

1 Analyzing dialect variation in historical speech corpora^a

2
3 Margaret E. L. Renwick^b

4 Linguistics Program, Gilbert Hall, University of Georgia, Athens Georgia 30602, USA

5
6 Rachel M. Olsen

7 Linguistics Program, Gilbert Hall, University of Georgia, Athens Georgia 30602, USA

8
9
10 RUNNING TITLE:

11 Analyzing dialect variation in historical speech corpora

12
13
14 DATE OF RESUBMISSION: December 21, 2016

15
16
17
18
19

^a Earlier versions of this work were presented in “Voices of Coastal Georgia,” an oral presentation at the 170th Meeting of the Acoustical Society of America, Special Session: Advancing methods for dialect variation, Jacksonville, FL, November 5, 2015; and in “Intradialectal Phonetic Variation in Southeast Georgia: Evidence from the *Linguistic Atlas of the Gulf States*,” an oral presentation at the 83rd Southeastern Conference on Linguistics, New Orleans, LA, March 29, 2016.

^b Address email correspondence to: mrenwick@uga.edu

ABSTRACT:

The *Linguistic Atlas of the Gulf States* is an extensive audio corpus of sociolinguistic interviews with 1121 speakers from eight southeastern U.S. states. Complete interviews have never been fully transcribed, leaving a wealth of phonetic information unexplored. This paper details methods for large-scale acoustic analysis of this historical speech corpus, providing a fuller picture of Southern speech than offered by previous impressionistic analyses. Interviews from 10 speakers (~36 hours) in southeast Georgia were transcribed and analyzed for dialectal features associated with the Southern Vowel Shift and African American Vowel Shift, also considering the effects of age, gender, and race. Multiple tokens of common words were annotated ($N = 6085$), and formant values of their stressed vowels were extracted. The effects of shifting on relative vowel placement were evaluated via Pillai scores, and vowel dynamics were estimated via Functional Data Analysis and modeled with linear mixed-effects regression. Results indicate that European American speakers show features of the Southern Vowel Shift, though certain speakers shift in more ways than others, and African American speakers' productions are consistent with the African American Vowel Shift. Wide variation is apparent, even within this small geographic region, contributing evidence of the complexity of Southern speech.

Keywords: Speech; Vowels; Dialect; Databases; Data analysis

I. Introduction

Audio corpora are valuable resources for exploring dialectal speech patterns. Historical corpora offer otherwise unavailable insight into speech communities, as well as how synchronic speech patterns may have developed over time. The *Linguistic Atlas of the Gulf States (LAGS)* is an extensive historical, sociolinguistic audio corpus containing approximately 5300 hours (Montgomery and Nunnally, 1998) of conversational interviews with 1121 speakers in eight U.S. Gulf states, recorded from 1968 – 1983. It is maintained and publicly available via the *Linguistic Atlas Project (LAP)* at the University of Georgia (Kretzschmar, 2011).

LAGS' original purpose was to survey lexical diversity by eliciting specific target lexical items, which were analyzed by *LAGS* fieldworkers and trained scribes. This corpus is valuable to modern sociophonetics because it includes detailed demographic information for each speaker. While the accompanying *Protocols* (Pederson, 1981) include handwritten close phonetic transcriptions of target lexical items, these items were not time-aligned with *LAGS* audio files, making it impossible to easily navigate the audio. Furthermore, within these conversational interviews, *only* the target lexical items were transcribed; thus, all non-target words uttered by the speakers – totaling thousands of hours of speech – remain to be transcribed and analyzed.

Our study takes the first step in this analysis. It uses full interviews recorded between 1969 – 1979 and originally transcribed by Renwick and Olsen (2016), from one *LAGS* speaker area (labeled *AK*) spanning five contiguous counties (Brantley, Camden, Charlton, Glynn, and Ware) in southeastern Georgia, and applies modern methods of acoustic analysis to them. This approach provides a method for bringing dialect analysis more closely into line with current trends in corpus linguistics. In traditional analysis of sociolinguistic corpora, individual lexical items are studied, while the remainder of the corpus is neglected. However in other realms of

phonetics and phonology, large amounts of increasingly spontaneous speech are annotated in their entirety (e.g. the Audio British National Corpus, Coleman et al., 2012) and made freely available to researchers. Full annotation of sociolinguistic corpora permits the intersection of modern acoustic analysis with demographic information about the speakers, while incorporating other factors like lexical frequency and prosody, to study speech in a fuller sociolinguistic context.

