Laurel MacKenzie*
University of Manchester

This paper examines the variable weakening and deletion of /s/ in the Languedocian variety of Modern Occitan, with particular attention to how this has affected the system of plural marking in noun phrases. Using data from linguistic atlases, I demonstrate that /s/-lenition in this variety involves a stage of vocalization to [j]. I find that, where /s/ on the definite article has vocalized to [j], the immediately-preceding vowel of the definite article has undergone concomitant raising to [e]. This raising appears to preserve the difference between singular and plural despite the plural’s weakening /s/. I argue that these results support Labov’s (1994) hypothesis that the meaning of a weakening element may be transferred to a stable, co-occurrent one.

1. Introduction

This paper examines the variable weakening and deletion of /s/ in the Languedocian variety of Modern Occitan, with particular attention to how this has affected the system of plural marking in noun phrases. As is common in other Romance languages, the dialect of Occitan under study marks plurality on a noun phrase by affixing an -s plural morpheme to the article, the noun, and any adjectives, as in (1).

(1) [la-s ‘kambio-s pe’lydo-s] DEF.FEM-PL leg-PL hairy-PL
‘hairy legs’

However, as in Spanish and Portuguese (see, e.g., Guy, 1981; Hernández-Campoy and Trudgill, 2002; Poplack, 1980; Ranson, 1991; Terrell, 1977), the /s/ of the plural can variably undergo weakening and deletion. This paper uses data from dialect atlases (Ravier, 1978, 1982, 1986, 1993) to examine what happens to the singular-plural distinction in Occitan when /s/ is lost: is there a tendency to preserve meaning? By what route?

I demonstrate that /s/-lenition in Languedocian involves a stage of vocalization to [j]. When the /s/ plural marking on the definite article vocalizes, the article vowel is raised to [e], preserving the difference between singular and plural despite the plural’s weakening /s/. I argue that these results support Labov’s (1994) hypothesis that the meaning of a weakening element may be transferred to a stable, co-occurrent one.

This paper is structured as follows. Section 2 provides background information on the language under study, including previous work on /s/-loss in Occitan, and the source of the data. Section 3 examines the weakening of the plural -s marker on nouns and definite articles, followed by an examination of concomitant effects on the vowel of the definite article. I conclude with a discussion of the role of meaning preservation in language change (Section 4).

2. Background

2.1. Language and Data

Occitan is the generic name used to refer to a collection of varieties, distinct from French, spoken throughout Southern France. Occitan is also known as Langue d’Oc and, historically, Provençal, though today this term refers properly only to a particular Occitan variety, that spoken in Provence. The variety of Occitan examined in this paper is Languedocian.

Ethnologue places both Occitan and French in the Gallo-Iberian branch of Romance languages (Lewis, Simons & Fennig, 2014); the two followed different paths of development beginning in the 5th century with the invasion of Gaul by foreign tribes. Northern France sustained invasions by Britons, Saxons, and particularly Franks, who lent their name to the nation, while Southwest France, with the exception of Gascony, which was invaded by the Basques, received minimal outside influence. The Latin language and Roman culture thus remained more intact in Southern than Northern France, given the different levels of occupation in each half of the country. By the 12th century, the vernaculars of Northern and Southern France were considered different languages, known as the Langue d’Oïl in the North and the Langue d’Oc in the South, after each language’s affirmative particle (Pope, 1952).
Anglade ([1921] 1977) placed the number of speakers of Occitan at 12 or 14 million; the 1999 French census found fewer than two million people who said that their parents had spoken to them at least occasionally in Occitan in their childhood, and only 250,000 who had spoken it to their own children (Héran, Filhon, and Deprez, 2002). Ethnologue places the number of speakers in France at only 110,000 (Lewis et al., 2014). As a reviewer observes, the difficulties inherent in polling speakers on their use of minority languages—particularly in France, where minority languages have been aggressively marginalized for centuries (Schiffman, 1996)—mean that we should interpret these numbers with some caution, in case some speakers of Occitan were reluctant to reveal themselves to census-takers. Still, sources agree in recognizing the language as endangered (Lewis et al., 2014; Moseley, 2010).

The data used in this paper come from the Atlas linguistique et ethnographique du Languedoc occidental (abbreviated henceforth as ALLOc; Ravier, 1978, 1982, 1986, 1993). The area sampled in ALLOc is bordered by the Dordogne river to the north and the Garonne river to the west; the eastern boundary of the sample was determined based on the sample sites used in other linguistic atlases (the Atlas linguistique et ethnographique du Massif Central, the Atlas linguistique du Languedoc oriental) as well as topographical, cultural, and ethnographic considerations (for instance, to avoid dividing what was felt by natives to be a coherent region). The sampled area encompasses, in whole or in part, the present-day departments of Ariège, Aude, Haute-Garonne, Tarn, Tarn-et-Garonne, Lot-et-Garonne, Aveyron, Lot, Gironde, Dordogne, and (with one point represented only) Corrèze. It includes the major city of Toulouse but not that of Bordeaux. Data were gathered between 1967 and 1972, collected from 131 localities 15–17 km (approx. 9–11 mi) apart, as plotted in Map 1. (Subsequent plots will zoom in on the region under study.)
Approximately 1200 lexical items were elicited for the atlas, with each lexical item elicited once per locality. 73% of the 265 primary informants were aged 60 or older, and 72% were male. Many worked in agricultural professions. This is, of course, reminiscent of the ‘non-mobile, older, rural males’ (NORMs) sought out as informants in many dialectological studies (Chambers & Trudgill, 1998:29), though the atlas gives no indication of whether this was a deliberate choice on the part of the researchers or an inevitability due to a difficulty in finding Occitan speakers in more urban areas. Indeed, a respondent from the city of Toulouse was included in the sample, as were women and young people (though these non-NORM informants are often identified as having been family members of the older males).

To elicit forms, researchers provided a French word and asked the informant for the Occitan translation, so all informants had some level of bilingualism. Data were collected by semantic category and were phonetically transcribed.

2.2. Linguistic Background

We can see a historical parallel to the /s/-lenition of Languedocian Occitan in a similar process that occurred in the evolution of French. Old French had two cases, a nominative and an oblique; the oblique case, which is the source of being lost due to /s/-deletion, the contrast was expressed only by the vowel of the definite article from singular /la/ (feminine) or /lə/ (masculine) to plural /le/ (both genders). The only place where /s/ still manifests itself is in the process of liaison before a vowel: for instance, the vowel-initial noun l’heure [lœʁ] ‘the hour’ has as its plural les heures [lez əʁ]. In the absence of liaison, though, plural meaning for nouns that form their orthographic plural in <s> is expressed only by the vowel of the definite article.

The chain of events that gave rise to this plural-conveying vowel alternation is not known. Seklaoui (1989:8) refers to the /e/ in the plural definite article as the result of “an obscure phonetic evolution.” Spence (1976) proposes that its source is as follows: The vowels of the Vulgar Latin plural definite articles los and las were reduced to [ə], which was then variably lengthened compensatorily after loss of the following /s/, and this [ə] then raised to [e]. He admits, though, that there is no clear phonetic mechanism by which this could have occurred, and that the need to disambiguate singular from plural must have been an influencing factor. In this paper, I propose that the patterns found in the related language of Occitan may elucidate the question.

This is not the first study of /s/-deletion and its effects on plurality in Modern Occitan. Eckert (1969, 1985) performed a similar study using data gathered from the *Atlas Linguistique de la France* (henceforth ALF), for which fieldwork was carried out between 1897 and 1901. She found evidence for two sound changes in Southern France: One deleting the plural /s/ marker, which was restricted only to the northern half of the region she examined, and a concomitant change of *a* to [o], which was more widespread, but which showed grammatical conditioning. Specifically, in the northern half of the region under study, where plural /s/ had been deleted, the *a* to [o] change had gone to completion in singulars only; in the area where /s/ was retained, the *a* to [o] change had gone to completion in both singulars and plurals (as sketched in (2)). Directly between the /s/-deleting North and the /s/-retaining South, Eckert identified an area of “number disturbance,” in which /s/ was variably retained; this region was bisected by the isogloss delimiting the change of *a* to [o] in plurals. The upshot of all this is that where the singular-plural distinction was in danger of being lost due to /s/-deletion, the contrast was maintained by the simultaneous *a* to [o] change.

<table>
<thead>
<tr>
<th>Region of /s/-deletion:</th>
<th>Region of /s/-retention:</th>
</tr>
</thead>
<tbody>
<tr>
<td>*a &gt; [o] in singulars only</td>
<td>*a &gt; [o] in both plurals</td>
</tr>
<tr>
<td>sg. lo bello pullo pl. la</td>
<td>and singulars sg. lo</td>
</tr>
<tr>
<td>bella pullo ‘the beautiful</td>
<td></td>
</tr>
<tr>
<td>hen(s)’</td>
<td>los bellos</td>
</tr>
<tr>
<td>plulos ‘the beautiful</td>
<td></td>
</tr>
<tr>
<td>hen(s)’</td>
<td></td>
</tr>
</tbody>
</table>

The data provided in ALLOC were collected 70 years after Eckert’s ALF data, and reveal that /s/-deletion has continued to spread, even into the area where the *a* to [o] change went to completion in both singulars and plurals. Floricic (2010), Barra-Jover (2012), and Sauzet (2012) all use the data in ALLOC (or as-yet unpublished data collected as part of the ALLOC survey) to investigate the status of plural marking in Languedocian in the absence of this “recourse” of final vowel quality. All three authors identify a range of phonological realizations of plurality, which I corroborate below. The present paper additionally expands on these works by providing a comprehensive geographical account of the situation in Languedocian, examining the realization of plural marking on multiple lexical items across the entire region sampled in the atlas. Doing so
allows us not only to examine the synchronic state of affairs, but, in comparing neighboring varieties, to hypothesize as to its diachronic development.

3. Data

The findings presented in this section are drawn from the realization of plural on the definite article and the noun of eight lexical items. These eight items were selected based on two criteria: For the majority of localities surveyed, they (a) are of feminine gender and (b) contain a noun beginning with a voiceless stop. These restrictions were made to control for processes independent of /s/-lenition which are known to affect the phonological realization of the definite article. For instance, as in other Romance languages, Languedocian Occitan marks noun gender in the vowel of the definite article, so nouns were restricted to a single gender (here, the feminine) in order to ensure that all would have the same definite article. Additionally, because the initial segment of a Languedocian noun has been shown to induce allophonic variation in the final segment of the definite article that precedes it (Barra-Jover, 2012; Ronjat, 1932; Sauzet, 2012), the manner and voicing of the initial consonant were restricted as a control (here, to voiceless stops only).

Any data point in which the noun was not feminine, not voiceless stop-initial, or not plural (due to, for instance, a location having a collective singular term for an item) was omitted from this study, and is represented in the maps below as an ‘N/A’ point. When an ALLOc informant provided two possible realizations for
a locality, the first of these was selected for analysis, based on the remark in ALLOC’s front material that “in a very general fashion, the order of forms on the map is the order in which they were provided by the informant” (Ravier, 1978:xiv, translation mine).

Findings below will frequently be exemplified with a single lexical item, but reference will be made to that item’s typicality vis-à-vis the other seven items examined.

### 3.1. Realization of /s/ on the Noun

Map 2 plots the realization of /s/ on the noun only, with one point for each location surveyed in ALLOC, for the lexical item les traces ‘animal tracks’ (ALLOC’s map 352). Points represented with a blue dot show no marker of plurality on the noun, e.g. [pjado] (point 24.30), which I term ‘deletion.’ Points represented with a yellow dot are those in which /s/ on the noun surfaces as [j], e.g. [pe’nadʃ] (point 46.17), which I term ‘vocalization.’ Points represented with a red dot are those in which /s/ on the noun surfaces as [s] (or, less frequently, some other fricative with a buccal articulation, such as [ʃçx]), e.g. [pe’zados] (point 81.05), which I term ‘retention.’ Points represented with a black ‘x’ were discarded as not applicable for the reasons elucidated above (i.e., the informant’s response was not feminine, not plural, or did not contain a voiceless stop-initial noun), or because the informant was unable to provide a response.

