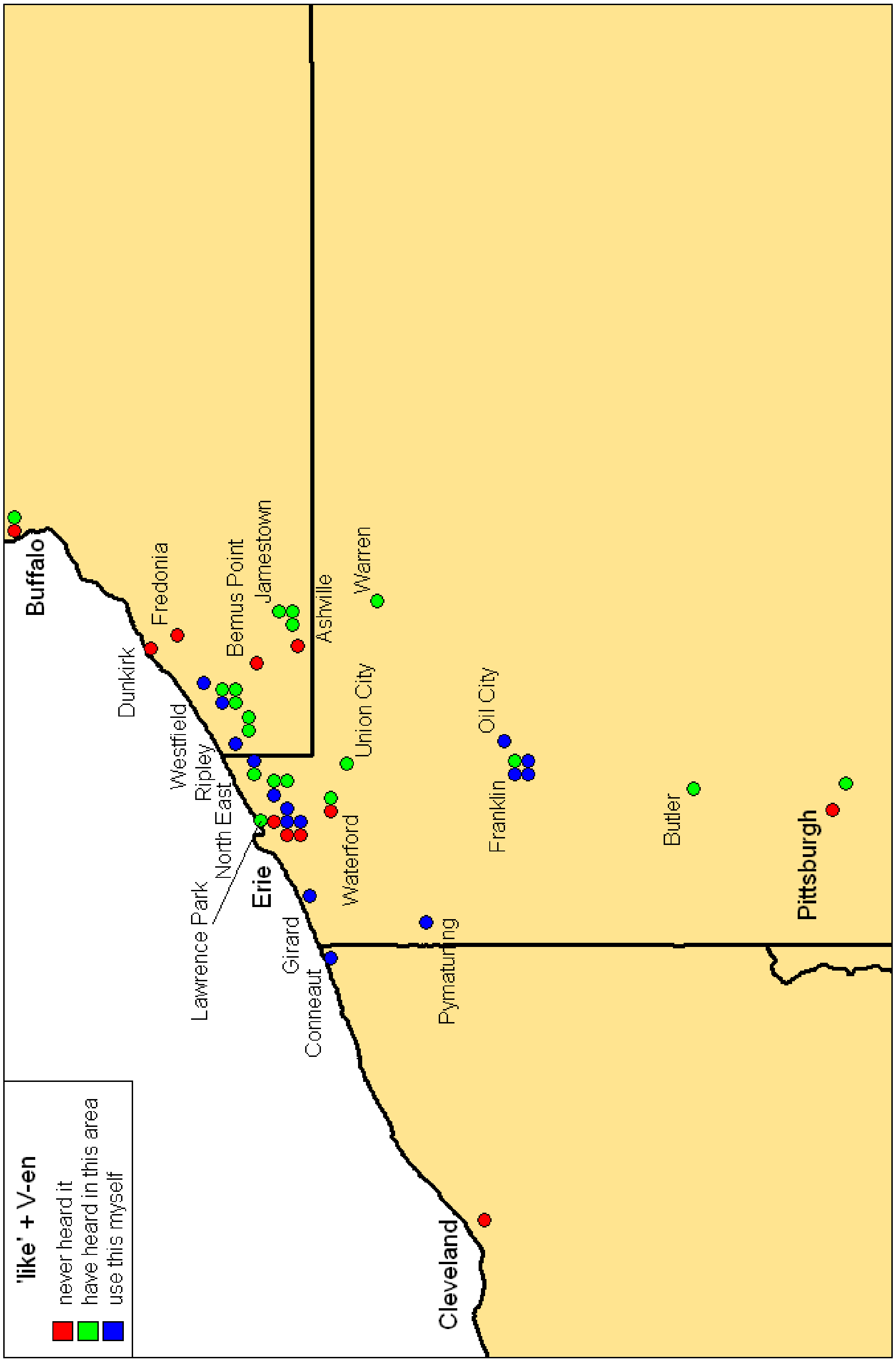

Figure 8.12: Acceptability responses for the sentence *Every newborn baby likes cuddled.*

Chapter 9

Settlement Patterns

The first recorded inhabitants of the coastal area along Lake Erie were the Erie Indians.[1] Their tribe was defeated in battle in 1654 by the Senecas, who killed many of the Eries and scattered those remaining alive among neighboring tribes. The Senecas remained the main inhabitants of the area up until the turn of the 19th century. The British and French each established forts in the area, and were vying for control of this strategically important bridge between the eastern and western settlements. However, neither the French or the British established non-military settlements.

American control over the area started in 1784 when Pennsylvania acquired the rights to the land through a treaty with the Six Nations. The British did not leave their military forts immediately, but were out-maneuvered by the American military who were able to establish alliances with the Senecas. The first American settlers arrived in 1795, aided by inexpensive land grants through the Pennsylvania Population Company. However, relations with the Senecas cooled drastically as American settlers began to move in, and the frequent violent raids discouraged settlement in the first years. However, military reinforcements were quickly sent to the area, and they dealt harshly with the hostile Senecas. This paved

---

[1] The brief summaries in this section are drawn from more detailed descriptions in Sanford (1894), Muller (1991), and Lechner (1994).

the way for a rapid increase in the population in the early 19th century.

Erie was incorporated as a borough in 1805, and experienced a sizeable growth in population throughout the 19th century. Table 9.1 shows the population growth in the city of Erie and Erie County for this time period (Sanford 1894, FWP 1938).