Complete transcriptions of LAGS recordings provide a detailed dataset to compare to traditional dialectal depictions of speech in the South. While present-day studies of variation, including Southern speech, often employ acoustic analysis (Clopper et al., 2005; Feagin, 2003, *inter alia*), such measures have rarely been used to analyze *LAGS* or other historical corpora within the *LAP*. Important exceptions are Thomas (2001), who presented aggregate vowel spaces based on *LAGS* data, and McCarthy's (2010) work on the Northern Cities Shift in the *Dictionary of American Regional English (DARE)*. This project builds on their work by assessing multiple repetitions of words within and across speakers to seek evidence of the Southern Vowel Shift (SVS) (Labov et al., 2006) and the African American Vowel Shift (AAVS) (Thomas, 2007) in historical speech. Tools for vowel analysis include Pillai scores (Hall-Lew, 2010) and Functional Data Analysis (FDA) (Koenig et al., 2008; Risdal and Kohn, 2014).

The current work is presented in the context of a larger project whose goal is to make up to 3400 hours of *LAGS* recordings accessible to researchers. This will permit a more precise, fine-grained view of dialect variation in the United States, and will facilitate the mapping and modeling of variation and sound change by incorporating greater amounts of data from each individual speaker.

II. Background

A. Vowel quality in Southern speech

Vowel quality in Southern speech has been widely studied and described. One of its best-known characteristics is the rearrangement of certain vowels in the vowel space, known as the Southern Vowel Shift. Labov, Yaeger, and Steiner (1972) describe many of its features based on interviews with speakers from the North Carolina Outer Banks, East Atlanta, and Central Texas; these were revised and expanded in later work to codify the SVS in its present form (Labov, 1991; Labov et al., 2006; cf. Thomas, 2003). Monophthongization, or glide weakening, of /aɪ/ is argued to trigger the shift, followed by lowering and backing of tense /i eɪ/, and raising and fronting of lax /ɪ ɛ æ/. These changes may be phonologically conditioned: in the Southern region explored here, /aɪ/ is expected to weaken only word-finally and before a voiced consonant (Labov et al., 2006). Along with these front vowel movements, fronting of /oo u/ is additionally a characteristic of Southern speech (Clopper et al., 2005; Labov et al., 1972), while Thomas (2003) describes upgliding of /ɔ/ towards [ɔʊ], weakening of the glide in /aʊ/, as well as /ɔɪ/ monophthongization particularly before /l/. These qualities support stereotypes of the “Southern drawl” (Dorrill, 2003), explored acoustically by Feagin (2003).

Age is relevant to the SVS, because it may determine whether a speaker participates in the shift. Weakening of the glide in /aɪ/ (PRIZE) among European American (EA) speakers was widespread by the 1930s (Thomas, 2003), having begun in the late 19th Century; however, upgliding of /ɔ/ in THOUGHT began as much as a century earlier, while fronting of /oo/ in GOAT was not common until after World War II (Thomas, 2005). Labov et al.’s (1972) description featured interviews with speakers spanning generations, often from a single family. They note that fronting of /u/ (GOOSE) and /oo/ occurs earlier than front vowel shifts, in

particular lowering of /eɪ/. Whereas fronting of /u/ and even /oo/ was found in older speakers, /eɪ/ lowering was only found in young speakers (born in the 1940s), but never in older speakers (born as early as the 1800s). Across generations, /u/-fronting eclipsed /oo/-fronting, a finding echoed both in earlier work describing fronting of /u/ but not /oo/ (Kurath and McDavid, 1961), and in later work arguing that /oo/-fronting is a relatively new feature of Southern speech common among younger European Americans throughout the South (Thomas, 2001). An approach comparing the speech of older speakers to that of younger speakers, to trace potential language change, is an Apparent Time analysis; this is a widely applied technique in sociolinguistic studies (Labov et al., 2013). We examine speakers born from 1894 – 1954, with certain Apparent Time comparisons.

A further important consideration in exploring Southern speech is ethnicity. African Americans (AA) are not expected to participate in the SVS, but rather may participate in the African American Vowel Shift (AAVS). The AAVS shares characteristics with the SVS, but also has important differences (Thomas, 2007). Generally, it involves fronting of the vowel /ɑ/ in LOT, and raising and fronting of /æ ɛ ɪ/ (TRAP, DRESS, KIT). Regarding the relationship between the AAVS and SVS, both patterns can include /u/ and /oo/ fronting; /u/-fronting is more likely than /oo/-fronting for both EA and AA speakers, but synchronically AAs are 10% less likely to front (Thomas, 2007). The /i ɪ/ switch is rare in African American English (AAE), while the /eɪ ɛ/ (FACE/DRESS) switch is somewhat more common, partially via lowering of /eɪ/, albeit not to the extent seen in the SVS. Glide weakening of /aɪ/ is documented in AA speech, finally and before voiced consonants, particularly liquids, while pre-voiceless weakening is less common than among EA speakers. The vowel /ɔɪ/ (CHOICE) may lower in AAE, but its glide is

not necessarily weakened. Southern AAs may also exhibit a less diphthongal /aʊ/ (PLOW) than do EAs, although the same vowel is also more fronted among EA speakers.