Map 2 reveals that vocalization of /s/ to [j], while not particularly prevalent in the area under study,
is well-represented in a tight cluster of localities in the northeast sector. Vocalization is not attested as a stage of /s/-loss in Ferguson’s (1990) typology of [s] > [h] changes, and likewise is not mentioned as a variant in those synchronic studies of /s/-weakening in Spanish and Portuguese cited earlier. In Occitan, it is frequently identified as an allophonic variant of /s/ which appears before voiced consonants (Barra-Jover, 2012; Ronjat, 1932; Wheeler, 1988), but the fact that the data points in Map 2 have been restricted to only those in which the article precedes a voiceless stop-initial noun means that [j] must be a stage on the /s/-weakening trajectory. Indeed, Seklaoui (1989) does list vocalization as a possible, though rare, outcome of /s/-weakening, citing examples from Italian as well as Occitan.

To show that the patterns of /s/-realization depicted in Map 2 are not specific to the lexical item les traces, I calculated the percentage of each variant of /s/ for each locality across all eight lexical items examined. The most common variant for each locality is plotted in Map 3, with a color saturation indicating how common that variant was across the eight lexical items (the darker the color, the more well-represented the variant). For instance, a locality for which [s] was retained on the noun in all eight lexical items examined will be plotted as a dark red dot; a locality which had retention of [s] for one item, vocalization to [j] for four items, and deletion of /s/ for three items will be plotted as a light yellow dot (indicating that [j] was the most common variant across the eight items, but that it was only

Map 4. Realization of /s/ on the article of les traces.
represented in 50% of the observed forms). As in Map 2, buccal fricatives other than \([s]\) (e.g. \([\dot{\textsf{ʃ}}\ \text{x}]\)) have been plotted as instances of retention. Map 3 indicates that, for all but a few border towns, the pattern demonstrated by *les traces* holds for the other lexical items examined.

Map 3 allows us to identify the northwest sector of the region under study as being the most advanced in /s/-weakening (having no overt plural marking on its nouns), the central/southern portion as being the most conservative (having not yet begun to weaken /s/ in any way), and the northeast sector as being somewhere in between (having lenited its realization of /s/ to \([\j]\) but not gone so far as to delete it). These three realizations could be viewed as points on a single trajectory (with /s/ first weakening to \([\j]\) and then deleting entirely); alternatively, there could be two different trajectories of /s/-lenition, one in which /s/ deletes with no intermediate stage of \([\j]\), and another in which /s/ vocalizes (potentially to delete later on). The present data don’t obviously argue for one or the other, and the existence of a unified trajectory of /s/-lenition is not crucial to the present study, so I remain agnostic on this point.

### 3.2. Realization of /s/ on the Article

Map 4 shows the realization of /s/ on the article of *les traces*. In addition to retained (e.g. [las pe’zados], point 81.05), vocalized (e.g. [lej ‘pjado], point 47.02), and deleted\(^*\) (e.g. [la: pj’da], point 24.10) variants—plotted as in the previous maps—aspiration of /s/ to \([\h]\)
(e.g. [lah ‘pjados], point 82.12), plotted in green, is observed; /s/ may also fully assimilate to the following consonant (e.g. [lap ‘pjado], point 47.10), resulting in a geminate (plotted in purple).

Map 4 shows that the southern half of the region under study again behaves conservatively, with [s] retained on the article as it was on the noun. Also, more generally, /s/ on the article has not advanced as far in the weakening process as had /s/ on the noun, with only one point showing total deletion of /s/. More advanced weakening on nouns than articles has been similarly found in studies of synchronic /s/-lenition in Spanish and Portuguese (Cedergren, 1973; Guy, 1981; Poplack, 1980).

Map 5 shows the same type of plot as Map 3, with the most attested variant across the eight lexical items plotted in a color saturation indicative of how common it is. There is somewhat more inter-item inconsistency in the realization of article /s/ than there was for noun /s/, as indicated by the greater number of lighter-colored points in Map 5 than in Map 3; indeed, the mean proportion of use of the most common variant is 0.977 for noun /s/ and 0.891 for article /s/. Still, the pattern evident in les traces is largely shared by the other lexical items under examination.

Again, the northeast, northwest, and southern regions show differing realizations, but again, there is no obvious indication of whether they represent successive stages on a continuum of [s] > [h] > [j], or distinct paths of lenition ([s] > [h]; [s] > [j]).

3.3. Article Vowel Quality

The final factor to examine is the quality of the vowel in the definite article. As outlined in Section 2.2, definite article vowel quality plays a crucial role in disambiguating singular from plural nouns in the related language of French.
Map 6 shows the vowel quality of the singular definite article in the region under study (based on ALLOC’s map 1), revealing that the region can be bisected into a [lɔ] region in the northeast and a [la] region in the southwest. The default assumption is that this vowel quality will be maintained in the plural, as it is in traditional dialects of Occitan (Wheeler, 1988) and in other Romance languages.

Map 7 superimposes isoglosses indicating the observed vowel quality of the plural definite article on the plot from Map 4 depicting the realization of /s/ on the article of les traces. The black isogloss marks [lɔ]/[la] split in the singular definite article. Map 7 reveals that this same split is maintained throughout much of the region in the plural article as well, with the crucial exception of the northwest sector, where the vowel of the article is, exceptionally, [e]. This region of article [e] coincides almost perfectly with the region in which /s/ on the article has vocalized to [j]. The observed vowel raising is thus almost certainly due to coarticulation of the vowel [a] and the following palatal glide.

Further data reveal, however, that vowel raising in Languedocian plural articles is not purely a mechanical process. Maps 9 and 10 present data for the lexical item les fleurs du vin ‘yeasts that develop on the surface of wine’ (map 754 in ALLOC). Map 8 shows that the manifestation of /s/ on the noun is nearly identical to what it was for les traces in Map 2.
By contrast, Map 9 reveals that the realization of /s/ on the article of "les fleurs du vin" differs in a crucial way from that of article /s/ for "les traces" (Map 4): Much of the southern sector of the region, which retained [s] in the article of "les traces", vocalizes /s/ in the article of "les fleurs du vin". This southern strip of vocalization has a phonological source: It corresponds almost perfectly with the region in which the noun for this lexical item begins with [f] rather than [k] (as it does elsewhere in the region). This is thus an instance of the *allophonic* /s/-vocalization mentioned earlier which has heretofore been omitted from analysis. The area in which the noun is fricative-initial is surrounded by the green isogloss in Map 10; it corresponds well to the southern area which vocalizes /s/ on the article to [j].

But unlike the vocalization depicted in Map 7, this allophonic vocalization does not trigger vowel raising. Map 11 shows a clear difference between /s/-vocalization in the northwest sector and /s/-vocalization in the south: In the south, vocalization has no accompanying vowel change from [a] to [e].

4. Analysis

I propose that these findings are interpretable via Labov’s (1994) observations on the maintenance of meaning in systems in which a grammatical marker is being lost. Specifically, Labov proposes the following:

When an element of the system is found to co-occur frequently with the signal being deleted, it is increasingly eligible to assume the distinctive
Labov cites as an example of this the loss of the negative marker *ne* in French, which has been all but replaced with the particle *pas*. He proposes that this situation has its source in the following development. Given that *ne* was involved in a process of weakening and deletion and *pas* was not strictly required for a negative sentence, in the history of French, sentences would inevitably surface which were intended by a speaker as negative but which bore no overt negative marker. These sentences were ambiguous between negative and positive (Labov’s “unsupported zeroes,” 1994:589), and some learners may have misinterpreted them as positive. Over time, as learners attempted to match their percentage of *ne* use to that of the negative sentences they received in their input, the unambiguously-negative input they received became more and more skewed toward tokens containing *pas*, since those that didn’t contain *pas* were prone to be factored out as positives. This eventually resulted in *pas* taking over as the negative marker, as the loss of *ne* eventually went to completion.

If an analogous process occurred in Languedocian, we could explain the differing article vowel qualities between the northwest sector and the south in cases like *les fleurs du vin*, where both regions vocalize */s/ on the article but only the northwest shows concomitant vowel raising. In the south, */s/ is consistently retained on the
noun (Map 8), meaning that there is no possibility of a plural form being mistaken for a singular, even when the \(/s/ of its article has weakened to [j]. In the northwest, by contrast, the loss of /s/ on the noun means that the singular comes to bear a close resemblance to the plural: Only the [j] on the article separates them. If this [j] were itself to be variably deleted, singular-plural ambiguity would result and those tokens in which the vowel had not raised to [e] could be misclassified as singulars. At the same time, a variably retained [j] would trigger vowel raising, meaning that the number of forms unambiguously identified as plurals could eventually shift to those that bore [e] in their article, despite having no consonantal plural marking. In the south, where there are no ambiguous forms, learners would not have any more evidence of plurality from forms in which the vowel was raised than from forms in which it wasn’t. As a result, any coarticulation-induced raising would fail to be cemented as part of the linguistic system, instead remaining merely a phonetic variant of the unraised vowel.

This analysis crucially depends on there being ambiguous forms in the northwest to be misrepresented as singular. I have assumed here that such forms arise when [j] is variably deleted, but there is no evidence from ALLOc of such [j]-deletion actually occurring. This is the downside of working with linguistic atlases: we have no access to the synchronic variation present.
within a speaker or a community at the time the data were collected. Nevertheless, the atlas has a clear advantage over other linguistic methods in allowing rapid comparison between multiple regions. In the present study, such rapid comparison has revealed data that appears to support the importance of ambiguous or near-ambiguous forms for meaning preservation in the face of a process of lenition.

5. Conclusions

This paper has added data from the Languedocian dialect of Occitan to the vast typology of cases of /s/-lenition in the world’s languages. The dialect atlas data considered here have revealed that /s/-lenition in Languedocian includes a stage of vocalization to [j], which, when it occurs on a definite article, appears to trigger raising of a preceding vowel. This raising, a change which is mechanically actuated, apparently becomes used to preserve plurality in regions where /s/-weakening is advanced.

This study raises a number of further questions. Did /s/-loss in French go through a stage of vocalization as well, and is this the source of the mysterious /e/ in plural definite articles in that language? What has happened to the realization of plural on verbs in Languedocian, and can that be shown to have had any effect on plural marking in the noun phrase? After all,
though the noun phrases in ALLOc were uttered in isolation, natural speech would of course provide additional sources of plural marking. Finally, preliminary analysis of masculine plural articles reveals that those in the advanced northwest sector take the shape [lu] in the singular but [lej] in the plural. Is this also (perhaps less articulatorily likely) a result of vowel mutation triggered by the following [j], or is this the result of analogy with the feminine article?

As is the case with many dialect atlases, ALLOc was conceived primarily to study lexical variation and was not strictly designed to examine sound changes: For instance, no attempt was made to target minimal pairs. Nevertheless, it has revealed a reasonable path by which the otherwise-unexplained /e/ in French plural articles may have come about, and has provided data relevant to theories of homophony avoidance and the maintenance of semantic distinctions.

Notes

1 The eight lexical items examined were les traces 'animal tracks' (ALLOc’s map 352), les courroies du joug ‘the bands of a yoke’ (map 516), les ridelles ‘walls of a cart’ (map 552), les fleurs du vin ‘yeasts that develop on the surface of wine’ (map 754), les tranches de pain ‘slices of bread used to garnish soup’ (map 1058), le pantalon ‘trousers’ (map 1131), les jarretières ‘garters’ (map 1136), les incisives ‘incisors’ (map 1175).

2 Ravier (1978:xiii, translation mine) transcribes this sound in ALLOc notation as [j] and describes it as the “weak element” of a “decreasing diphthong” which does not form its own syllable. I follow other researchers (Floricic, 2010; Sauzet, 2012) as interpreting this as IPA [j].

3 Consonants noted by ALLOc as being “weakly articulated” are not differentiated in this or subsequent maps from consonants that were articulated normally.

4 One might expect to see included in this list of realizations of /s/ its aspiration to [h], a common stage of /s/-lenition across languages (Ferguson, 1990; Seklaoui, 1989). Indeed, Map 4 will show that aspiration is a common realization of /s/ on the definite article. The lack of aspiration observed on the noun is most likely due to a prohibition on phrase-final [h] (which would appear on the noun); the lexical items elicited for ALLOc were produced in isolation, with nothing following the noun.

5 Calculations of percentages were done after N/A tokens were omitted, so for localities where N/As were recorded for one or more lexical items, the denominator of the fraction was less than eight.

6 Only one locality (24.10) ever shows deletion of /s/ on the article; it is accompanied by lengthening of the article vowel.