Year	City Pop.	County Pop.
1800	81	1,468
1810	394	1,358
1820	635	8,553
1830	1,329	17,041
1840	3,412	31,344
1850	5,858	38,742
1860	11,113	49,697
1870	15,516	
1880	27,737	
1890	40,634	

Table 9.1: Population growth in the city of Erie and Erie County, 1800–1890

In considering the effect of settlement patterns on the subsequent linguistic system in Erie, it is necessary to consider the geographical origins of the early settlers. It has been known for some time among cultural geographers that a small group of the earliest settlers in a region can have a profound and lasting impact on the culture of the region. This was formulated clearly by Zelinsky (1973:13–14) as the Doctrine of First Effective Settlement:

> Whenever an empty territory undergoes settlement or an earlier population is dislodged by invaders, the specific characteristics of the first group able to effect a viable, self-perpetuating society are of crucial significance for the later social and cultural geography of the area, no matter how tiny the initial band of settlers may have been...Thus, in terms of lasting impact, the activities of a few hundred, or even a few score, initial colonizers can mean much more for the cultural geography of a place than the contributions of tens of thousands of new immigrants a few generations later.

Recent studies have also shown that this is true for linguistic structure: Mufwene (1996) formulated a similar idea which he called the Founder Principle, and demonstrated its applicability to creole genesis; Labov (2007) provides an explanation of the diffusion of the

New York City short-/æ/ system to Cincinnati based on early settlement data; and Dinkin (2008) correlates the early preponderance of Dutch settlers in Amsterdam and Oneonta in eastern upstate New York to the fact that these two towns are not participating in the Northern Cities Shift.

As Chapter 2 has shown, the earliest linguistic evidence indicates that Erie originally patterned with the North. Thus, we would expect a large portion of the earliest settlers to have arrived from sources similar to the ones that settled nearby cities in the North such as Buffalo and Rochester. Published sources indicate that there were two main sources of early settlers to Erie County: New England and Southeastern Pennsylvania. The New Englanders who arrived in Erie were for the most part of British origin, and came to Erie through New York state primarily from Massachusetts and Connecticut. On the other hand, the settlers arriving from Southeastern Pennsylvania were of either Scots-Irish or German descent.

However, published sources do not provide the information necessary to account for the early linguistic patterns, namely, the proportion of the two groups among the early settlers. FWP (1938:23) does state that most of the settlers prior to 1800 came from New England and New York, and that subsequent migrations were also from the same sections. However, the authors provide no specific data to support this claim. On the other hand, Sanford (1894) claims that the early settlers were a mix of New Englanders and Scots-Irish: "the first settlers in Erie County were mostly...from moral, thrifty, intelligent New England; or...perhaps a more numerous class, of the illustrious, historic race of Scotch-Irish." However, she also does not provide any specific numbers to support this claim.

Two primary sources were consulted in an attempt to document the early settlement history of the region more accurately: a publication documenting the location of the burial sites for all of the "Revolutionary Patriots" interred in Erie County, and an early history of the county with biographical information for the prominent early settler.

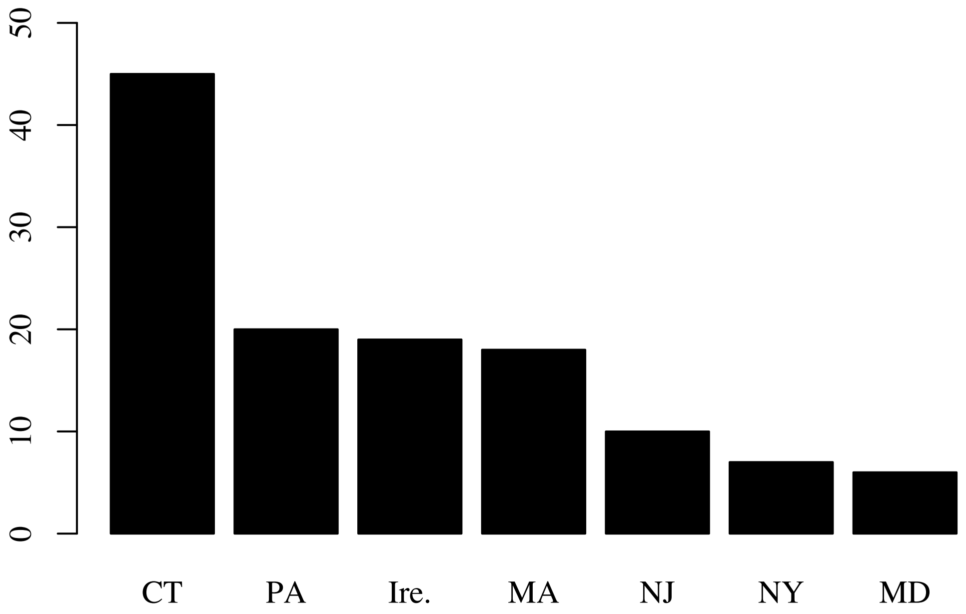

Figure 9.1: Place of birth of Revolutionary Patriots from Erie County

The first source is entitled *Revolutionary Patriots in Erie County, PA*, and was published by the Sons of the American Revolution as an aid in genealogy. The National Association of the Sons of the American Revolution defines a Revolutionary Patriot as an individual who has given "acceptable service to this nation." Examples of such service include (but are not limited to) such actions as: signing the Declaration of Independence, serving in the Revolutionary army between April 19, 1775 and Nov 26, 1783, serving in the Continental Congress, etc. The Erie Chapter of the Sons of the American Revolution compiled a list of all known Revolutionary Patriots who resided in and were buried in Erie County. Out of the 212 such Patriots, the place of birth is known for 143. Figure 9.1 summarizes this information by listing the most common birth places for the Revolutionary Patriots. By far the most common place of birth was the state of Connecticut, which claimed 31% of them (45 out of 143). Among foreign countries, Ireland was the only one with a sizeable number of Revolutionary Patriots, with 19 (England and Germany produced 2 each, Holland and Wales produed 1).