Generational differences may occur among AA speakers. As shown by Thomas (2007), AAs in the Upper South born in the mid- to late-nineteenth century, as was a speaker in the present study, had more frequent monophthongal variants of FACE, GOAT, and THOUGHT vowels than EAs from the same geographic area. However, monophthongal variants of FACE and GOAT vowels are not evident in the speech of younger African Americans born after World War I, as was another AA speaker examined here.

A variety of studies have shown that the SVS does not occur uniformly across all Southern speakers, nor are Southern features limited to those discussed above. Studies of specific areas including Memphis, Alabama, and Charleston have produced conflicting findings on the presence of the SVS within those regions (Feagin, 2003). Furthermore, speakers within Georgia do not adhere identically to dialect features: recent sociolinguistic and survey results from the Atlanta area indicate varying rates of monophthongization and realization of SVS features, across racial and socioeconomic lines (Kretzschmar, 2015; Prichard, 2010). Findings of variation in shift participation even among siblings have made the strong point that intradialectal differences are expected (Fridland and Kendall, 2012; Kendall and Fridland, 2012), and some researchers have grouped speakers according to their degree of participation in the SVS or AAVS (Koops, 2014; Risdal and Kohn, 2014). It is anticipated that speakers' adherence to the SVS may also vary in Southeast Georgia, the source of data in our study.

B. Methods for analyzing Southern speech

1. Impressionistic analyses of Southern speech

Studies of Southern speech preceding and concurrent with *LAGS*' collection were largely impressionistic, or involved a limited number and network of speakers. While impressionistic analysis offers useful information about salient speech patterns, it may fail to capture individual and intradialectal variation. *LAGS* itself includes *Protocols* which are detailed phonetic transcriptions made without the benefit of acoustic analysis tools. This follows in the tradition of other dialectal surveys such as the *Linguistic Atlas of the Middle and South Atlantic States* (*LAMSAS*), which included speakers from the region studied here (specifically Glynn, Camden, and Charlton counties), interviewed in 1947 (Kretzschmar et al., 1994); unfortunately those interviews were not recorded. Similar impressionistic protocols from the *Linguistic Atlas of the North Central States* (*LANCS*) were used in an Apparent Time study comparing Central Ohio speakers born in the 19th Century to subsequent generations (Durian, 2012).

The *Atlas of North American English* (*ANAE*), compiled much later than *LAGS* in the 1990s, similarly relies principally on impressionistic transcriptions of surveys given to speakers over the telephone (Labov et al., 2006). While select speakers were recorded for acoustic analysis, few speakers were analyzed per geographic area, and individual variation is not discussed. The region of interest to the current study, in southeast Georgia, is not represented in the *ANAE*; the nearest acoustically analyzed speakers were two women of similar age and ethnicity to one another in Jacksonville, FL, over 40 miles away (Labov et al., 2006). Both speakers monophthongized /aɪ/ word-finally and before voiced codas, though the degree of glide weakening is not captured.

2. Acoustic studies of linguistic atlas data

Thomas (2001) conducted acoustic measurement of vowels in *LAGS*, as well as other archival audio corpora including *DARE*. In his work, formant measurements were taken at monophthong

midpoints, and at two points for diphthongs. Vowel spaces were compiled using means of 7-10 productions per vowel, per speaker (Thomas, 2001, p. 10). Thomas' closest speaker to *LAGS* area *AK* was a EA female born in 1890 from Moultrie, Georgia (2001, p. 113), a resident of the Piney Woods region, which is linguistically distinct from the Atlantic Coastal Plain represented in area *AK*. She demonstrated SVS features such as weakened /aɪ/ before both voiced and voiceless consonants (not predicted in area *AK* based on Labov's description), and /æ/-raising. Thomas' other Georgia speaker was an African American male born in 1844 from Skidaway Island, who spoke Gullah. This speaker showed evidence of AAE in his monophthongization of pre-voiced /aɪ/.

Linguistic atlas data from outside the South has also been analyzed acoustically. McCarthy (2010) explored the ordering of the Northern Cities Vowel Shift in Chicago, using archival audio collected in *DARE*. Thomas (2010) examined longitudinal language change by comparing *DARE* speakers born from 1880-1908, to speakers born in 1970-1994 to establish whether the boundaries between the North and Midland dialects in Ohio had evolved over time. Acoustic data from *DARE* and, to a small extent, *LANCS* have also shed light on language change in central Ohio (Durian, 2012).