7 Sauzet (2012) similarly notes the correspondence between the region of vowel change in the definite article and the region of /s/ deletion on the noun, and briefly mentions a possible role of article-final [j].

Acknowledgements

Many thanks to William Labov, who provided invaluable comments at every stage of this work, audiences at the University of Pennsylvania and at NWAV38, and two anonymous reviewers for their comments.

References


Dialect perceptions in real time: A restudy of Miami-Cuban perceptions

Gabriela G. Alfaraz*

Michigan State University

Perceptual dialectology investigates nonlinguists’ beliefs about their own and other varieties. This paper fills a gap in longitudinal research in this area with a restudy of the perceptions of Miami Cubans carried out twelve years after the first study. Perceptions are examined in relation to social and demographic changes with a sample of 84 participants of Cuban origin who responded to a questionnaire about the correctness of regional varieties of Spanish. The results showed that perceptions of non-Cuban varieties remained relatively stable over time, continuing to correlate with race and poverty. Perceptions toward the Cuban Spanish of the Miami community were also stable and, as in the earlier study, were highly positive, reflecting strong beliefs in its correctness-status. In contrast, perceptions of Cuban Spanish on the island were significantly more negative; it was ranked the least correct of the regional varieties evaluated. Factors underlying perceptions are examined in relation to demographic changes, political ideology, and beliefs about race and poverty. This paper highlights the contribution of the longitudinal study of dialect perceptions to the understanding of language attitudes, intergroup relations, and language change.

1. Introduction

This paper presents a longitudinal study of the perceptions of Cubans residing in Miami-Dade County, Florida, where 65% of the population is Latino, and over half of Cuban origin, toward the correctness of varieties of Spanish, including the Cuban one. The study was modeled on an earlier one carried out in 1998 with the same population. The aim of this second study was to examine whether perceptions had changed or remained stable in relation to social and demographic changes in the community. In what follows, the goals of perceptual dialectology are summarized, the findings of the earlier study are outlined, and the Cuban community in Miami-Dade County is described.

1.1. Perceptual Dialectology

The area of study known as perceptual dialectology grew from Preston’s research on nonlinguists’ attitudes toward regional varieties of U.S. English (Preston, 1986, 1988, 1989, 1996, 1999; Long & Preston, 2002). Since its inception, studies have been carried out in a wide range of communities with diverse sociolinguistic situations and language ideologies, including English in the United States (i.e. Benson, 2003; Bucholtz, Bermudez, Fung, Vargas & Edwards, 2010; Fridland & Bartlett, 2006; Hartley, 2005; Preston, 1986, 1988, 1996) and England (Pearce, 2009), French in Canada (Evans, 2002) and France (Kuiper, 1999, 2005), Spanish in Madrid (Moreno Fernández & Moreno Fernández, 2002), at the US-Mexico border (Martínez, 2003) and in Miami (Alfaraz, 2002), German after the reunification (Dailey-O’Cain, 1999), as well as Hungarian (Kontra, 2002), Turkish (Demirci, 2002; Demirci & Kleiner, 1998, 1999), Japanese (Long, 1999), and Korean (Long & Yim, 2002), to mention some varieties and regions.

At the heart of research in this area is an interest in understanding the beliefs and attitudes of the folk, or everyday people, toward language varieties and their speakers (Niedzielski & Preston, 2000). Preston (1999: xxv) noted that among the reasons for investigating what the folk believe about regional and social dialects is that it can reveal “folk dialect areas where there are none scientifically and vice versa.” Moreover, Preston pointed out that “instances of language change and so-called language attitudes...might be profoundly influenced by folk beliefs about language, particularly beliefs about the status of language varieties and the speakers of them” (p. xxiv). Labov (2001: 191) noted that “covert attitudes and beliefs” contribute to the transmission of changes “if speakers feel that their adoption of the linguistic form will lead others to attribute to them the positive traits of the given group and allow them to share in the privileges of that group.” Studies on perceptions of dialects can help estimate the likelihood of convergence or divergence (Giles, 1973), and thus, the direction of linguistic changes in dialect contact situations (Auer, Hinskins & Kerswill, 2005; Trudgill, 1986).

Perceptual dialectology studies have generated a body of work that has contributed to our understanding.
of dialect perceptions, but longitudinal studies represent a gap in the research. With the exception of Preston (2011) and the research reported on in this paper, studies have not addressed perceptions over time, particularly in relation to social changes. Preston (2011) compared findings from twenty-some years of work on perceptual dialectology in Michigan. At the outset, he asked:

William Michigander’s language attitudes look the same twenty years later? Surely more than two decades of greater media exposure to regional varieties, particularly their use by obviously well-educated speakers, and perhaps some growing sense that severely downgrading the speech of others might not be the nicest thing one could do will have ameliorated these harsh views.

His findings led him to conclude that the group’s perceptions had not changed: “[P]rejudice against the South and perceptions of Michigan as the site where the most correct English is spoken have not changed in twenty years” (p. 7). Moreover, rather than having mitigated their evaluations of other regions, Michiganders gave harsher ratings to some regions than they had in earlier studies.

The description of findings from the first study of Miami Cubans, discussed in the next section, will show that Cubans share with Preston’s Michiganders a very high regard for the correctness-status of their own variety. The evidence Preston found for stability in the perceptions of Michiganders provides a backdrop for the study of perceptions among Miami Cubans. It motivates the question of whether the perceptions of Cubans have remained stable because of their strong investment in the correctness-status of their variety, similar to the Michiganders in Preston’s study, or whether Cuban perceptions have changed, despite the high regard for their own variety, in response to local social and demographic changes.

1.2. The First Study of Miami-Cuban Perceptions

The original study, conducted in Alfaraz (2002), examined the perceptions of Cuban-origin individuals residing in Miami, Florida. It focused on regional varieties of Latin American Spanish, presented as countries, following a common classification model that disregards that isoglosses do not coincide with national boundaries (Lipski, 1994; Penny, 2004), but which, based on evidence from a pilot study that included drawing dialect regions on maps, reflects how Spanish speakers think about regional dialects. Along with the varieties of Latin American, Peninsular Spanish was included as a variety because it is widely accepted as the standard—Spain is the seat of the institution governing Spanish, La Real Academia Española, which oversees and approves standard grammar and lexicon published in their official dictionary and grammar books.

Fieldwork was conducted in Miami in 1998, on the heels of a major wave of immigration from Cuba. The study included a stratified sample of 148 individuals who had immigrated at different time periods, or who had been born in the United States and actively used Spanish. The study and all instructions were in Spanish. Using a questionnaire, we asked participants to rate on a seven-point scale the correctness, pleasantness, and degree of difference of the varieties of Latin American and Spain. Although given the opportunity to add regions other than the countries listed, only a few participants added Andalusia, reflecting a general north-south division in Peninsular Spanish. Rather than eliciting perceptions of the Cuban variety by listing the name of the country, Cuba was listed as two varieties: One representing the variety of the diaspora (Cuba 1), and the other the variety of the island (Cuba 2). This separation of Cuban Spanish into two varieties was based on observing widely held beliefs about intragroup differences based on the dimensions of space—here (Miami-United States) versus there (Cuba)—and time—then (pre-revolution) versus now (post-revolution). When taking the survey, participants did not comment on the division and rated them without hesitation. Indeed, it was surprising to see automatic responses given to their own variety. The ratings of other varieties were straightforward, both of those with a stronger demographic presence in Miami and those not commonly encountered, as, for instance, Bolivian or Paraguayan Spanish.

The findings for regional dialects in Figure 1 confirms that the distinction made for the two varieties of Cuban Spanish captured a perceptual dialect boundary that was revealed in their ratings. The diaspora variety (Cuba 1) received the highest rating (6.0), second only to Spain (6.34), whereas the island variety (Cuba 2) received among the lowest ratings (4.03). The significance of the separation between the varieties was shown with a cluster analysis in which Cuba 1 and Spain were in one cluster, Cuba 2 in a second, and all the other regional varieties in a third. Although a linguistic boundary between the diaspora and island varieties had not been demonstrated with empirical data, there was a perceptual-dialect boundary whose function was to strengthen intragroup boundaries. Not only did the established immigrants give Cuba 1 higher ratings than Cuba 2, the newest arrivals also gave the island variety lower ratings than the diaspora one. Regardless of time of arrival to Miami (or birth there), all groups noted differences between Cubans here and Cubans there. For newcomers, the perceptual boundary was spatial (here versus there), whereas for established
groups, the boundary was both spatial and temporal (then versus now).

While these results for Cuban Spanish most obviously stemmed from political ideology, they also revealed ideologies of economics and race. For the Cuban varieties, wealth and race correlated with correctness: The island was perceived as poor, black, and nonstandard, whereas the diaspora was wealthy, white, and standard. These beliefs about wealth and race surfaced in the evaluation of other regional varieties as well. A correlation was found between positive ratings and degree of white population and, similarly, between positive ratings and greater wealth or economic development, operationalized as gross domestic product per capita. Thus, higher ratings were given to regions with larger proportions of white population and higher gross domestic product per capita, for instance, Spain (6.34) and Argentina (5.16), which ranked third after Cuba 1. On the other hand, correctness ratings were lower for varieties from poorer regions that had larger proportions of non-white or mixed-race speakers, particularly (and paradoxically), the Spanish of their two Caribbean island neighbors, Puerto Rico (3.83) and the Dominican Republic (3.73). In sum, this first study confirmed that Cubans in Miami believed in the high degree of correctness-status of the diaspora (their) variety. Moreover, the study showed a link between correctness-status, race, and economic development. Whether it continues to influence perceptions will be examined in this restudy of the community.

1.3. Cuban Presence in the Community

The Cuban enclave in Miami-Dade County predates the Cuban Revolution, but it was this event in 1959 that triggered large waves of immigration from the island to the United States and settlement of heavy concentrations of Cubans in Miami. According to the 2010 Census, over two-thirds of Cubans lived in the state of Florida and nearly half the Cubans in the United States lived in Miami. In the Miami enclave, Cubans made up more than half (53%) of the Hispanic or Latino population. In the decade from 2000 to 2010, the Cuban population increased by 26%, and immigration from Cuba, as well as from other Latin American countries fueled growth of the Hispanic population.

There were four major waves of immigration from Cuba from 1959 to 1996, with continuous immigration of smaller numbers in the interludes. The largest number of Cubans arrived in the 1960s and early 1970s: The first wave (1959–1962) brought 248,000 Cubans, the second (1965–1973) brought 297,318 more, and in the three-year interlude (1962–1965) between these waves, 56,000 more arrived. During the Mariel Boatlift in 1980, 124,776 Cubans arrived by boat, and during the last major wave, the so-called Rafter Crisis (1994–1996), 80,000 Cuban rafters and visa holders settled in the United States.

Growth of the Cuban population in Miami continued at a rapid pace in the ten-year period between 1995 and 2005. During this time, over 200,000 Cubans were estimated to have immigrated to the U.S., among them rafters (balseros), people smuggled in on speedboats (boteros), and winners of the U.S. government visa lottery (el bombo) (Henken, 2005). According to the 2010 Census, the number of Cubans in the U.S. increased from 1.2 million in 2000 to 1.8 million in 2010. While agencies resettled Cuban arrivals to various parts of the United States, an overwhelming number settled in Miami-Dade County or migrated there from the places they had been sent during the resettlement process.

The Cuban population in Miami-Dade numbered 856,007 at the time of the 2010 Census. The attraction of the area for newcomers, as García (1996: xi) described it during the mid-1990s, continues to explain its appeal...
to Cubans, whether new arrivals or established immigrants:

The new immigrants who will settle in Miami will find the city familiar and yet alien. Miami is certainly not a Cuban city, but it is home to the second-largest Cuban population in the world, and new immigrants will find it enough like home to want to settle there— or return there, if things do not go well for them in other parts of the United States. Miami will serve as a prism through which to interpret the U.S., and hopefully a buffer to shield them from the harsher aspects of the adaptation process. They will be among compatriotas who know what it’s like to leave one’s homeland in search of better options.

Although Miami-Dade has historically been an enclave for Cuban immigrants, it attracted immigrants from Latin American countries as economic instability and political turmoil increased migration from the region. The 2010 U.S. Census reported that, between 2000–2010, the number of South Americans grew by 77% and the number of Central Americans by 65%. Like Cubans, immigrants from other Spanish-speaking countries often prefer to settle in Miami rather than venture beyond the bilingual city to other regions, where they will, most certainly, need to learn English in order to work and carry out public interactions.