When the data in Figure 9.1 are grouped into three representative regions (North, Mid-Atlantic, and Europe), it is clear that settlers from the North far outnumbered those from other areas. For these purposes, the North is defined as New England plus New York state, and the Mid-Atlantic region includes Pennsylvania, New Jersey, and Maryland. Grouped together this way, the majority of the Revolutionary Patriots came from the Northern states, 57.3%, while a sizeable minority came from the Mid-Atlantic states, 25.2%. 17.5% of the Revolutionary Patriots came from Europe (as mentioned above, nearly all of these were from Ireland).

The average year of birth for all of these Revolutionary Patriots is 1754, and the average year of death is 1833. Thus, this group must have been among the very first settlers in the region, since Erie's first non-indigenous settler arrived in 1795.

A further source on the settlement of Erie County comes from a history of the county written in 1884, whose second volume contains biographies of 1,077 important residents of the county at the time of writing (Sanford 1894). Of these, 615 were born outside of Erie County, and can thus provide evidence for the geographic origin of a larger group of early settlers than the Revolutionary Patriots. The places of birth for these 615 prominent Erie County residents are presented in Figure 9.2 (excluding those locations that are represented by 10 or fewer settlers).

The data from this group of prominent residents represents a later stage of immigration than the Revolutionary Patriots data—the average year of birth for the residents considered in the 1884 book is 1825, a good three generations later than the earlier group. However, the proportion of Northern settlers to Mid-Atlantic and European ones remains quite sim-

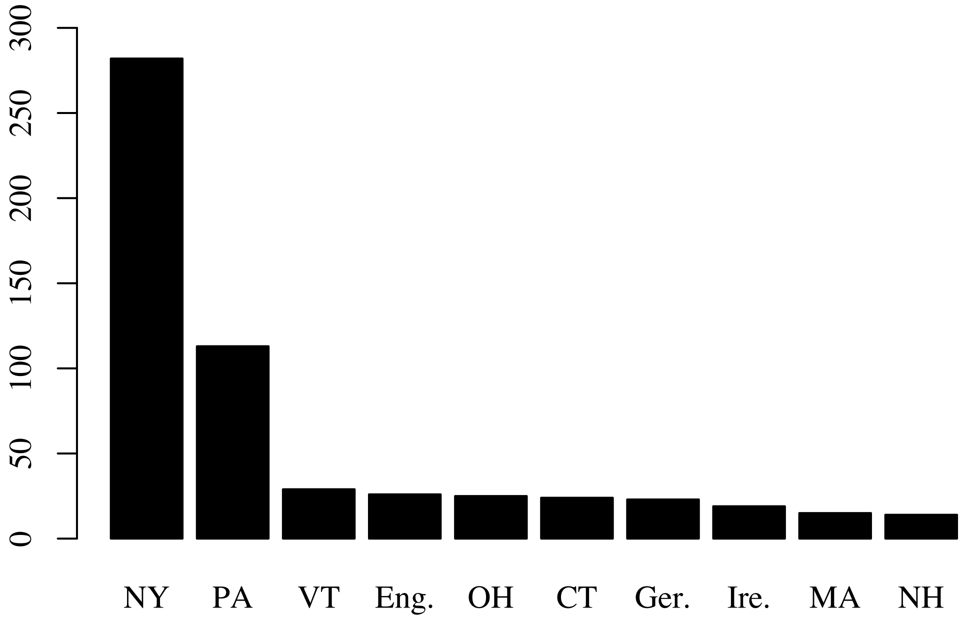

Figure 9.2: Place of birth of prominent 19th century Erie County residents

ilar: 60.1% to 19.9% to 13.0%, respectively. Again, the North contributes the majority of settlers, but a sizable portion also comes from Pennsylvania. The Northern settlers in this later group are dominated by those born in New York (as opposed to Connecticut for the Revolutionary Patriots). This reflects the massive westward migration from New England into New York state in the early / mid-nineteenth century, due, in large part, to the opening of the Erie Canal in 1825.

These two sources of evidence for the early settlement history of Erie (the Revolutionary Patriots and the list of prominent Erieites contained in Sanford (1894)) thus agree in locating the area of origin for the majority of the settlers in the North. The early Northern lexical and phonological patterns exhibited by the speakers from Erie County in PEAS (see

Chapter 2) are thus a result of this settlement history. On the other hand, both sources also demonstrate that Erie County has always had a sizeable contingent (greater than 40%) of non-Northern settlers, especially the Scots-Irish, often via other original settlement locations in Pennsylvania. This mixed early settlement history likely explains Erie's receptiveness to non-Northern linguistic features.