3. *Analysis of dialectal vowel quality*

There are common practices for measuring, representing, and analyzing variation in dialectal vowel quality in terms of their position in acoustic space, the distances between vowels, and their dynamic trajectories. A coarse-grained, but effective, technique to provide vowel coordinates is to take F1 and F2 formant measurements at the midpoint of monophthongal vowels, ideally across multiple tokens, and present the results as means per dialect area (Clopper et al., 2005) or per speaker (Thomas, 2001). This degree of summary obscures the potential for both

intradialectal and intraspeaker variation in vowel production, leading us to avoid this level of data summary. Euclidean distances between vowels are employed by some, particularly to ascertain whether a vowel has shifted (Fridland et al., 2014; Fridland and Kendall, 2012; Kendall and Fridland, 2012). Distances are typically calculated between vowel centroids, but this technique cannot quantify the degree of overlap between clouds of data points that are relevant for determining e.g. the presence of a vowel merger (Hall-Lew, 2010). Instead we calculate Pillai scores, as this statistic takes variation into account to evaluate the relationship between specific pairs of vowels on an individual speaker level. Scores result from a comparison of two vowels' distributions, based on a MANOVA that may incorporate phonological variables, and reflects the degree of distinction (or merger) between them (Amengual and Chamorro, 2015; Hall-Lew, 2010; Hay et al., 2006).

Dynamics within the vowel are also relevant to dialect distinctions. Diphthongization has been quantified using Euclidean distance *within* each vowel token: the more a vowel changes over its time course, the more diphthongal it is (Fabricius, 2007; Haddican et al., 2013; Maclagan and Hay, 2007; Reed, 2014; Schützler, 2015). As illustrated by Fox and Jacewicz (2009), Euclidean distance does not capture the full scope of within-vowel formant movement, which varies across dialects. Euclidean distances are easily interpretable and have the advantage of collapsing both F1 and F2 movements into a single value, but they remain unaffected by any curves in formant trajectories. Trajectory Length (Fox and Jacewicz, 2009) does take formant movement into account, as it sums distances between multiple points in the vowel trajectory. A related measure, spectral rate of change (Fox and Jacewicz, 2009), incorporates vowel duration to pinpoint dynamic sources of variation.

Even when combined, trajectory length and duration cannot fully capture the shape of formants' paths. Vowels' spectral and temporal dynamics do differ both across and within dialects, characteristics that recent work has captured via Functional Data Analysis (FDA) of variation in /æ/ among speakers from Houston (Koops, 2014), and EA vs. AA speakers in North Carolina (Risdal and Kohn, 2014). To holistically decompose and describe the height, shape, and direction of formant trajectories, we employ FDA to generate a polynomial whose coefficients take curvature into account.

Duration is also relevant to dialectal vowel variation: speakers from Tennessee were found to produce tense/lax vowel pairs with a *reduced* durational difference precisely when they overlapped *more* in spectral quality (Fridland et al., 2013). We do not conduct a systematic durational comparison, because prosodic context is uncontrolled, and it is likely that durational differences due to prosody would outweigh intra-variety differences.

C. The Current Study

Quantifying vowel productions in a historical speech corpus like *LAGS*, using methods that incorporate the potential for variation, contributes to knowledge of Southern speech patterns and can be applied to other dialects and datasets. With these methods, vowel shifting in southeast Georgia is investigated, including the following research questions: Do front lax vowels undergo the raising and overlap patterns expected in both the SVS and AAVS? If so, increased overlap is predicted between /i ɪ/; /eɪ ε/; and also /ε æ/, due to /æ/-raising; however, greater overlap among /i ɪ/, /eɪ ε/ is predicted for EA speakers, since their tense vowels may lower and centralize while lax vowels raise and front. Second, are back vowels /oo u/ fronted? If so, greater fronting is expected for EA than AA speakers (Thomas, 2007). Third, is there evidence of glide weakening for /aɪ ɔɪ aʊ/, and is it phonologically conditioned as described by Labov et al. (2006)?

Throughout, sociolinguistic variables including race, age, and gender are taken into account. With regard to age, based on previous findings we predict that because the EA speakers interviewed in area *AK* were all born before 1929, they might not exhibit strong /oo/ fronting, and their degree of /u/-fronting and /aI/-weakening may vary. The methods follow in III, the results are presented in section IV, and section V describes their implications for models of Southern speech.

III. Methods

A. Data selection and preparation

This initial study of *LAGS*' acoustics focuses on one speaker area within Georgia, area *AK*, defined by *LAGS* as a unit on the basis of geographical, historical, and social factors (Pederson et al., 1986). This area lies in far southeastern Georgia, just north of the Florida border and Jacksonville, and includes the coastal city of St. Marys, the inland city of Waycross, and several barrier islands on which Gullah Geechee is spoken.¹ Area *AK* falls near the edge, but firmly within, the South according to Labov, Ash, and Boberg in the *Atlas of North American English* (*ANAE*) (2006). It may be within a distinctive subarea of the Lower South described by Kurath and McDavid (1961), encompassing the coast of Georgia and Florida that adjoins the South Carolina Low Country. Interviews were recorded between 1969 – 1979 in the counties of Camden, Charlton, Glynn, Ware, and Brantley (although no interviews were conducted in the latter). Audio totaling approximately 36 hours for ten speakers (of twelve interviewed) is preserved, though the poor audio quality of one speaker's interview precludes his inclusion in the results presented here. These speakers (4 male, 5 female; mean age 63.7 years) vary widely in age, gender, race, occupation, socioeconomic status, education level, and worldview (determined in conversation by *LAGS* fieldworkers). Table I provides information about each speaker, as well