The immigration of Spanish-speaking groups has contributed to a dialect contact situation in which Cubans are increasingly in contact with diverse varieties of Latin American Spanish. Moreover, it can be said that heavy recent immigration from Cuba has created a situation in which Cuban Spanish is in contact with itself, in the sense that the variety of earlier immigrants is in contact with the variety of newcomers.

1.4. Goals

The goal of this research was to examine dialect perceptions in real time. To this end, the first study was replicated with a twelve-year time depth to investigate perceptions in relation to changes in the community, including significant growth in the number of immigrants from Latin American countries and an influx of newcomers from Cuba that contributed to the growth of the Cuban population in Miami-Dade County by 44% between 2000 and 2010. In the remainder of this paper, the status of perceptions toward regional varieties of Spanish and toward varieties of Cuban Spanish are described and later compared to the findings of the earlier study.

2. Design and Analysis

Classic perceptual dialectology studies have used a variety of techniques to elicit evaluations, including hand-drawn maps, which have been a rich source of data about folk beliefs (Hartley & Preston, 1999; Preston, 1986, 1996), and questionnaires about the degree of difference, and the correctness and pleasantness of varieties, terms which correspond to the notions of status and solidarity (Lambert, 1967); in fact, in Preston (2011), we find correctness expanded to correctness-status and pleasantness to pleasantness-solidarity.

As noted above, the original study of perceptions among Cubans in Miami used a questionnaire, and the same questionnaire was used in the restudy. The questionnaire listed the Spanish-speaking countries of Latin America and Spain alphabetically, with the order reversed on half the questionnaires, and, as in the first study, Cuba was listed as Cuba 1 and Cuba 2. A seven-point scale was used to rate correctness and pleasantness, presented separately, but with correctness being presented first. As in the initial study, participants completed the task without questioning the options. Once again, the separation of Cuba 1 and Cuba 2 was not questioned. Participants clearly enjoyed filling out the questionnaire and most talked aloud as they gave ratings, explaining their choices and justifying them with impressions and anecdotes, providing a rich source of qualitative data similar to that obtained when the think aloud method (van Someren, Barnard & Sandberg, 1994) is used. In both this and the first study, the confidence with which participants gave ratings was striking. Although the actual numbers given may have varied somewhat, the talk surrounding the ratings was similar and demonstrated the cohesiveness of the group’s perceptions.

The questionnaire was written and delivered in Spanish. Participants were recruited if they actively spoke Spanish in their daily lives, had acquired Spanish as a first language, or had acquired it simultaneously with English. Verbal invitations to participate, questions, and explanations were in Spanish, or in Spanish and English, depending on whether the participant was monolingual or bilingual. The study was conducted by a Cuban-American bilingual whose variety during data collection can be described as a standard variety of US Cuban Spanish, with a fair degree of leveling of marked phonetic features, and accommodation to participants’ varieties, particularly popular US Spanish (Otheguy, 2009, 2010) when situations became more casual. Participants were recruited in public spaces, including retail stores, as well as through acquaintances, friends, and colleagues. As in the first study, participants were often hesitant to give social information, even after being assured of their anonymity.

The sample contained 84 Cuban-origin individuals. As in the previous study, an attempt was made to survey Cubans of varying ages and socioeconomic statuses who had arrived from Cuba at different periods of time, as well as Cuban-Americans born in the US.
of Cuban parents, but these were only included in the study if they reported actively participating in the Spanish-speaking community. The social factors studied included gender, age, socioeconomic status (SES), social network, and year of arrival from Cuba. Gender was considered a binomial category made up of male and female. Age groups were established according to ten-year spans, with five groups created: 20–29, 30–39, 40–49, 50–59, and 60+. Socioeconomic status reflected a combination of schooling and type of employment; four groups were created representing the upper middle class, lower-middle class, upper-working class, and lower-working class. Network groupings were based on participants’ reports about the number of Cuban friends and coworkers they interacted with, resulting in four groups: All Cuban, mostly Cuban, half Cuban, or contains few or no Cubans.

We also asked the year in which participants arrived in the US. Four groups were created according to age of arrival from Cuba. These groupings generally followed the major waves of immigration described earlier. The first group contained immigrants who arrived between 1959 and 1978, during the first and second major waves, which occurred between 1959–1962 and 1965–1973, and in the slow but continuous immigration that occurred between and after the major waves. The second group consisted of Cubans who immigrated during the years between 1979–1992; this group had people from the third major wave, through the Mariel Boatlift in 1980, as well as others who filtered in during the remainder of the 1980s, often from third countries. The third group represented immigrants who arrived during the fourth major wave, the Rafter Crisis, and covered the period 1994–1999. In this restudy, a fourth group was added that included immigrants who arrived between 2000 and 2010, after the four major waves.

Statistical analysis of the quantitative data from the questionnaires was carried out using SPSS, with a variety of statistical tests run on the data, including t-tests, correlations, multidimensional scaling, and k-means cluster analysis. In the following section, the results for Cuban Spanish and the other regional varieties studied are presented separately. The results for other varieties are presented first, followed by those for Cuban Spanish. After these results, a comparison of the findings for correctness from this study and the earlier study is discussed.

3. Results

3.1. Regional Varieties

The means of the correctness-status ratings for regional varieties are shown in Figure 2. It was not surprising that Peninsular Spanish received the highest rating (6.31), which confirmed that it was regarded as the most prestigious variety, representing a model of correct usage. None of the Latin American varieties reached the six-point range in their rankings; in fact, the next highest scores were in the low five-point range: Argentina (5.16), Chile (5.11), and Costa Rica (5.07). These were followed with varieties in the four-point range: Venezuela (4.83), Colombia (4.80), Panama (4.30), Uruguay (4.23), Peru (4.08), and Ecuador (4.04). The rest of the regional dialects received ratings in the three-point range, including Paraguay (3.99), El Salvador (3.92), Mexico (3.90), Honduras (3.81), Bolivia (3.73), Guatemala (3.64), Nicaragua (3.49), Puerto Rico (3.33), and the Dominican Republic (3.23).

These perceptions of correctness followed a general regional trend that is evident in the multidimensional scaling in Figure 3. The prestige-correctness of Peninsular Spanish is reflected in the placement of Spain in the upper left quadrant, set off from other regions. The Caribbean varieties of Puerto Rico and the Dominican Republic, rated the two least correct, are furthest to the right on the x-axis. In comments about these regions, participants indicated that the degree of incorrectness stemmed from their association with large mulatto and

Figure 2. Ratings of regional varieties without Cuban Spanish.
African-origin populations. The Central American varieties of Costa Rica, Panama, El Salvador, Honduras, and Guatemala are in the upper section, but Nicaragua, another Central American variety, is in the lower right section. It is, however, different from the other Central American varieties in that it has had a strong presence in the community. The US’s largest group of Nicaraguans resides in the city of Sweetwater in the Miami-metropolitan area, where signage in Spanish reflects the group’s presence, and a public park and public middle school are named after a famous Nicaraguan poet, Rubén Darío. Sweetwater, or Pequeña Managua (Little Managua), is surrounded by other suburban metro cities with heavy concentrations of Cubans. Plotted along with the Central American regional varieties are the South American varieties of Bolivia and Ecuador; their being grouped with Central American varieties reflects a perception of incorrectness that, as with the insular Caribbean varieties, participants attributed to large concentrations of mestizo and indigenous populations. In contrast, the varieties of the Southern Cone, Argentina, Uruguay, and Paraguay, are in the lower section, and Chile is toward the bottom of the upper one. As noted above, Argentina and Chile received the two highest correctness ratings among Latin American regions. The plotting of Cuban Spanish, discussed below, will provide further insight on this spatial arrangement of these varieties.

To understand the influence of social factors on perceptions, age, gender, socioeconomic status, social network, and year of arrival were analyzed using t-tests. All but gender were found to be significant. The results for age (p < .001) showed that the oldest participants, 60 years and older, gave the highest ratings (4.54), and there was a trend for ratings to become lower as age decreased: 50–59 years 4.38, 40–49 years 4.20, 30–39 years 4.08, and 20–29 years 3.55. Despite an expectation that the oldest groups would give lower ratings, demonstrating less tolerance to varieties other than their own, it was the youngest group that gave the harshest ratings. Linguistic evaluations can be subject to age-effects, according to Garrett, Coupland, and Williams (2003:85-86), who noted that age may influence the judgments of younger adults, whose attitudes are still developing; furthermore, younger people may be inclined to oppose the views of the dominant group. The lower ratings of younger people may be related to age, but they can also indicate a change in the community toward harsher evaluations of non-Cuban dialects.

The results for socioeconomic status (p < .001) indicated that participants from the highest social status group (upper middle) gave the highest ratings (4.57), whereas those from the lowest social status group (lower working) gave the lowest (3.92). The groups in the middle had scores in between these two ranges: Lower middle (4.28) and upper working (4.35). Competition for employment and resources may explain why participants in the lowest socioeconomic group gave the lowest ratings. In the struggle to secure employment as unskilled laborers, these Cubans may find themselves competing with immigrants from other parts of Latin America, which may cause them to regard less favorably Latin America varieties in general. Along with this economic component, the linguistic insecurity of this group may contribute to it downgrading other varieties in relation to its own.

Along with age and socioeconomic status, the composition of the social network significantly influenced ratings of regional varieties (p < .01). Although the differences here were small, the results showed that Cubans with more closed networks had more favorable views of the correctness of other varieties than Cubans with more open networks. The most favorable ratings were given by people who reported networks that were entirely made up of Cubans (4.52), and the second highest were from participants with mostly Cubans in their network (4.33). Varieties were judged more harshly if the network was half Cuban (4.00), or if it had few or no Cubans (4.09). Thus, rather than mitigating negative views of other dialects, these findings suggest that regular interactions with speakers of other varieties contributed to negative evaluations. The implications for language change are clear: Open networks are not likely to result in linguistic accommodation that diverges from the Cuban norm. The findings for Cuban Spanish, discussed below, will add another layer to this prediction of divergence from other regional varieties in this community.

Year of arrival or birth in the U.S. also showed significant differences in ratings (p < .001). The group born in the U.S. gave the lowest ratings (3.92); the two groups of immigrants that arrived after 1994 had similar ratings—2000–2010 (3.99) and 1994–1999 (4.00)—and earlier immigrants gave higher ratings—arrivals from 1959–1978 (4.45) and 1979–1992 (4.59).
Viewed in terms of competition among groups in the community, the more favorable ratings of established immigrants reflect their greater social and economic stability, whereas the newer arrivals are more likely to be in direct competition for jobs and resources with newcomers from other regions, but this does not necessarily explain the low ratings of US-born Cubans, who are represented across socioeconomic groups. The evaluations of this group are discussed further in relation to the results for the two Cuban varieties.

3.2. Cuban Spanish

Including the results for Cuban Spanish in the ranking of regions, shown in Figure 4, reveals that Cuba 1 had the second highest score (6.12) after Peninsular Spanish (6.31), which made it the highest ranked for correctness of the Latin American varieties. The other Cuban variety, Cuba 2, on the other hand, received lower evaluations (3.14) than any of the other varieties. These results confirmed the community’s high regard for the correctness-status of the variety represented in Miami and its low regard for the variety of the island.

Multidimensional scaling revealed interesting patterns in the spatial representation of the two Cuban varieties in relation to other regional varieties. In Figure 5 we see that Cuba 1 is on the far left below Spain on the x-axis, whereas Cuba 2 in the far upper right, directly above the Dominican Republic and Puerto Rico, and closest to Bolivia. This plot also indicates with circles the clusters derived from a K-means cluster analysis with three clusters specified. The first cluster is made up of Spain and Cuba 1, positioned to the far left, set off from other regions. The second cluster includes all the non-Cuban varieties. The single member of the third cluster is Cuba 2. These results suggest that perceptions center on the opposition between the two varieties of Cuban Spanish. The prestige of Cuba 1, measured against Peninsular Spanish as the prescriptive norm, contrasts with the low status of Cuba 2, whose distinctness sets it off from other Latin American varieties. Thus, the often cited belief that Cuban Spanish on the island has become increasingly impoverished over time underlies a perceptual boundary that marks the political and ideological separation of Cubans in Miami from Cubans on the island.