Under this explanation, Erie was never completely a Northern city, as an initial examination of the evidence from Kurath and McDavid (1961) and Kurath (1949) would suggest. As was discussed in Section 6.10 the two LAMSAS speakers who provided the data from Erie County for these two atlases were from small farming communities near the border with New York. It is quite likely that their speech was not representative of the speech in the county as a whole, especially the urbanized portion. The archival evidence from H.O. Hirt and the Seasonal Workers in Viticulture oral history project support this: the merger of /o/ and /oh/ spread to the city of Erie before it spread to the small, rural communities elsewhere in the county.

Two distinct groups thus co-existed in Erie County during the first portion of its history, coinciding roughly with the 19th century. From the onset, the contingent of merged settlers was likely large enough that their children did not aquire the distinction. This corresponds to Stage 2 of the Migration Hypothesis proposed by Johnson (2007:425), under which the proportion of natively-merged children entering the peer group is high enough that children entering school meet enough merged peers to be able to retain it (and not acquire the distinction from their unmerged peers). The proportion of merged Erieites then eventually grew large enough towards the end of the 19th century that natively-unmerged children came into contact with enough merged children in their peer group that they began to acquire the merger as well. This situation corresponds to Stage 3 of the Migration Hypothesis (Johnson 2007:426). The proportion of early Erie settlers originating from merged regions is consistent with the empirical data provided by Johnson (2007:436) for the spread of the merger to Seekonk and South Attleboro, MA. Furthermore, it is also greater than the figure of 22% suggested by a statistical model based on population structure to be the minimum necessary to cause the spread of the merger (Yang 2009).

This hypothesis thus proposes that the settlement history of Erie was qualitatively different than the settlement history of other Northern cities, such as Buffalo and Cleveland. Those cities almost certainly had larger percentages of their original settlers from the North, and smaller percentages from the Midland. Crucially, the early contingent of Scots-Irish settlers in these other cities is hypothesized to be much smaller. Further archival research into the settlement histories of these cities will be necessary to confirm this hypothesis.

Chapter 10

Conclusion

The goals of this dissertation were threefold: empirical, methodological, and theoretical. This final section will briefly summarize how these three goals were met, and discuss a few areas where further research is warranted.

First of all, the methodological aim of producing a valid sociophonetic analysis of vowels through completely automated means was achieved. Using the technique of transcription and subsequent forced alignment, I was able to quickly create a time-aligned corpus in which examples of all words and phonemes of interest can be extracted quickly and easily. This methodology is easy to implement, and the necessary software is freely available. The fields of sociophonetics, sociolinguistics, and dialect geography would benefit greatly from the adoption of the methodology of forced alignment, since it would enable the analysis of much larger amounts of data. With the continual increase in publicly available speech corpora, the widespread use of forced alignment and automatic phonetic analysis would bring about an dramatic increase in the types of analyses that could be conducted and the number of speakers that could be considered.

As Chapter 4 showed, however, several error-reduction techniques need to be applied concomitantly with a fully automated acoustic analysis in order to ensure that the results are comparable to a careful manual analysis. Further research in this area will hopefully standardize and improve upon these techniques. A particularly fruitful line of research will be the use of robust statistics for error exclusion and outlier detection.

On the empirical side, this dissertation makes a contribution to the ever-growing body of research available about the merger of /o/ and /oh/ in North America. Specifically, it demonstrates that the merger took place in the city of Erie before 1900, and gradually spread outward to the smaller towns and rural areas throughout the county. My field work also shows that the merger is continuing to spread into areas previously on the other side of the North / Midland boundary, and has gone to completion for the youngest generation in the town of Ripley, NY. An important direction for follow-up research is the area in New York past Ripley to the North on the way to Buffalo. Unfortunately, the field work for this dissertation did not include any speakers from the younger generations from these towns, so it is impossible to say with certainty what the status of the merger there is. It is possible that it is also in the process of spreading beyond Ripley, but the lack of apparent time evidence from these other towns along Lake Erie in New York means that we can not know for sure.

On the other hand, my field work shows that the city of Erie is located on the Northern side of the boundary with respect to the back upgliding vowels. The Midland pattern of strong fronting of /ow/ and /Kuw/ is not exhibited by speakers as far North as Erie: it appears to be more specific to Pittsburgh and the other areas of Western Pennsylvania traditionally associated with the Midland.

With respect to the theoretical questions, the empirical evidence shows that dialect boundaries for different subsystems of vowels can fall in different locations. Additionally, the types of boundaries associated with each subsystem can be different. The evidence for the merger of /o/ and /oh/ presented in Chapter 6 showed a sharp boundary: there are very few speakers on the Northern side of the boundary with any evidence for the merger, and no speakers on the Midland side of the boundary with evidence for the distinction. On the other hand, the boundary for the back upgliding vowels is more diffuse. For example, Figure 7.6 showed how there are a few speakers in Erie and North East who show the strong fronting of /ow/ that is characteristic of Pittsburghers, whereas there are also a few speakers from Western Pennsylvania south of Erie County who have mean values of /ow/ less than 1300 Hz. This evidence indicates that the Northern and Midland systems overlap to some degree in the boundary area. This, in turn, suggests that sharp boundaries occur when the feature in question is categorical in nature (such as a phonemic merger), and that less discrete boundaries occur when the feature is phonetically gradient in nature (such as vowel fronting). Further research into other boundary regions with distinct boundary locations for different vowel subsystems will hopefully shed more light on this.