as the number of hours of audio available, and the number of words transcribed in their interview (see III.B, below). Demographic factors and speaker numbers were provided by *LAGS* and coded following *LAGS* conventions: age (years); gender (male = M, female = F); socioeconomic status (SES: lower class = L, middle class = M, upper class = U); race (African American = AA, European American = EA).

TABLE I. Speakers, ordered by Race, Generation (oldest – youngest), then Gender²

Speaker	County	Generation	Birth Year	Recorded	Age	Race	Gender	SES	Audio (Hr:Min)	Word Count
195	Camden	1	1894	1974	80	EA	M	M	9:00	74,083
199	Ware	1	1899	1977	78	EA	M	L	4:20	15,466
198	Charlton	1	1896	1972	76	EA	M	M	2:10	7,787
197A	Charlton	1	1903	1974	71	EA	M	L	2:50	4,812
197	Charlton	1	1897	1969	72	EA	F	M	1:10	4,740
201A	Glynn	2	1916	1974	58	EA	F	M	3:10	4,673
202	Glynn	2	1919	1974	55	EA	F	U	3:15	5,360
200	Glynn	1	1900	1974	74	AA	F	L	4:00	10,532
201	Glynn	3	1954	1977	23	AA	F	M	3:15	2,850

Biographies of *LAGS* speakers include several major factors across which dialect differences emerge: race, geography, gender, and age (Thomas, 2001). Our analyses reflect these divisions, focusing on the effects of race and age on speech patterns. Accordingly, speakers were grouped into three generations: 1 (born 1894-1903), 2 (born 1916-1929), and 3 (born 1954).³ The sample from *LAGS* area AK is highly, but unavoidably, unbalanced with respect to all demographic variables: for example, it includes no black males, no upper-class males, and the distribution of gender is not equal across races, classes, or generations. As more speakers' interviews are transcribed, geographical coverage will expand and a more balanced sample will emerge, allowing the investigation of differences owing to, for example, varying overt and covert prestige norms across social classes within each race and gender.

The interviews were originally recorded on reel-to-reel and audio cassette tapes (Pederson et al., 1986), which were not intended for archival audio: most tapes were retained for reference until phonetic transcription was completed, and then recorded over with new interviews, meaning we are fortunate to retain these. Fieldworkers did not necessarily aim to obtain high-quality recordings, and some tapes degraded with time. The recordings were digitized to .wav format by the Linguistic Atlas Project (Kretzschmar, 2011) at a sampling rate of 48kHz. The original audio files are lengthy (~30-40 minutes each), so to permit analysis with Praat (Boersma and Weenink, 2015), they were cut into pieces of 3-4 minutes, on which transcription was carried out. These were also treated in Praat with a low-pass Hann band filter (with 100 Hz smoothing) above 15000 Hz, to eliminate a high-frequency noise.

B. Transcribing the data

Manual transcription of the interviews was necessary because current tools for automatic speech recognition are not trained for these data. Likewise, forced alignment techniques common in phonetic analysis are only usable with a full orthographic transcription, which was unavailable. Transcription was carried out in SPPAS (Bigi and Hirst, 2012), using its tools for silence detection, TextGrid creation and manual orthographic transcription (IPU Transcribe). Full interviews were orthographically transcribed, including only speech of the interviewee, not the interviewer. Transcribed TextGrid intervals correspond approximately to individual utterances, separated by silence and typically less than 10 seconds in length. The TextGrids containing these full transcriptions were used to create, via a Perl script, a searchable index of over 132,300 tokens, which linked word labels to utterance-level audio time stamps (multiple words share the same time information). This index was used to identify commonly spoken words within and across speakers: each word token occupies one row, listing a unique identifying code, the

speaker's ID, the total number of tokens of that word produced by the speaker, and the time interval of its location in a particular audio file and TextGrid. Importantly, this marked the first time that *LAGS* interviews have been fully transcribed and systematically time-aligned to an audio signal.