In contrast to the influence of social factors on the ratings of non-Cuban regional varieties, t-tests showed that only age (p < .001) and year of arrival (p < .05) were significant for Cuba 1, and none were significant for Cuba 2. The evaluation of Cuba 1 across age groups showed that the youngest group (20–29) gave it the lowest rating (5.0) and the oldest group gave it the highest (6.49). Other age groups were between these two extremes: 30–39 (6.09), 40–49 (5.67), 50–59 (6.00). The same age distribution was found for the rating of the non-Cuban regional varieties: The oldest group gave the highest, and the youngest, the lowest scores. As suggested above, this pattern may be age-graded, related to changes in individuals over time, as noted in Garrett et al. (2003), or it may reflect a change in perceptions in the community. If this is, in fact, age-graded, then younger
participants are likely to change their attitudes as they age. On the other hand, if these ratings for Cuba 1 reflect community change, then they are indicative of less favorable attitudes encroaching on perceptions of the highly esteemed variety of the Cuban diaspora community in Miami.

For younger Cubans who arrived as children or were born in Miami, Cuba 1 is a variety of U.S. Spanish that interacts with English, and their less favorable evaluations of it may be based on a negative attitude toward contact features and salient bilingual practices, such as borrowing and codeswitching. Similarly, younger speakers who immigrated from Cuba agree that Cuban Spanish in Miami is prestigious, but for many of these speakers, particularly the ones who are not bilingual, new features derived from contact with English that they have encountered are highly salient and mark differences between their native variety (Cuba 2) and the variety in Miami (Cuba 1). These differences figure prominently in conversations about language and language learning in the U.S. In some of the interviews, newcomers reported, in a humorous tone, that they needed to learn to speak *Spanglish* in Miami, and that they were forgetting to speak Spanish but had not yet learned English.

The results for year of arrival revealed that Cuba 1 received the highest ratings from the newest arrivals from 2000–2010 (6.38) and the oldest from 1959–1978 (6.37). Following these were arrivals from 1979–1992 (5.93), 1994–1999 (5.64), and the US-born group (5.63).

While it was not surprising that the longest-established residents gave Cuba 1, their own variety, more favorable ratings, it was unexpected that the newcomers, who cannot claim it as their native variety, gave it similarly high ratings. In the ratings for Cuba 2, although not significantly different, Cubans born in the U.S. gave the lowest ratings (2.75), followed by arrivals from 1959–1978 (3.08). The ratings were somewhat higher for the other year of arrival groups: 1979–1992 (3.40), 1994–1999 (3.18), 2000–2010 (3.56). In contrast to Cuba 1, and to other regional varieties, there is consensus across groups about the status of Cuba 2.

### 4. Real-Time Comparisons

#### 4.1. Regional Varieties

A comparison of the ratings of correctness for regional varieties from the 2010 and 1998 studies revealed a generally stable pattern. The results in Figure 6 show that ratings were somewhat lower in the second study than in the first—fourteen of the nineteen varieties had lower scores—and that the regions with the lowest scores were the ones that received harsher ratings in the second study, including Bolivia, Guatemala, Nicaragua, Puerto Rico, and Dominican Republic. Increases in correctness ratings at the second time period were minimal and limited to regions whose correctness had been rated more highly in the first study. Overall, however, differences in ratings over time were not significant.

Interesting differences between varieties were found when ratings were converted into regularized scores. This was done based on a percent calculated from the highest rating, given to Spain, and a rate change score derived from the difference between the percentage points for the two time periods. The regularized scores in Figure 7 are of changes in the ratings between 1998 and 2010. This figure also shows demographic changes in Miami during the period studied as a percent reflecting differences in the total population of individuals who claimed regional origin in the country (from the U.S. Census Bureau figures for 2000–2010).

The regularized rating change confirms that rating changes were negative for all varieties except those of Panama, Costa Rica, Chile, Venezuela, and Argentina.
Negative rating changes were largest for countries whose correctness-prestige was lowest in the first study: Bolivia (−11.08), Guatemala (−8.40), Nicaragua (−7.94), Puerto Rico (−7.64), and the Dominican Republic (−7.64). Negative changes were also found for Ecuador (−6.01), Colombia (−5.0), Uruguay (−3.94), Paraguay (−3.49), El Salvador (−3.02), Honduras (−2.4), and slightly less negative scores were given to Mexico (−1.28) and Peru (−1.11).

By far the greatest rate change was found for Bolivia, whose correctness score dropped 11.08 percent. This has clear political underpinnings: The ascent of Evo Morales, who had strong ties to Cuban leader Fidel Castro and the deceased Venezuelan leader, Hugo Chávez, was an event that ran directly counter to the ideological position of the group. Why then was Venezuelan Spanish given a slightly more positive rating in the restudy, or at least not given a harsher one, when it also experienced political changes toward the left with Hugo Chavez, elected to office in 1999, at the helm? The answer is that Cubans are sympathetic to the plight of the tens of thousands of Venezuelans who arrived in Miami after Chavez became president. According to the U.S. Census, the number of Venezuelans in Miami grew 117%, from 21,593 to 46,851, between 2000 and 2010. Venezuelan politics are followed as closely in local, Cuban-dominated, Spanish-language media outlets as the political events in Cuba have been since the settlement of the diaspora community.

Although we can point to the growth of the Venezuelan community in Miami in the years between the first and second study as a factor that mitigated the political one, demographic factors alone did not significantly account for ratings of regional varieties. When changes in ratings were examined in relation to changes in the population of groups in Miami, no significant link was found between them ($R^2 = .05$, $F(1, 16) = .83, p = .38$). Demographic presence accounted for only 5 percent of the change in perceptions observed between 2000 and 2010. Thus, in the dialect contact situation, there were factors beyond the interplay of ingroup and outgroup dynamics, evident in demarcation of group boundaries and competition for resources, that influenced attitudes. Political ideology was clearly an important factor, and it interacted with ideologies of race and wealth to shape perceptions towards regional varieties.

The earlier study reported that economic development influenced perceptions. A positive correlation was found between evaluations and gross domestic product per capita ($r(19) = .67, p < .01$). To examine whether changes in economic development continued to correlate with ratings of the regions, a correlations test was run on the 2010 ratings and recent GDP per capita, which had increased over the last decade in all regions except Argentina. Once again, a strong positive correlation ($r(19) = .61, p < .01$) was found between perceptions of regional varieties and wealth of the region.

Racial composition of regions was another factor correlating with perceptions in the previous study ($r(19) = .58, p < .01$). To determine whether race continued to be a relevant factor, a correlations test was carried out on the 2010 mean scores and racial composition, calculated as the proportion of white versus non-white population. Again, the results showed a positive correlation ($r(19) = .61, p < .01$) between perceptions of varieties and degree of racial diversity in the region. As in the earlier study, countries that were predominately indio or mestizo had lower ratings, and countries where large segments of the population were of African origin received the lowest ratings. Thus, the low ratings given to some Central American varieties,
such as El Salvador, Honduras, Guatemala, and Nicaragua, some South American varieties, namely Paraguay, Ecuador, and Bolivia, and the Caribbean varieties of Puerto Rican and Dominican Spanish, can be explained as a downgrading based on race. As will be discussed below, race and poverty are factors that, along with politics, also played a role in shaping beliefs about the correctness-status of both Cuba 1 and Cuba 2.

4.2. Cuban Spanish

Comparing the perceptions of the Cuban varieties in the two studies, in Figure 8, shows that Cuba 1 held its place as the second most correct variety. Indeed, the rating of Cuba 1 remained generally stable, increasing slightly from 6.00 in 1998 to 6.12 in 2010. Cuba 2, on the other hand, was ranked higher among the varieties in the earlier study, where it ranked 6th ahead of Nicaragua, Mexico, Honduras, Puerto Rico, and Dominican Republic, but in the restudy, it ranked last. Ratings of Cuba 2 dropped from 4.03 in 1998 to 3.14 in 2010, when it received the lowest correctness ratings of all varieties.

The ranking of the change in ratings over time for all the regional varieties in Figure 9 shows differences in the stability of perceptions of the Cuban varieties. First, it is clear that Cuba 1, compared to itself over time and compared to most other varieties, had remained generally stable, with only a small increase of 2.35%. Thus, the belief in the prestige and status of Cuba 1 had held strong, its resistance to change firmly grounded in its ideological function. In contrast, Cuba 2 had the largest change of all varieties, decreasing 13.8 percentage points. Only Bolivia, which experienced a political move to the left, had a double-digit decrease in the ratings. It is clear that Cuba 2 was heavily downgraded because it is on the opposite side of the ideological divide separating Miami-Cubans from their homeland.
Moreover, in Miami these perceptions also function to mark boundaries between established Cubans and newcomers. Downgrading the variety of newcomers, Cuba 2, is an overt rejection of that group. García (1996:xi), whose book on Cuban immigrants in Miami was published during the Rafter Crisis when immigration from camps in Guantanamo was in progress, forecast tensions that exist today between established immigrants and new arrivals as a result of their racial, sociocultural, and political differences:

This is not to say that there will not be conflict. Emigrés of the first wave (1959–62), disproportionately white and middle class, will find it difficult to relate to the new immigrants, whom they consider rough, poor, and uneducated. The fact that many of the new immigrants are black or of mixed racial heritage, and were once the faithful revolutionary proletariat, widens the cultural chasm. Already, the term balsero has become a pejorative among the older emigrés, a way of differentiating themselves from the new arrivals, just as the term marielito (for the Cubans of the third wave) acquired currency in the early 1980s for many of the same reasons. The arrival of more than twenty thousand new immigrants each year, most of whom will probably stay in south Florida, will also exacerbate ethnic tension in Miami.

Thus, language serves to mark boundaries between established Cubans and newcomers. Led by the older immigrants from the early first wave, for whom Cuba 1 is the native variety, established Cubans point to the variety spoken by new arrivals (Cuba 2), with its abundance of unfamiliar lexical items, and perhaps with phonetic and grammatical changes that advanced over the course of forty-some years, and note it a corrupt form of the language they knew. Certainly, this is not a surprising discovery, given that older generations tend to believe their language is deteriorating in younger generations. What makes this an interesting finding is that contributes to our understanding of group relations is the separation of the groups in time and space and their reencounter in one location, the unfamiliar new group having encroached on the physical and linguistic space of the older group, which deploys its power and prestige to maintain its position and uses its beliefs about the correctness-status of its language to fortify the boundaries separating the two.

What is remarkable in the evaluation of the Cuban varieties is that newcomers are complicit in upgrading the diaspora variety and downgrading the island one. In the earlier paper, I noted that this “appears to reflect their desire to disassociate from the island in order to avoid the negative sentiments” of established Cubans; meanwhile, they “signal their loyalty to the Miami group by accepting its attitudes toward Cuban Spanish on the island.” Whether as an outcome of hegemony, or perhaps driven by their own linguistic insecurity, newcomers embrace the correctness-status of Cuba 1. Newcomers’ attitudes toward Cuba 1 signal their acceptance of the community’s predominant ideology and alignment with the values of established immigrants, not necessarily political, but in terms of the aspirations of immigrants in general, including economic stability, home ownership, access to education for oneself and one’s children, and other indicators of successful adaptation.

In a comparison of the means by year of arrival at the two time periods, shown in Figure 10, we see that the stability of Cuba 1 contrasts with changes in perceptions of Cuba 2. Ratings for Cuba 2 of Cubans who arrived during and after the Rafter Crisis (1994–1999) were considerably higher in 1998 (5.1) than in the restudy (3.18). They appear to have accommodated to the perceptions of earlier immigrants. When the first study was conducted, this group had only recently arrived. It was a large group with strong connections, forged in planning and leaving Cuba or in the camps at Guantanamo. They also had strong ties to Cuba—many were the first in their families to emigrate, leaving parents, spouses, siblings, and children behind. Over time, however, they brought their families and settled into life in Miami in the same way earlier immigrant groups had done years earlier. A decade or so later, the members of this group had less in common with new arrivals from the island than with the established groups, who also had homes, families, and generally shared the same values and lifestyles. Interview data confirmed that members of this group believed they were more like Cubans who arrived in earlier waves than like the newcomers. As for the newest arrivals, the findings of this study showed that they have
appropriated the diaspora variety and expressed negative perceptions of the variety on the island.