The evidence presented in Chapther 8 for the lexical and morphosyntactic variables shows that the speakers that pattern with the Midland with regard to the merger of /o/ and /oh/ generally also use the Midland variants for positive *anymore* and *need* + Past Participle. Furthermore, Erie is also differentiated from the Northern region in the neighboring area of New York by the lexical variables *elementary* and *redd up*. While this evidence does seem to suggest that these lexical and morphosyntactic variables pattern together with the status of /o/ and /oh/, a more likely explanation for their correlated boundaries can be found in the settlement patterns.

The Midland lexical and morphosyntactic variables under discussion are all attributable to the large influx of Scots-Irish settlers. Additionally, the early and complete merger of /o/ and /oh/ in the area of Western Pennsylvania around Pittsburgh is likely attributable to Scots-Irish settlement in the area. Thus, it seems likely that the spread of these features to the city of Erie also can be attributed to the relatively large presence of Scots-Irish settlers in Erie in the 19th century. This, then, relates to explanation proposed in Chapter 9 for Erie's unique dialectological history: its switch from the North to the Midland was brought about by a large early presence of non-Northern (especially Scots-Irish) settlers. As discussed in Chapter 9, further research into the settlement histories of Northern towns such as Cleveland and Buffalo is necessary to verify this claim.

Appendix A

Key to Vowel Symbols

ANAE	Wells (1982)	Arpabet
/i/	KIT	IH
/e/	DRESS	EH
/æ/	TRAP	AE
/o/	LOT	AA
/ʌ/	STRUT	AH
/u/	FOOT	UH

ANAE	Wells (1982)	Arpabet
/iy/	FLEECE	IY
/ey/	FACE	EY
/ay/	PRICE	AY
/oy/	CHOICE	OY
/aw/	MOUTH	AW
/ow/	GOAT	OW
/uw/	GOOSE	UW
/ah/	PALM	AA
/oh/	THOUGHT	AO

The vowel symbols used in this dissertation follow the notation used in the ANAE (Labov et al. 2006:11–15). The two tables shown above present the equivalent vowel symbols used in two other popular notational systems: Wells (1982) and Arpabet (Fisher et al. 1986).[1]

---

[1] Note that Arpabet does not distinguish between /o/ and /ah/, and simply uses the symbol AA for vowels in both lexical classes.

In addition, several ANAE symbols used to denote vowels in specific allophonic contexts were adopted for this dissertation. These are listed in the following table, along with example words and a description of the symbol's meaning.

Symbol	Examples	Description
/æ/	*ham*, *manager*, *rang*	/æ/ before a nasal consonant
/ayV/	*ride*, *buy*	/ay/ occurring before a voiced coda or word-finally
/ay0/	*fight*, *rice*	/ay/ occurring before a voiceless coda
/Tuw/	*two*, *soon*	/uw/ occurring after a coronal onset
/Kuw/	*food*, *boot*, *who*	/uw/ occurring after a non-coronal onset

Appendix B

List of Minimal Pairs Tested

Minimal Pairs

pin vs. pen

hoarse vs. horse

cot vs. caught

Mary vs. merry

merry vs. marry

fool vs. full

whale vs. wail

poor vs. pour

collar vs. caller

pool vs. pull

ferry vs. furry

don vs. dawn

which vs. witch

barn vs. born

stock vs. stalk

tour vs. tore

berry vs. bury

near Minimal Pairs

father vs. bother

nearer vs. mirror

spa vs. paw

on vs. Don

Appendix C

Word List

hood
bag
here
news
today
dangle
toe
creek
duck
awe
knot
pen
Janet
goal
poor
huge
merry
mole
food
core
den
better
high
Oklahoma
lift
cot
toy
Ed
bitter
left
pot
cable
party
writer
witch
lost
both
bird
manager
sorry
pin
coffee
hammock
boat
man
height
Moe
go
found
ferry
hide
fool
soon
Sunday
bus
Dan
pal
bat
open
house
forty
tock
orange
hug

out
cloth
laughed
hay
spider
beer
Gothic
marry
don’t
had
boy
odd
downtown
gone
hurt
mother
Tuesday
spa
caught
Dawn
sad
Mark
send
talk
full
boss
on
Don
butter
now
planet
down
bother
sew
horrible
bee
understand
deck
sack
bad
farm
made
cut
wash
head
agree
bike
hoe
dog
bit
Mary
who’d
elementary
tire
up
collar
heed
ham
good
route
off
boot
how
put
four
sang
dad
began
hid
know
caller
class
hospital
pour
path
sin
fairy
ran
hope
which
month
hammer
dude
happy
roof
name
cap
bet
thought
classic
copy
home
tiger
father
song
rider

Appendix D

Sentences for Judgment Elicitation Task

Rate the following sentences as 1 ("I could say a sentence like this."), 2 ("I wouldn't say this, but I've heard people around here say something like it.") or 3 ("I've never heard anything like this before—it sounds like bad English."):

1) I was a pitcher when I was young, but now I don't play baseball anymore. 1 2 3

2) My cat looks really hungry. I think he wants fed. 1 2 3

3) Ticket prices are so high anymore, I never go to the movie theater. 1 2 3

4) I really should redd up the living room before the guests come over. 1 2 3

5) Anymore, there's too much crime in this neighborhood. 1 2 3

6) Every newborn baby likes cuddled. 1 2 3

7) I got into an accident last week, and now my front bumper needs repaired. 1 2 3

8) John eats fast food so much anymore, it's no wonder that he's becoming overweight. 1 2 3

Circle the word that sounds most natural in each sentence:

1) I drove through a big, muddy puddle yesterday. Now my car needs washed. / washing.