C. Words and vowels collected

While previous analyses of *LAGS* have focused only on target lexical items transcribed in the *Protocols*, this project acoustically analyzes multiple repetitions of words within and across speakers. Two immediate benefits of this approach to analyzing atlas data are that by collecting the same words across speakers, inter-speaker comparability is increased; and by collecting multiple tokens per speaker, stronger estimates of central tendencies are available, while intraspeaker variation can be estimated more accurately. Tokens were gathered for the set of American English vowels listed in Table II. From the index we selected words produced by as many speakers as possible, preferably with multiple repetitions, which also contained a stressed target vowel that was unlikely to be reduced. For example, all participants uttered *house*, *fence*, *wood*, and *call* at least once, so they were analyzed for all speakers. Because natural language is unbalanced, some words were produced by only some speakers (e.g. *sun* was uttered by three speakers); the remaining tokens of /ʌ/ were gathered via other words containing *sun* (*Sunday*, *sundown*, *sunrise*) or via tokens of *stuff* (also *stuffed*, *stuffing*). We labeled at least five tokens of each vowel, per speaker; loquacious speakers like 195 provided more data points (due to greater repetitions of words) than less talkative interviewees like 201; we analyzed 24 and 5 tokens of *house*, respectively, from those two speakers. Words were identified and segmented manually in Praat. Not all tokens of each target word were usable: tokens were not segmented if close inspection revealed them to be unsuitable for acoustic analysis, due to e.g. noise in the recording,

reduction, or talker overlap. 6085 tokens of 239 unique words were labeled for further analysis; Table II shows the number of tokens gathered per vowel, with example words.

Each target word was annotated for the following phonological characteristics: canonical American English vowel quality, syllabic structure (open/closed), and manner and voicing of preceding and following segments (when present within the word). Some common words differed in minor ways, but shared a lexical root; thus all were also annotated by lexical root word for ease of analysis (i.e. *high*, *higher*, and *highly* were grouped as *high*). The data set contains 79 unique root words.

TABLE II. Number of tokens analyzed for each vowel or diphthong, with example words.

IPA	N	Words
i	412	<i>people, cheese, three</i>
ɪ	311	<i>six, fifth, kitchen</i>
eɪ	269	<i>eight, paper, place</i>
ɛ	409	<i>yellow, red, guess</i>
æ	662	<i>black, have, dad</i>
ɑ	593	<i>father, pot, closet</i>
ʌ	147	<i>sun, stuff</i>
ɔ	839	<i>dog, saw, taught</i>
ʊ	327	<i>wood, put, bull</i>
u	146	<i>room, school, due</i>
aɪ	738	<i>white, five, high</i>
aʊ	193	<i>cow(s), house</i>
ɔɪ	367	<i>oysters, boy, poison</i>
oʊ	672	<i>oak, coat, road</i>

D. Acoustic analysis and normalization

Acoustic characteristics of all target vowels were automatically extracted in Praat. The formant detection algorithm (Burg method) was manually optimized for each speaker; for females, typically 5 formants were extracted with a ceiling of 5500 Hz, and for men, 5 formants were extracted with a ceiling of 4500 Hz. For two speakers (1 male), 4 formants and a ceiling of 5000 Hz were used; and for one female speaker, 5 formants and a ceiling of 5000 Hz fit the data best.

Formant values were extracted in two batches. First, F1 and F2 values at the vowel midpoint were extracted, hand-checked, and corrected by visual inspection of the spectrogram. We also extracted F1, F2, and F3 values at 10% intervals of each target vowel's duration. It is not currently feasible to hand-check this larger data set. While the time-interval data allow evaluation of spectral change over the course of the vowel, the midpoint data are traditional in dialectal analyses and are immediately comparable to the predictions of the SVS and AAVS models.

Formant values, both those measured at the vowel midpoint and those measured in deciles, were separately normalized to z-scores (Lobanov, 1971), in order to render data comparable across speakers and genders while retaining characteristics that are vowel-specific, regionally-driven or sociophonetic in nature (Adank et al., 2004). This normalization method is speaker-intrinsic, vowel-extrinsic and formant-intrinsic, resulting in output centered on a mean of (0, 0).

E. Data analysis

Formant data were analyzed using two methods. Vowel midpoint data are plotted for visualization, and analyzed using Pillai scores, to reveal overlap in F1/F2 space among front vowels, and fronting of back vowels. Spectral trajectories were studied with Functional Data Analysis. Both analyses include statistical modeling to test for effects of sociolinguistic factors on vowels' acoustics.