4.3. Language Change

To return to the significance of perceptions for language change, it was noted in the introduction that the beliefs of nonlinguists contribute to formation of language attitudes that can influence linguistic change (Preston, 1999). Based on the findings described here for regional varieties and for Cuban Spanish, predictions can be made about the direction of changes in the Miami-Cuban speech community. The implications for movement toward leveling in the dialect contact situation are clear: It is unlikely that Cubans will accommodate linguistically to the fast-growing groups of speakers of other regional varieties in the community. In fact, one possible outcome is that the Cuban group will move in the opposite direction, diverging from the others with features, frequencies, and constraints that differentiate Cuban Spanish from other varieties. Nonetheless, multidimensional scaling and cluster analysis, in both this and the earlier study, showed that Cubans are not very concerned with other regional varieties, but rather, that they focus their attention on Cuban Spanish, on the varieties of the diaspora and the island. Given the harsh downgrading of the island variety, it is doubtful that the group will accommodate through convergence to the variety of newcomers. It is very likely, however, that the new arrivals will accommodate to the variety of the diaspora community, similar to what was found for the perceptions of arrivals from 1994–1999. Testing this prediction will depend on uncovering the linguistic features, beyond lexical ones, that serve to mark differences in the varieties of Cuban Spanish of established groups and newcomers. Taken together, the results outlined above showed a robust and stable belief in the prestige of the diaspora variety that may inhibit accommodation and leveling, despite ongoing social and demographic changes in the community.

5. Conclusion

Longitudinal research on perceptions in the Cuban community in Miami showed patterns of stability and change. Its findings suggest that perceptions may remain stable in communities in which they are firmly rooted in ideology, whether sociopolitical, as shown for Miami Cubans, or linguistic, as Preston has demonstrated for Michiganders. The restudy of perceptions in the Cuban community in Miami raises questions for future research. The most obvious, perhaps, is whether the distinctions Cubans make between varieties of Cuban Spanish is a perceptual dialect boundary that reflects an ideological one separating the diaspora from the island, or whether it is an actual dialect boundary separating established immigrants from the original speech community. Thus, the relationship between perceptions and production needs to be investigated to determine whether the perceptual dialect differences observed here correspond to a linguistic reality. Research on production data needs to be carried out to understand linguistic changes in the diaspora speech community and compare them to those of the sending community. Future research needs to explore further the findings highlighted here about the influence of political, racial, and economic ideologies on the perception of dialects. Moreover, longitudinal research is needed to explore the linguistic and social factors that motivate stability or change in perceptions over time in different communities. Studies in real time, coupled with production studies, will contribute to our understanding of the influence of folk perceptions on linguistic changes.

References


Northern dialect evidence for the chronology of the Great Vowel Shift

Hilary Prichard*
University of Pennsylvania

This paper demonstrates how the tools of dialect geography may fruitfully lend a new perspective to historical data in order to address the lingering questions left by previous analyses. A geographic examination of Survey of English Dialects data provides evidence in favor of a push-chain analysis of the Great Vowel Shift, in which the Middle English high-mid long vowels raised before the high long vowels were diphthongized. It is also demonstrated that the so-called “irregular” dialect outcomes, which have previously been cited as evidence for a lack of unity of the Great Vowel Shift, are no longer problematic when viewed in the light of a theory of dialect contact, and can in fact refine our understanding of the chronology and geographic extent of the shift itself.

1. Introduction

After over a century of debate, the chronology of the Great Vowel Shift (GVS) is still unclear. This paper represents a departure from past lines of argumentation in its approach to the questions surrounding the GVS; rather than arguing purely from theory, or from the evidence of a few isolated data points, the geographic patterns of dialect data from across the entire north of England are examined here with the goal of testing previous hypotheses of the GVS chronology. Applying arguments from modern variationist studies to historical data lends a new perspective to the questions surrounding this sound change. The results are twofold: Broadly speaking, this research presents evidence that unexpected dialect outcomes need not challenge a unified chain shift model of large-scale sound change, while in particular, the data presented here supports a push-chain analysis of the GVS, à la Luick (1896).

Over the course of the past century, countless scholars have relied heavily upon differing interpretations of these nonstandard dialect outcomes. While Luick posited a unified chain shift model of the GVS in which this unshifted ME ə found in some northern dialects was seen as the key to understanding the ordering of the changes, some later scholars have tended to interpret the dialect data as posing a problem for a unified model of the shift (e.g., Stockwell & Minkova, 1988; Smith, 2007). Their general contention is that these nonstandard reflexes are the exceptions which disprove the rule, and they have accordingly supplied alternate theories which do not require all the different changes involved in the GVS to be dependent upon one another.

The underlying assumption behind these objections seems to be that, in order for the GVS to be considered a unified phenomenon, it must have proceeded via the same mechanism in every dialect where shifted forms are found. However, we assume that there are several different possible mechanisms of change. Variationist studies have identified diverse changes which are argued to be the product of such mechanisms as levelling, or lexical diffusion, rather than regular Neogrammarian sound change. This paper is based in a framework proposed by Labov (2007), which views sound change as the result of two distinct mechanisms, transmission and diffusion. Within this framework, we will explore whether the nonstandard northern dialect forms may be viewed as the unsurprising product of diffusion. This is a point previously raised by Dinkin:

Stockwell and Minkova’s citing of regions of England in which some GVS shifts took place but not others does not in principle vitiate the traditional model of chain shifting—the conclusion is merely that those may be regions to which
the GVS diffused, rather than ones where it originated as a chain shift. (2012:755)

The distinction between transmission and diffusion will be further discussed below, but the key idea is that highly structured changes—such as the GVS—retain their internal structure when they are the product of transmission, but break down and tend to show irregularities when they are diffused. Given this, an examination of the dialect data should be able to reveal where the GVS was transmitted as a highly structured chain shift, and where diffusion has resulted in a breakdown of this structure.

2. Background

2.1. The Great Vowel Shift Debate

The ongoing debate surrounding the status of the Great Vowel Shift originated with the work of Luick (1896) and Jespersen (1909). While it was Luick’s work which in essence sparked a century-long debate, Jespersen’s conception of the shift has since become more well-known and widely accepted (see, e.g., Johnston, 1992; Lass, 1976:57; Stockwell & Minkova, 1988). Both saw the shift as a series of connected, interdependent changes, but differed in their view of the order in which the changes took place chronologically. Jespersen, pointing to parallel changes in other Germanic languages, posited what is now termed a pull- or drag-chain model, in which the initial event which caused the shift was the diphthongization of the high vowels (Figure 1a).

In contrast, Luick argued that the mid vowels ē and ō initiated the shift by raising and displacing the high vowels, a push-chain. The crux of Luick’s argument rests on the fact that certain modern English dialects retain ME ā while still showing other parts of the shift. He took this as evidence that the shift cannot have been initiated by the high vowels, since such a model cannot account for why ā only diphthongized in some dialects and not others. He demonstrated that a push-chain model better accounts for this fact; the reason ā did not shift in these dialects is that they had undergone an earlier change in which ME ō was fronted (Figure 1b). Under a push-chain model, this prior ō-fronting effectively prevented the back vowels from participating in the chain shift by not supplying the pressure necessary for ā to diphthongize.

Much work on the GVS has followed Luick and Jespersen in positing that the shift proceeded as a coherent, “unitary” change; in which the back and front vowel shifts are necessarily related and the individual front and back vowel changes are structurally dependent upon one another. Such work has been united in attempting to explain the underlying phonological mechanism of the shift from this basic premise, while differing in whether the individual changes are chronologically ordered or occurred simultaneously. Indeed, several scholars have presented analyses of the GVS using one or more structured rules (e.g., Chomsky & Halle’s 1968:256 ordered exchange rules, Carter’s 1975 abstract schema, Lass’s 1976 rules and metarules), which, depending upon their implementation, may be interpreted as giving rise to either simultaneous or ordered changes. In spite of these analyses, it is not obvious that the GVS was ever implemented as a synchronic rule, or that the structured rules of Chomsky and Halle, for example, should be interpreted as such. Furthermore, the simultaneity argument may be driven more by the needs of phonological theory than the available evidence (see Lass, 1976:65). This argument hinges on the idea that any temporally-ordered changes would necessarily have led to vowel class mergers rather than the overall shift which did occur (Stockwell & Minkova, 1988:366). However, real-time studies of current chain shifts in progress have demonstrated that such shifts need not lead to merger (e.g., Labov et al., 2006). Lass (1976:71) argues that this type of change may even be accommodated in a rule-based approach by stipulating a “no collapse” output condition.

While the proponents of such rule-based accounts have had to redefine what it means for a shift to be “unitary” when no single traditional rule can satisfactorily account for all of the changes involved, they have still largely sought to preserve the idea of unity and coherence in their analyses of the GVS; Lass (1992:148) in particular defends Luick’s concept of innere Zusammenhang (internal coherence), and sees the traditional conception of the GVS as a pattern “too valuable to discard.” By contrast, a significant minority of scholars do not agree that the unitary status of the GVS is above question, calling attention to irregularities in the dialect data in order to challenge the idea that the shift was in any sense unitary. Stockwell and Minkova (1988) put forth a scathing review of previous work, concluding that the GVS is “the linguist’s creation through hindsight.” Their objection which is most relevant here is what they term “the dialect problem.” Briefly, their contention is that “[t]he vowel shift did not proceed along the same lines in the dialects as it is assumed to have done in London,” and

![Figure 1. The Great Vowel Shift (a) and the GVS in the north (b).](image-url)
the wide range of apparently irregular outcomes of the shift found in the dialect data is evidence against a unitary chain shift model of the GVS. Stockwell and Minkova are not the only scholars who have pointed out irregularities in the dialect data; Western (1912) and Boisson (1982) have both cited examples which seem problematic for a unitary model of the GVS. As we shall see below, it is irrefutably the case that the Middle English vowel reflexes are different in the north than they are in the south; however, this fact does not necessarily rule out a unified account.

A rather different take on the question of the status of the GVS in the north comes in the form of Smith’s (2007) contention that the GVS, as it is usually conceived, did not happen in the north at all. Rather, he argues for two separate shifts—a Southern Shift (the GVS as we know it), and a separate Northern Shift (which was in essence just the front half of the GVS)—noting that,

[[It is worth recalling that similar outcomes in the history of sounds can often derive from diverse origins. It is thus quite possible that the northern and southern shifts, though in some ways similar, could be differently triggered […]. (Smith, 2007:140)

This is certainly a fair point, and would be interesting if it could be proven; however, Smith’s argument for the independent triggering of these two shifts is at best weakly supported by the sources he cites. It is equally true, as demonstrated below, that these similar outcomes can be explained as deriving from the same source; why then posit two shifts, when one will do the job?

2.2. A Variationist Perspective

Labov (2007) proposes a resolution to the tensions between the family tree and wave models of linguistic change in the form of two different mechanisms of change. The first, transmission, is linguistic descent of the type modeled by the family tree; in its prototypical form, it consists of faithful transmission from generation to generation via first language acquisition in children. By contrast, diffusion occurs in contact situations between adults, and is thus expected to show more irregular outcomes than transmission. Labov demonstrates this distinction clearly in his examination of the difference between the regular outcomes of transmission of the Northern Cities Shift throughout the Inland North dialect area versus irregularities resulting from the diffusion of this chain shift along the St. Louis corridor. Further work by Dinkin (2012) on the diffusion of the Northern Cities Shift in New York supports the idea that when chain shifts diffuse, the structural relationships cease to hold, resulting in irregularities.

The research presented here will investigate whether the GVS dialect irregularities noted in the literature may similarly be explained as the result of the diffusion of a unitary chain shift. In particular, the maps presented below are analyzed in terms of the geographic relationships between the modern reflexes of the ME vowels, drawing on the common isogloss relations—bundling, complementation, and nesting—found in dialect geography. As the GVS is traditionally conceived of as a chain shift, the nesting relationship, in which “the spatial distribution of one feature is contained entirely within that of another, establishing an implicational relationship” (Labov et al., 2006:44), will be of particular importance.