2) I haven’t cleaned my kitchen in weeks. The floor really needs mopped. / mopping.

Appendix E

DARE's version of "Arthur the Rat" (Cassidy and Hall 1985:xliii)

Once upon a time there was a young rat who couldn't make up his mind. Whenever the other rats asked him if he would like to come out hunting with them, he would answer in a hoarse voice, "I don't know." And when they said, "Would you rather stay inside?" he wouldn't say yes, or no either. He'd always shirk making a choice.

One fine day his aunt Josephine said to him, "Now look here! No one will ever care for you if you carry on like this. You have no more mind of your own than a greasy old blade of grass!"

The young rat coughed and looked wise, as usual, but said nothing.

"Don't you think so?" said his aunt stamping with her foot, for she couldn't bear to see the young rat so cold blooded.

"I don't know," was all he ever answered, and then he'd walk off to think for an hour or more, whether he should stay in his hole in the ground or go out into the loft.

One night the rats heard a loud noise in the loft. It was a very dreary old place. The roof let the rain come washing in, the beams and rafters had all rotted through, so that the whole thing was quite unsafe.

At last one of the joists gave way, and the beams fell with one edge on the floor. The walls shook, the cupola fell off, and all the rats' hair stood on end with fear and horror.

"This won't do," said their leader. "We can't stay cooped up here any longer." So they sent out scouts to search for a new home.

A little later on that evening the scouts came back and said they had found an old-fashioned horse-barn where there would be room and board for all of them.

The leader gave the order at once, "Company fall in!" and the rats crawled out of their holes right away and stood on the floor in a long line.

Just then the old rat caught sight of young Arthur—that was the name of the shirker. He wasn't in the line, and he wasn't exactly outside it—he stood just by it.

"Come on, get in line!" growled the old rat coarsely. "Of course you're coming too?"

"I don't know," said Arthur calmly.

"Why, the idea of it! You don't think it's safe here any more, do you?"

"I'm not certain," said Arthur undaunted. "The roof may not fall down yet."

"Well," said the old rat, "we can't wait for you to join us." Then he turned to the others and shouted, "Right about face! March!" and the long line marched out of the barn while the young rat watched them.

"I think I'll go tomorrow," he said to himself, "but then again, perhaps I won't—it's so nice and snug here. I guess I'll go back to my hole under the log for a while just to make up my mind."

But during the night there was a big crash. Down came beams, rafters, joists—the whole business.

Next morning—it was a foggy day—some men came to look over the damage. It seemed odd to them that the old building was not haunted by rats. But at last one of them happened to move a board, and he caught sight of a young rat, quite dead, half in and half out of his hole.

Thus the shirker got his due, and there was no mourning for him.

Bibliography

Anufryk, Volha. 2008. Effects of language ability on phonetic and phonological variation of English intonation by German speakers. In *Proceedings of the 8th International Seminar on Speech Production*, 233–235.

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

Bailey, Guy, Tom Wikle, Jan Tillery, and Lori Sand. 1993. Some patterns of linguistic diffusion. *Languave Variation and Change* 5:359–390.

Baker, Adam. 2005. Quantifying diphthongs: A statistical technique for distinguisthing formant contours. Paper presented at NWAV 35, Columbus, OH.

Baranowski, Maciej. 2006. Phonological variation and change in the dialect of Charleston, South Carolina. Doctoral Dissertation, University of Pennsylvania.

Bishop, Christopher M. 2006. *Pattern recognition and machine learning*. Springer.

Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.

Boberg, Charles. 2005. The Canadian Shift in Montreal. *Language Variation and Change* 17:133–154.

Carver, Craig M. 1987. *American regional dialects: A word geography*. Ann Arbor: University of Michigan Press.

Cassidy, Frederic G., and Joan Houston Hall. 1985. *Dictionary of American Regional English, Volume I*. Cambridge, MA: Harvard University Press, Belknap.

Cassidy, Frederic G., and Joan Houston Hall. 1985–2002. *Dictionary of American Regional English*. Cambridge, MA: Harvard University Press, Belknap.

Cassidy, Frederic G., and Joan Houston Hall. 2002. *Dictionary of American Regional English, Volume IV*. Cambridge, MA: Harvard University Press, Belknap.

Chambers, J.K., and Peter Trudgill. 1999. *Dialectology (2nd Edition)*. Cambridge University Press.

Chen, Nancy F., Wade Shen, Joseph Campbell, and Reva Schwartz. 2009. Large-scale analysis of formant frequency estimation variability in conversational telephone speech. In *Proceedings of Interspeech*.

Crozier, Alan. 1984. The Scotch-Irish influence on American English. *American Speech* 59:310–331.

Davis, Steven B., and Paul Mermelstein. 1980. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28:357–366.

Deng, Li, Xiaodong Cui, Robert Pruvenok, Jonathan Huang, Safiyy Momen, Yanyi Cheng, and Abeer Alwan. 2006. A database of vocal tract resonance trajectories for research in speech processing. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing* .

Dinkin, Aaron. 2008. Settlement patterns and the Eastern boundary of the Northern Cities Shift. Paper presented at the American Dialect Society annual meeting, Chicago, January 2008.

Dinkin, Aaron. 2009. Dialect Boundaries and Phonological Change in Upstate New York. Doctoral Dissertation, University of Pennsylvania.

DiPaolo, Marianna. 1988. Pronunciation and categorization in sound change. In *Linguistic Change and Contact: NWAV XVI*, 84–92. Austin: Dept. of Linguistics, University of Texas.