1. Pillai scores

Effects of the SVS and AAVS on the acoustics of front vowels and certain back vowels can be quantified via the acoustic distance, or overlap, among certain vowel pairs. Pillai scores were calculated to test for and compare overlap across speakers, using the following vowel pairs: /i ɪ/,

/eɪ ε/, /ɛ æ/, /æ α/, /u i/, /oo i/. The Pillai score is a test statistic with a value between 0-1, accompanied by a p -value, based on a multivariate analysis of variance (MANOVA), which is appropriate for unbalanced data sets from corpora. When the formant values of two vowels are modeled by the MANOVA, a low Pillai score indicates high overlap between them (in some cases, a merger), while a high Pillai score indicates distinct distributions (Hall-Lew, 2010; Hay et al., 2006). The score itself is interpreted as a relative measure of the distance between two vowel clouds. A small p -value ($p < 0.05$) indicates a statistical distinction between the vowels, e.g. that the factor of vowel type is a significant predictor of formant values in the MANOVA. Pillai scores were calculated, based on normalized data, from a MANOVA that also included the manner and voicing of the prevocalic and postvocalic consonants (if any), as well as the syllable structure (open vs. closed) surrounding the target vowel. These factors were taken into account because they, particularly postvocalic context and syllable structure, are known to affect vowels' acoustics.

2. *Functional Data Analysis*

The phonetic realization of vowels over time was evaluated by creating a model of their formant trajectories from start to finish, using Functional Data Analysis (FDA) (Ramsay et al., 2009). Besides its applications to Southern vowel distinctions (Koops, 2014; Risdal and Kohn, 2014), the technique has had other recent applications in phonetics, to model differences in fundamental frequency contour shape (Grabe et al., 2003, 2007; Kochanski et al., 2005), variability in fricative production (Koenig et al., 2008), vocal effort (Mooshammer, 2010), and acoustic intensity (Renwick et al., 2014). FDA was carried out on both F1 and F2 measurements. Across all ten Lobanov-normalized F1 or F2 measurements for each vowel token, an orthogonal cubic polynomial function $i = at^3 + bt^2 + ct + d$ was fitted, using the *stats* package within R (R Core

Team, 2000). For each of the terms a coefficient is produced, indicating a characteristic of the formant's curve: **a**, the cubic term, shows how "S-shaped" the curve is; **b**, the quadratic coefficient, indicates the curve's degree of parabolic shape, and whether it is convex or concave; **c** is the linear term, or overall slope of the curve; and **d** is the curve's intercept, showing its mean value. Each vowel's trajectory is reduced to these four coefficients, which are subjected to statistical modeling: factors including sociolinguistic and acoustic characteristics (e.g. duration) are used to predict coefficient values, to show how different aspects of vowels' dynamics are affected by each factor.

IV. Results

The Results highlight two sets of vowels: first, the Pillai scores of pairs implicated in the SVS and AAVS are compared using midpoint formant values. Second, FDA is used to model the dynamics of vowels' formant trajectories. The latter includes SVS and AAVS vowels, as well as diphthongs that may undergo glide weakening in Southern varieties. The analyses take into account sociolinguistic factors of race, age, and gender, which are shown to have significant effects on Pillai scores and the output of FDA.

A. Vowel spaces in *LAGS* area AK

Setting the stage for quantitative comparisons within the vowel space, Figs. 1 and 2 depict vowel spaces of four *LAGS* speakers, generated using one F1/F2 point per token, measured at the temporal midpoint (cf. Fig. 7 of Renwick and Olsen, 2016). Figure 1 shows the oldest EA male, speaker 195, and the youngest EA female, speaker 202. Impressionistically, speaker 195 has the "least Southern" speech within this sample; he shows substantial overlap between /i ɪ/, but otherwise his front vowels /eɪ ε/, /ε æ/ overlap only slightly. Speaker 202, by contrast, has more

overlap among mid front vowels: /i eɪ ε/ all overlap in her data, and /æ/ is raised; she also has a fronted /u/. In Fig. 2, a generational comparison is possible between the two AA speakers, both female (200 and 201). Speaker 200, the older AA female, displays little of the overlap between /i ɪ/ or /eɪ ε/ that characterizes the AAVS, but has considerable overlap of /ε æ/, suggesting the low front vowel is raised; additionally, many of her /ɑ/ tokens are fronted and overlap with /æ/, also diagnostic of the AAVS. Speaker 201, a generation younger, shows raising of /æ/ and /ε/, although the latter is highly variable; in her productions, /i ɪ/ also overlap. Both the AA speakers have back /u oo/, compared to the EA speakers in Fig. 1.