3. Data & Methods

This project makes use of data taken from Eduard Kolb’s 1966 work, the Linguistic Atlas of England: Phonological Atlas of the Northern Region (PANR). The data for this atlas was collected between 1950 and 1961, as part of the Survey of English Dialects (SED) project. Kolb hand-picked raw data from the SED interviews in order to present a cohesive picture of the patterns of variation in northern phonology. In the PANR, SED data is presented as a series of maps, which each display the phonetic realization of a single lexical item in the eighty different northern SED locales. The maps are organized into groups based on the Middle English vowel class of the word, making the SED data more readily accessible for the purposes of this project. The data is drawn from the six traditional northern counties in addition to northern Lincolnshire and the Isle of Man. The former is not typically thought of as part of the north, but is included by Kolb for linguistic reasons. As Wells notes,

‘Northern’ in this sense might more precisely be glossed ‘midlands or northern’. We cross from the south to the linguistic north at the point where we pass the northern limits (in broad local accents) of the FOOT-STRUT Split and of BATH Broadening. (1982:II.349)

This linguistic line often does not correspond perfectly with the popular conception of the north, as revealed, for example, by the work of perceptual dialectologists (see e.g., Montgomery, 2007). The Isle of Man is not included in this analysis, due to its distinct linguistic history.

The SED used traditional data collection methods: Fieldworkers hand-recorded close transcriptions of informants’ responses to a standard questionnaire, which contained over 1000 questions and typically took at least four days to complete. The recording sites are all small rural towns, with the exception of York
Map 1. PANR base map.

and Leeds. Map 1 is a base map of the towns included in the atlas, with historical county borders for reference. SED numbers are given, and the online and SVG versions of these maps also display the town names.

Informants were chosen on the basis of how well they represented the local traditional dialect, and were typically male, aged 60 or above, and engaged in agricultural work (Orton, 1962). Each point on the map represents the responses of anywhere from 1-5 informants. In most cases only one variant per word was recorded in each location; however, in the few cases where different variants were present, the additional variants were usually included either as an additional point or diacritic on the map, or listed in a key if they were considered by Kolb (1966) to be unusual or otherwise unrepresentative. For the purposes of this project, only the variants recorded on each map itself are used, as they are assumed to be the most reliable data.

This paper is written very much in the spirit of the original work of Eduard Kolb (1966) upon which it is based, in that it recognizes that a carefully-selected subset of the available data can in itself be enlightening. Thus, of the considerable amount of data contained in the PANR, thirty-three of the ME long vowel words were transferred to digital format, corresponding to the modern PRICE, FLEECE, FACE, MOUTH, GOOSE, and GOAT classes (Wells, 1982). A limited amount of re-categorization was done, the goal being to reduce the number of keyed variants to a manageable amount by eliminating phonetic minutia, in the hopes of attaining a clearer identification of the broad phonological patterns. For example, diacritics indicating fine differences in vowel quality (e.g., \( \ddot{a}, \dddot{e} \)) were ignored, and variants with the same nucleus but superscripted on- or off-glides which differ slightly in quality are generally mapped together (e.g., \( \ddot{u}, \dddot{u} \)).

There have been several past efforts to map the SED vowel data in a more theoretically interesting way than previous lexical atlases, and each has wrestled with the problem of representing multidimensional data in traditional, often black-and-white, print media. Kolb
(1966) chose to create a separate map for each word, demonstrating in rich detail the extent of the phonetic variation in each, but making it difficult to arrive at phonological generalizations across an entire vowel class. Anderson (1987) took the opposite approach, mapping the geographic distribution of each reflex of a vowel class. This technique makes it easier to see the distribution of any one variant of interest but creates difficulty in identifying the geographic relationships between different variants. Other notable approaches include Ogura’s (1987) work on the role of lexical diffusion in the spread of the GVS, and Labov’s (1994:476ff) review thereof.

Thanks to the new possibilities provided by online publication, we no longer have to limit our examination of the SED data to one dimension at a time. The mapping method used here affords a more comprehensive view of the vowel system, as well as closer scrutiny of the patterns of diffusion, including previously obscured nesting effects of the variants. Each map displays all of the variants found in an entire vowel class; the points are color-coded according to which variant was the most frequent in the words sampled, and the tooltips (which are displayed when hovering the mouse over any point) contain a list of every other variant found at that location.4 In this way it is possible to achieve a wider view of the geographic relationships between the most common reflexes of each Middle English vowel class, without losing the rich phonetic detail and lexical differences in the process.

4. Results

The thirty-three digitized PANR words are presented here in a series of nine maps, reflecting Kolb’s ME long vowel categories: ī, ē, ē̄, ā, ō. The vowels ā and ō are both divided into two categories, indicated by subscripts, due to their being reflexes of different Old/Early Middle English etymological classes. We will first examine the relatively unproblematic front
vowels and consider the evidence they provide for a specific chronology of the GVS, before turning to the more controversial and irregular back vowel reflexes.

4.1. The Front Vowels

Map 2 shows the most common reflexes of ME ĭ in the following lexical items, which are part of the modern PRICE class: sky, Friday, time, knife, and writing. It should be noted that this vowel class is subject to the Scottish Vowel Length Rule (SVLR) in Northumberland and Tyneside, unlike other parts of England. Thus, in these locations all words in this set except for sky should show the short allophone predicted by the SVLR. Milroy (1996:214) reports that this short allophone is “regionally marked and has a high-mid nucleus near [ei] rather than the mid (central) nucleus usually described for the SVLR.”

This map shows a clear progression of the change from [iː] > [aɪ] from south to north; the expected reflex [ai] is found in the northeast, and further advanced forms [aɪ, aː; æ] are found in the south-central area. In Northumberland we find the expected SVLR variant, [ɛɪ], in all words except sky, which shows [aɪ]. Wells (1982:I.149–150) notes that back [aɪ] variants are characteristic of the urban south, while glide weakening or monophthongization is prevalent in Manchester and Leeds; it is indeed in the southwest Manchester area that we see the heaviest concentration of monophthongal forms in Map 2.

Based on this map, it appears safe to say that the diphthongization of ME ĭ has gone to completion in the north; there are no pre-shift or partially-shifted reflexes to be found in this data. The shift is furthest advanced in the southernmost counties, exactly what we would expect to see from a change which is generally thought to have originated in the south.5

Moving on to the next vowel involved in the front half of the GVS, Map 3 shows the most common reflexes of ME ɛ (modern FLEECE class) in the following lexical

Map 3. Reflexes of ME ɛ.
items: three, wheel, goose, and sheep. This is by far the most uniform map reproduced here, from which it is clear that the $\hat{e} > [i:]$ shift has gone to completion. The only part of the north which does not have [i:] as the most frequent reflex is the patch of [a] in the Yorkshire Dales. As the Dales are a mountainous and sparsely populated area, it is tempting to identify this as a relic area; however, [a] is not terribly likely to be an intermediate step between [e:] and [i:]. Fortunately, Lass (1976:90ff) demonstrates how such a reflex might be analyzed as innovative, taking Dentdale, Yorkshire as a case study. This particular form is indeed more plausible as a further development of the shift to [i:] since as Dobson (1968:659) observes, the shift from [i:] > [ii] > [ai] (which he proposes as the trajectory for ME $\hat{a}$) is attested in more modern Cockney and Australian varieties. Thus while there may be a handful of older forms still in use throughout the north (possible candidates include the [e] in three found in Du4 and Du5), this change has uniformly gone to completion.

Map 4 shows the outcomes of ME $\hat{e}$, in the words team and wheat. This vowel class derives from Old English (OE) $\hat{a}$, $\hat{e}$a and was raised first to [e:] via the GVS, and later to [i:] via the 18th century FLEECE merger. Wells (1982:1195) notes that this later change didn’t happen everywhere, and pre-merger pronunciations are found in much of the north. On the whole it would appear that the $[e:] > [e:] > [i:]$ shift did happen throughout the far north and southwest; the glaring exception is the wide band of locations where this vowel is realized as an ingliding [a]. Anderson (1987:65) says of this reflex, “It is reasonable to regard this type as representing the oldest vernacular type to have developed from ME /e/.” That being the case, it is difficult to explain why the far north does not share this older type, but rather has the fully-shifted pattern. Anderson’s (1987) suggestion that this vowel was “caught” between ME $\hat{a}$, which raised early, and ME $\hat{e}$, which raised slowly, is confusing at best, and does not explain how this pressure would have caused ME $\hat{e}$ to raise in the far north.
Map 5 meanwhile shows the outcomes of ME $\ddot{e}_2$, which is derived from OE/early ME $\ddot{e}$ after Open Syllable Lengthening (OSL). In most English dialects, this vowel followed the same path as $\ddot{e}_1$ and also finished as part of the modern FLEECE class. Indeed, we see rather similar outcomes between the two classes, save for the patch of $[\ddot{e}i]$ in the southwest found here. This reflex occurs in the part of England where ME $\ddot{e}_2$ did not merge with ME $\ddot{e}_1$, but rather remained distinct, developing an upgliding diphthong (Anderson, 1987:82). Thus Maps 4 and 5 are identical for the locales where these two $[e]$ classes merged but show a contrast for most of Lancashire and south Yorkshire, where they did not. This data comes from the words eat, steal, and speaks (from OE etan, stelan, specan, subject to OSL).

Map 6 shows the last vowel involved in the front half of the GVS, ME $\ddot{a}$. The lexical items this map shows are spade, grave, bacon, and gable. The history of this vowel class, the modern FACE class, is somewhat complex. The GVS resulted in $\ddot{a}$ raising to $[e]$, which Dobson (1968:594) says must have begun in the 15th century and progressed first through $[æː]$. Thus ME $\ddot{a}$ and $ai$ merged in $[eː]$ sometime in the 16th century and continued to raise together (Dobson, 1968:779). By the 17th century, $[e]$ had already begun to occur, and was standard by the 18th century. The modern $[eɪ]$ pronunciation dates from around 1800. So in sum, $[a] > [æː] > [eː] > [e] > [eɪ]$. There is also evidence of an $[eə]$ reflex in northern dialects, attested in Gil (1619), Smith (1568), and Cooper (1685) (quoted in Dobson, 1968:603).

This extended series of changes between ME and Modern English leaves more room for variation in the modern reflexes, and indeed a good deal more variation is seen here. Many points show some variant of the expected [e]-nucleus, but some also appear to be further advanced to an [i]-nucleus, while most do not appear to have undergone the later shift to [eɪ] at all. Wells (1982:II.357) states that the $[e]$ form only developed in the midlands and perhaps the urban middle north; if so, this must be a quite recent change, as no trace can be found in this data.
On the basis of this map, it is safe to say that the first “step” as it were in the GVS raising of ME ā did occur throughout the north—there are no locales which retain [aː] or [æː]. Furthermore, in describing the modern variation for this vowel, Beal (2010:19-20) identifies the ingliding diphthongs found throughout most of the north as “traditional” and observes that they are now largely lost in favor of the more “pan-Northern norm” of monophthongal [eː] or the diphthongal [ei] which is now spreading from Manchester and Liverpool. Indeed, Kerswill (2003) reports a dramatic decline in the use of the traditional [aː] in favor of [eː] among young speakers in Newcastle. This agrees with the pattern in Map 6, where we do indeed see [eː] in the southeast Manchester-Liverpool area, with the traditional ingliding diphthongs dominating the rest of the north.

Thus far, we have seen each piece of the front half of the GVS robustly represented in the north. The shift from ē to [iː] even showed a small nesting pattern, with the reflexes in the southeast displaying progressively advanced forms (later backing and/or monophthongization changes) nested inside the expected [aː] reflexes to the north and west. Next, the shift from ē to [iː] is uniformly represented across the entire north, with the exception of the innovative [aː] reflexes in the Yorkshire Dales. The shift from ē to [iː] in both the ē₁ and ē₂ classes proceeded as expected in the northernmost and southernmost counties, with the standard [iː] reflex beginning to encroach on the traditional [aː] form along the north Yorkshire border (see e.g., Du6, Y2). Finally, the shift from ā to [ɛː] certainly occurred throughout the north, but the region shows a much higher degree of variation in terms of what happened after that initial raising than it did for any of the other front vowel classes.

In terms of chronology, the patterns summarized here tend to support the view that the shift began with the upper half of the vowel system; that is, the raising of ē and diphthongizing of ā, since these two vowels show the most complete and consistent GVS reflexes. The lower
vowels ɛ and ą also show the expected GVS reflexes, but to a lesser extent as there are large areas of the North which still maintain older forms. Accepting then, for the moment, the Luick/Lass formulation of the GVS as being driven by mid-vowel raising, thus far we have seen no locales which might be described as having an irregular or incomplete form of the shift, leading to the conclusion that these changes are the product of transmission.