DiPaolo, Marianna. 1992. Hypercorrection in response to the apparent merger of (ɔ) and (a) in Utah English. *Language and Communication* 12:267–292.

Dunlap, A. R. 1945. Observations on American colloquial idiom. *American Speech* 20:12–21.

Evanini, Keelan, Stephen Isard, and Mark Liberman. 2009a. Automatic formant extraction for sociolinguistic analysis of large corpora. In *Proceedings of Interspeech*.

Evanini, Keelan, Joshua Tauberer, Jiahong Yuan, and Mark Liberman. 2009b. Forced alignment using P2FA: A tutorial. Unpublished manuscript, University of Pennsylvania.

Fisher, William M., George R. Doddington, and Kathleen M. Goudie-Marshall. 1986. he darpa speech recognition research database: Specifications and status. In *Proceedings of DARPA Workshop on Speech Recognition*, 93–99.

Flanagan, James L. 1955. A difference limen for vowel formant frequency. *Journal of the Acoustical Society of America* 27:613–617.

Fogle, Deena. 2006. Expansion or approximation: The low-back merger in Indianapolis. Paper presented at NWAV 36, Philadelphia, PA.

Fogle, Deena. 2008. Indianapolis, Indiana: A prototype of Midland convergence. *University of Pennsylvania Working Papers in Linguistics: Proceedings of the 31st Annual Penn Linguistics Colloquium* 14:135–148.

FWP, Federal Writers' Project. 1938. *Erie: A guide to the city and county*. Philadelphia, William Penn Association: WPA.

Glenn, Meghan Lammie, and Stephanie Strassel. 2008. Shared linguistic resources for the meeting domain. In *Lecture Notes in Computer Science*, volume 4625, 401–413. Springer.

Gooden, Shelome, and Maeve Eberhardt. 2007. Local identity and ethnicity in Pittsburgh AAE. *University of Pennsylvania Working Papers in Linguistics* 13:81–94.

Gordon, Matthew J. 2006. Tracking the low back merger in Missouri. In *Language variation and change in the American Midland: A new look at "Heartland" English*, ed. Thomas E. Murray and Beth Lee Simon, 57–68. John Benjamins.

Hall-Lew, Lauren. 2009. Ethnicity and phonetic variation in a San Francisco neighborhood. Doctoral Dissertation, Stanford University.

Hazen, Kirk. 2005. Mergers in the mountains: West Virginia division and unification. *English World-Wide* 26:199–221.

Hermansky, Hynek. 1990. Perceptual linear predictive (PLP) analysis of speech. *Journal of the Acoustical Society of America* 87:1738–1752.

Herold, Ruth. 1990. Mechanisms of merger: The implementation and distribution of the low back merger in Eastern Pennsylvania. Doctoral Dissertation, University of Pennsylvania.

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

Hinton, Leanne, Sue Bremner, Hazel Corcoran, Jean Learner, Herb Luthin, Birch Moonwomon, and Mary van Clay. 1987. It's not just the Valley Girls: A study of California English. In *Proceedings of the 13th annual meeting of the Berkeley Linguistics Society*, 117–127.

Jenks, George F., and Fred C. Caspall. 1971. Error on choroplethic maps: Definition, measurement, reduction. *Annals of the Association of American Geographers* 61:217–244.

Johnson, Daniel Ezra. 2007. Stability and change along a dialect boundary: The low vowels of Southeastern New England. Doctoral Dissertation, University of Pennsylvania.

Johnson, Keith. 1997. *Acoustic and Auditory Phonetics*. Blackwell.

Konopka, Kenneth, and Janet Pierrehumbert. 2008. Vowels in contact: Mexican Heritage English in Chicago. In *Proceedings of the Symposium about Language and Society - Austin*, 94–103.

Kretzschmar, William A., Jr., Virginia G. McDavid, Theodore K. Lerud, and Ellen Johnson. 1993. *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. University of Chicago Press.

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

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

Labov, William. 1972. *Sociolingusitic patterns*. Phialdelphia: University of Pennsylvania Press.

Labov, William. 1973. Where do grammars stop? In *Report on the twenty-third annual round table meeting on linguistics and language studies*, ed. Roger Shuy, 43–88. Georgetown University Press.

Labov, William. 1974. Linguistic change as a form of communication. In *Human communication: Theoretical explorations*, ed. A. Silverstein, 221–256. Hillsdale, N.J.: Erlbaum.

Labov, William. 1991. The three dialects of English. In *New ways of analyzing sound change*, ed. Penelope Eckert, 1–44. New York: Academic Press.

Labov, William. 1994. *Principles of linguistic change, volume I: Internal factors*. Oxford: Basil Blackwell.

Labov, William. 2003. Pursuing the cascade model. In *Social dialectology*, ed. David Britain and Jenny Cheshire, 9–22. John Benjamins.

Labov, William. 2007. Transmission and diffusion. *Language* 83:344–487.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English*. Mouton de Gruyter.

Labov, William, Mark Karan, and Corey Miller. 1991. Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3:33–74.

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

LDC. 2004. Linguistic Data Consortium: RT-07 meeting quick transcription guidelines, version 1.3. `http://www.ldc.upenn.edu/Catalog/docs/LDC2004T13/MeetingDataQTRSpec-V1.3.pdf`.

Lechner, Carl B. 1994. *Erie: Link to the Great Lakes*. Erie, PA: Erie County Historical Society.