4.2. The Back Vowels

Map 7 shows the reflexes of the vowel at the heart of this century-long debate: ME ą, the modern MOUTH class. This map is based on the words how, house, clouds, about, and drought. There is a very clear distinction between the [uː]-retaining areas in red and the shifted areas, with a transition zone consisting of points with a small on-glide or a centralized nucleus. Similar to what we saw for ME ɪ, there is also a clear nesting pattern here, with the most advanced forms—[aə], [æʊ], etc.—clustered in the southwest and fanning out from there. Unlike ME ɪ, there is a large area which retains the older, unshifted ą, encompassing not only Northumberland, but also northern Cumberland and Westmorland, eastern Yorkshire, and part of northern Lincolnshire, roughly north of the Ribble-Humber line. Wells (1982:1.152) identifies the [æ ~ ɛ ~ e] nucleus variants with southern, innovative dialects, and the schwa variants as conservative and rural.

The outcomes of the next back vowel, ME ô, are shown in Map 8. The lexical items used are noon, boots, tooth, and moon, which belong to the modern GOOSE class. Wells (1982:1.185) says, “North of a line running from southern Cumbria to the Humber estuary, the present-day dialectal reflex of Middle English /ɔː/ is a front vowel, e.g. [gies] goose, while Middle English /uː/ remains monophthongal, e.g. [huːs] house.” The line Wells references is, of course, the Ribble-Humber line, an old and persistent bundle of isoglosses.
which has long distinguished the dialects to the north and south.

The chronology for this vowel was as follows, in the areas in which it fronted: \(\hat{o}\) fronted to \[\hat{o}:\], which was raised to \[\hat{y}:\], and later diphthongized to \[\hat{y}:\hat{u}\] etc. Wells (1982:1.186) identifies these fronted variants as now "sharply recessive," an indication of which may be seen in the two Northumberland locales displaying a geographically-isolated (presumably levelled) \[\hat{u}:\] reflex. Similarly, the \[\hat{u}:\] reflexes found in North Lincolnshire are likely late developments owing to diffusion from the south, as argued in Britton 2002, and not a GVS-induced merger between \(\hat{o}\) and \(\hat{u}\).

This is the second vowel which is crucial to Luick’s push-chain model, since it has been proposed that it was the fronting of \(\hat{o}\) prior to the GVS which was responsible for the lack of diphthongization of \(\hat{u}\) in parts of the north. The dependency relationship between these two vowels is examined more thoroughly in Section 4.3 below; for the moment, simply note that there are two different types of ME \(\hat{o}\) reflex being shown here: One which underwent fronting, and one which did not. Based on this map (Map 8), there appears to be a reasonably clean divide along the Ribble-Humber line (which agrees with Wells’s description) between the two types of reflex, with no obvious transition zone.

The next two maps, Maps 9 and 10, show the reflexes of ME \(\hat{a}\). The lexical items all belong to the modern GOAT class: clothes, both, oak (\(\hat{a}_1\)), and coal, foal, note (\(\hat{a}_2\)); however, the two groups have quite distinct histories in the north. The first set, \(\hat{a}_1\), belongs to the OE \(\hat{a}\) class, which was raised and rounded to \(\hat{o}\) in southern ME, but remained unchanged in northern ME, instead fronting and raising along with ME \(\hat{u}\) (Mossé, 1952:22). The second set, \(\hat{a}_2\), is derived from early ME \(\hat{u}\) with OSL and remained back, generally developing into an ingliding diphthong. In the south then, both these classes were raised to \[\hat{a}:\] by the GVS, after which \[\hat{u}:\] merged into the class, which was later raised and diphthongized to

Map 8. Reflexes of ME \(\hat{a}\).
the modern reflex, [oo]. In the north, only the \( \ddot{\alpha}_2 \) class followed this path, while the \( \ddot{\alpha}_1 \) class remained fronted and later raised along the front track to [e:]. Modern monophthongal variants reflect a lack of the Long Mid Diphthonging change, and are generally found in the north (Wells, 1982:146).

The modern reflexes in Map 9 clearly display the distinct history of this vowel class in the north. North of the Ribble-Humber line we see reflexes with front nuclei; to the south, the nuclei remain back and show GVS raising, precisely the distinct outcomes expected. The \( \ddot{\alpha}_2 \) class mapped in Map 10 shows a different pattern. The only fronted forms are found in Northumberland, in the Durham and Tyne and Wear area, and they are historically unrelated to the forms seen in Map 11. It is likely that these [\( \ddot{\alpha}: \sim \ddot{\alpha}_2 \)] reflexes represent a further innovation over the [\( \ddot{u} \)] reflexes found throughout most of the north, along the general lines suggested by Lass (1976:98ff.). Wales (2006:173) reports with some skepticism that Wakelin (1984) attributes this feature to “an attempt to ‘conform more closely’ to RP.”

Whereas the northern front vowel system showed the expected reflexes of the GVS, including varying degrees of advancement of the shift which formed nesting patterns, the back vowel system shows far less influence of the GVS. Outside of southeast Lancashire and Yorkshire, there are very few of the expected GVS reflexes to be found. The only vowel which shows a clear nesting pattern is ME \( \ddot{u} \), and even then, only within the southwestern area which had not undergone prior \( \ddot{u} \)-fronting; both \( \ddot{u} \) and \( \ddot{\alpha}_2 \) present much more complex pictures, owing to their distinct histories in the north. Overall, the picture seems to be one of regular transmission within the southeastern locales which followed the south in raising rather than fronting \( \ddot{u} \) and \( \ddot{\alpha}_1 \); however, the GVS back-vowel changes could not be transmitted in the areas north of the Ribble-Humber line which had prior fronting of these vowels.
4.3. Dependency Relationships

The last map presented here returns to the question of the dependency relationship between ò-raising and ú-diphthongization upon which Luick’s push-chain model relies. If his chronology is correct, and we assume that the GVS proceeded as a unitary shift in the north, then we might not expect to see any ú-diphthongization in the areas of the north in which ò-raising did not occur, due to prior ò-fronting. Given the possibility for diffusion, however, we expect that any locales which do show both ú-diphthongization and ò-fronting will be found in a transition zone where contact is likely between the dialects which fronted ò and those which didn’t.

As Map 11 shows, there is indeed a band running diagonally across the north where we see evidence of both ò-fronting and ú-diphthongization. Of this area, Anderson (1987:41) says,

It is hard to account for this intrusion [of ú-diphthongs] into an area which is generally highly conservative, and one can only assume that the change has been introduced from the Lancashire side or possibly from more southerly parts of the West Riding.

These apparently contradictory outcomes are precisely the sort of evidence that has been used by scholars in the past to argue against Luick’s chronology, or the very unity of the shift itself (specifically Western, 1912:3).

However, the complete shifts of ï and ë in the front vowel system across the north as a whole, and the complete shifts of the back vowels in southeastern Lancashire and Yorkshire tend to support the coherence of the chain shift. The fact that these contradictory points are located along the Ribble-Humber line, rather than say the far north, agrees with the diffusion hypothesis stated above and suggests that these diphthongs may not be a separate, spontaneous development. Therefore interpret this data not as a refutation of Luick’s theory, but as evidence that any standard GVS back vowel reflexes north of the
Ribble-Humber line are the product of diffusion rather than regular transmission, which would have proceeded as a unitary chain shift. This analysis leads to the following rough sketch of the history of the GVS in the north:

1. The GVS spread in a regular fashion across the north by means of transmission.
2. The $\delta$-raising change was not able to be transmitted north of the Ribble-Humber line, since $\delta$ in that region had undergone prior fronting.
3. Thus $\ddot{u}$ also did not diphthongize in those locations since the condition for diphthongization ($\delta$-raising) was absent.
4. Any apparently inconsistent modern data concerning these back vowels may be explained as the result of post-GVS diffusion, as I have demonstrated here for the cases where fronted $\delta$ and diphthongized $\ddot{u}$ coexist, and as Britton (2002) did for the areas in north Lincolnshire where both raised $\delta$ and undiphthongized $\ddot{u}$ can be found.

By viewing this data in the light of the transmission versus diffusion distinction, the apparent irregularities of the GVS in the back vowels of the north cease to be problematic for a unitary account of the GVS. Rather, these irregularities are seen as the natural result of the breakdown in structural relationships that occurred during the later diffusion of the back vowel changes from the south, across the dialect boundary running from the Ribble to the Humber.

In showing that these apparently contradictory reflexes are part of a larger pattern, and not truly random or problematic, this work demonstrates one way in which a new perspective may fruitfully contribute to a widely discussed question. While there are certain regular outcomes which are strong enough to appear even in a small subset of data, this analysis was naturally limited by the small amount of data used. Given only 2–5 words per vowel class, it was impossible to reliably observe the contextual effects which are known to affect the GVS, for example (e.g., those noted in

---

**Map 11.** Back vowel relationships.
Dobson, 1968). A useful extension of this project would be to similarly map all of the SED data, and perhaps some of the more reliable orthoepists’ notes, in order to arrive at a more nuanced picture of the patterns of change present. It was also not feasible to conduct an exhaustive study of the available historical evidence here; some of this history is examined elsewhere (Prichard, to appear), but much more could be done to trace the history of some of the more unusual modern reflexes noted here.

5. Conclusion

Based on the findings above, this study concludes that the problematic nature of the ME long vowel reflexes in the north is simply the result of the diffusion of fully-shifted forms into an area which had a rather different initial vowel system than was found in the south. I therefore agree with Smith’s (1996:99) assessment that although the processes involved in the Northern Shift are useful for our understanding of the process of triggering of the wider Shift, it is suggested here that the problem presented by these dialects is distinct from that manifested by the more southerly varieties.

But disagree with his conclusion that these problems are best accounted for by postulating a separate northern shift. This analysis accounts for the fact that regular outcomes and nesting patterns are found for the reflexes of the front vowels, while irregular outcomes and diffusion patterns are found for the back vowels: The prior fronting of ə in the north prevented the GVS from occurring in the back vowels, even as it proceeded regularly in the front vowels. The back vowel outcomes which appear to be the result of the GVS are the result of diffusion from areas which did undergo the full shift.

Acknowledgements

Thanks to Bill Labov, Don Ringe, Gillian Sankoff, Kyle Gorman, and audiences at NEW5 and NWAV41 for their helpful comments; to Aaron Ecay and Chris Montgomery for their contributions to the maps presented here; and to an anonymous reviewer especially for the time and effort they devoted to furthering my education in the historical intricacies of the Northern English dialects. This work is based on data provided through EDINA UKBORDERS with the support of the ESRC and JISC and uses Kain and Oliver historic boundary material which is copyright of the AHDS History [University of Essex], Humphrey Southall, Nick Burton and the University of Portsmouth.

Notes

1 ‘Irregular’, in the sense of unexpected, not the result of regular application of the GVS changes.

2 Further discussion of these arguments and a more detailed history of Northern English vowels may be found in Prichard (to appear).

3 Obsolete IPA characters have been replaced with their modern equivalents here.

4 To use the tooltip function, you must view the maps in a web browser, either by downloading the accompanying svg files, or by viewing the html version of this article.

5 I know of two possible exceptions to this consensus: Johnston (1992), who proposes a northern/midlands origin (but argues that the GVS is really two unrelated shifts), and Smith (2007), who similarly proposes that the northern vowel changes are completely separate from those that occurred in the south.

6 As a reviewer points out, Lass (1976:89) suggests that even these diphthongal reflexes of ME ə may be ‘late importations’. However, as he does not specify how late, it is difficult to know if these are best seen as recently-levelled varieties, or areas in which the shift was very slowly transmitted.

7 The area of diphthongization shown in Map 11 is not identical to that seen in Map 7. Since it is argued here that the diphthong is diffusing from the south, and this may of course proceed by means of lexical diffusion, Map 11 assumes that locations which reported a diphthong in any ME ə word, even if a majority of the data did not have a diphthong, provide evidence for the diffusion of the standard form.

8 That said, in Map 11 I have differentiated “full” diphthongs (the [au] type) from “mid” diphthongs (the [æ] type) as it is not clear that these are both the product of diffusion. The history of the mid diphthongs will need to be investigated elsewhere.

References


Milroy, James. 1996. Variation in /ai/ in Northern British English, with comments on Canadian Raising.