Lennig, Matthew. 1978. Acoustic measurement of linguistic change: The modern Paris vowel system. Doctoral Dissertation, University of Pennsylvania.

Lisker, Leigh. 1948. The distinction between [æ] and [ɛ]: A problem in acoustic analysis. *Language* 24:397–407.

Majors, Tivoli. 2005. Low back vowel merger in Missouri speech: Acoustic description and explanation. *American Speech* 80:165–179.

McCool, Sam. 1982. *Sam McCool's new Pittsburghese*. Pittsburgh, PA: Goodwill Industries.

Mermelstein, Paul. 1978. Difference limens for formant frequencies of steady-state and consonant-bound vowels. *Journal of the Acoustical Society of America* 63:572–580.

Miller, James D. 1989. Auditory-perceptual interpretation of the vowel. *Journal of the Acoustical Society of America* 85:2114–2134.

Montgomery, Michael. 1997. The Scotch-Irish element in Appalachian English: How broad? How deep? In *Ulster and North America: Transatlantic perspectives on the Scotch-Irish*, ed. Tyler Blethen and Curtis Wood, 189–212. Tuscaloosa: University of Alabama Press.

Montgomery, Michael. 2004. Solving Kurath's puzzle: Establishing the antecedents of the American Midland dialect region. In *Legacies of colonial English: Studies in transported dialects*, ed. Raymond Hickey, 311–325. Cambridge University Press.

Moonwomon, Birch. 1992. Sound change in San Francisco English. Doctoral Dissertation, University of California, Berkeley.

Mufwene, Salikoko S. 1996. The Founder Principle in creole genesis. *Diachronica* 13:83–134.

Muller, Mary M. 1991. *A Town at Presque Isle: A Short History of Erie, Pennsylvania to 1980*. Erie County Historical Society.

Murray, Thomas E., Timothy C. Frazer, and Beth Lee Simon. 1996. Need + past participle in American English. *American Speech* 71:255–271.

Murray, Thomas E., and Beth Lee Simon. 1999. Want + past participle in American English. *American Speech* 74:140–164.

Murray, Thomas E., and Beth Lee Simon. 2002. At the intersection of regional and social dialects: The case of like + past participle in American English. *American Speech* 77:32–69.

Nearey, Terrence. 1977. Phonetic feature system for vowels. Doctoral Dissertation, University of Connecticut.

Nycz, Jennifer, and Paul De Decker. 2005. A new way of analyzing vowels: Comparing formant contours using smoothing spline ANOVA. Paper presented at NWAV 35, Columbus, OH.

Peterson, Gordon E., and Harold L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24:175–184.

Pierrehumbert, Janet B., Tessa Bent, Benjamin Munson, Ann R. Bradlow, and J. Michael Bailey. 2004. The influence of sexual orientation on vowel production. *Journal of the Acoustical Society of America* 116:1905–1908.

Rabiner, Lawrence R. 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE* 77:257–286.

Sanford, Laura G. 1894. *The History of Erie County, Pennsylvania, from its first settlement, revised and enlarged edition*. Erie, PA: published by the author.

Sankoff, David, and Gillian Sankoff. 1973. Sample survey methods and computer-assisted analysis in the study of grammatical variation. In *Canadian Languages in their Social Context*, ed. Regna Darnell, 7–64. Edmonton, Alberta: Linguistic Research.

Shields, Kenneth. 1997. Positive *anymore* in Southeastern Pennsylvania. *American Speech* 72:217–220.

Stabley, Rhodes R. 1959. ‘needs painted’, etc., in western Pennsylvania. *American Speech* 34:69–70.

Steinlen, Anja K. 2002. *The influence of consonants on native and non-native vowel production*. Tübingen: Narr.

Talkin, David. 1987. Speech formant trajectory estimation using dynamic programming with modulated transition costs. *Journal of the International Phonetic Association* 82:S55.

Tauberer, Joshua, and Keelan Evanini. 2009. Intrinsic vowel duration and the post-vocalic voicing effect: Some evidence from dialects of North American English. In *Proceedings of Interspeech*.

Trudgill, Peter. 1974. Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society* 3:215–246.

Viterbi, Andrew J. 1967. Error bounds for convolution codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory* 13:260–269.

Watson, Catherine I., and Jonathan Harrington. 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America* 106:458–468.

Wells, J.C. 1982. *The accents of English.* Cambridge University Press.

Wetmore, Thomas H. 1959. *The low-central and low-back vowels of the English of the Eastern United States*. Tuscaloosa, AL: American Dialect Society.

Yang, Charles. 2009. Population structure and language change. Unpublished manuscript, University of Pennsylvania.

Youmans, Gilbert. 1986. Any more on anymore?: Evidence from a Missouri dialect survey. *American Speech* 61:61–75.

Yuan, Jiahong, and Mark Liberman. 2008a. Speaker identification on the SCOTUS corpus. In *Proceedings of Acoustics '08.*

Yuan, Jiahong, and Mark Liberman. 2008b. Vowel acoustic space in continuous speech: An example of using audio books for research. Paper presented at CatCod 2008, the First International Workshop on Cataloguing and Encoding of Spoken Language Data.

Yuan, Jiahong, and Mark Liberman. 2009. Investigating /l/ variation in English through forced alignment. In *Proceedings of Interspeech*.

Zelinsky, Wilbur. 1973. *The cultural geography of the United States*. Prentice Hall.