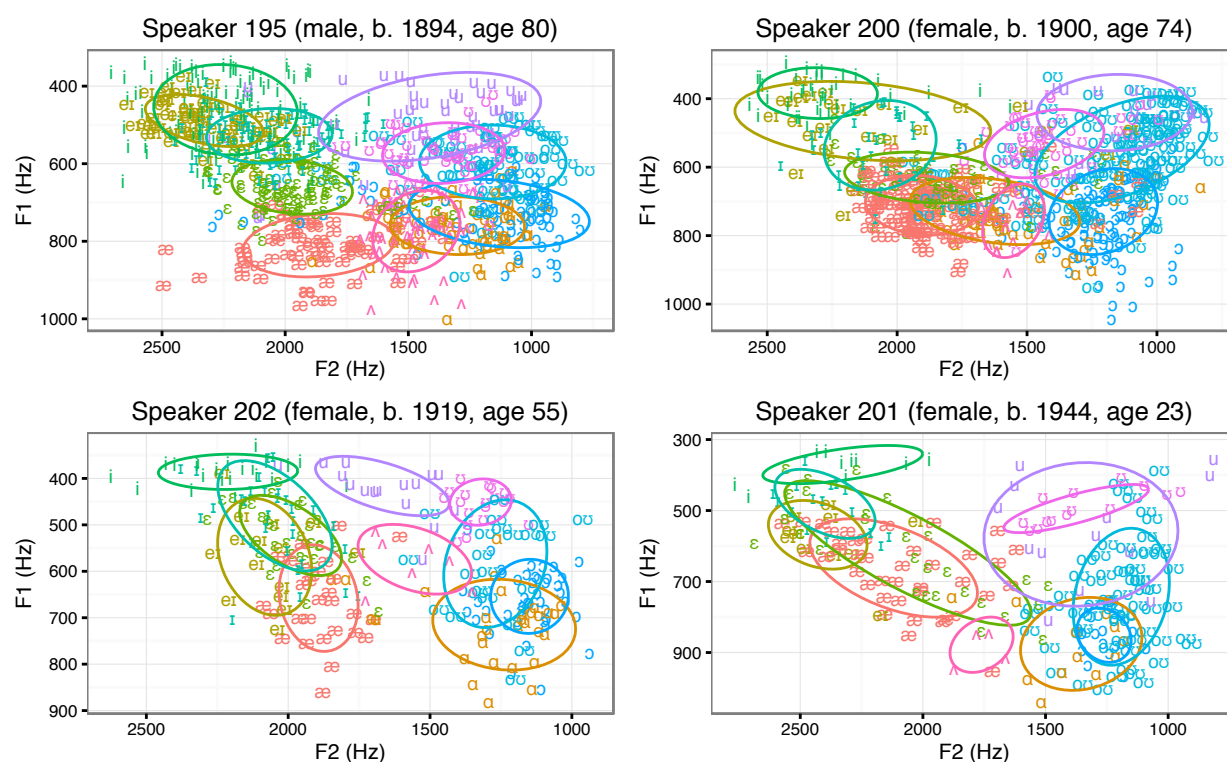


Figure 1. (left) Vowel spaces from EA speakers 195 and 202. (Color online.)

Figure 2. (right) Vowel spaces from the two AA speakers, 200 and 201. (Color online.)

B. Pillai scores

Pillai scores for six pairwise vowel comparisons are summarized in Table III, which is sorted from oldest to youngest speaker within each race group. For all pairs, considerable interspeaker variation occurs. Unless a Pillai score has a high p -value ($p > 0.05$), the MANOVA detected a significant difference across vowel types; a high p -value could indicate merger or near-merger (Hall-Lew, 2010). Additionally, the Pillai score indicates the degree of overlap between the vowels, where 0 would indicate total overlap, and 1 a complete separation in F1/F2 space. Below, the relationships between Pillai scores and race are tested. Race is hypothesized to affect several pairs' scores: because the SVS involves "switching" of front vowel pairs, greater overlap is predicted among EA than AA speakers for /i ɪ/, /eɪ ε/; however, some overlap is expected for AA speakers because of lax vowel raising, which also affects /ε æ/ in both the SVS and AAVS. Back vowel fronting among EA speakers is tested here with /u i/ and /oo i/, where a higher Pillai score would indicate less fronting.⁴ Finally, AAVS fronting of the vowel in LOT is tested here via Pillai scores between /ɑ æ/.

TABLE III. Pillai scores, sorted from oldest to youngest speaker within each race group (EA, then AA). Unless otherwise indicated, $p < 0.001$.

Speaker	/i ɪ/		/eɪ ε/		/ε æ/		/ɑ æ/		/u i/		/oo i/	
	Pillai	p	Pillai	p	Pillai	p	Pillai	p	Pillai	p	Pillai	p
195	0.315		0.823		0.731		0.693		0.717		0.875	
199	0.503		0.359		0.455		0.355		0.359		0.744	
198	0.096	< 0.05	0.248		0.509		0.722		0.556		0.824	
197	0.497		0.132	< 0.05	0.589		0.818		0.245	0.052	0.620	
197A	0.242	< 0.05	0.625		0.658		0.694		0.815		0.817	
201A	0.720		0.030	0.566	0.415		0.677		0.586	< 0.01	0.829	
202	0.492		0.319	< 0.01	0.597		0.867		0.751		0.933	
200	0.690		0.661		0.318		0.404		0.934		0.881	
201	0.614		0.580		0.107	< 0.05	0.791		0.947		0.913	

1. Testing for Vowel Shifts

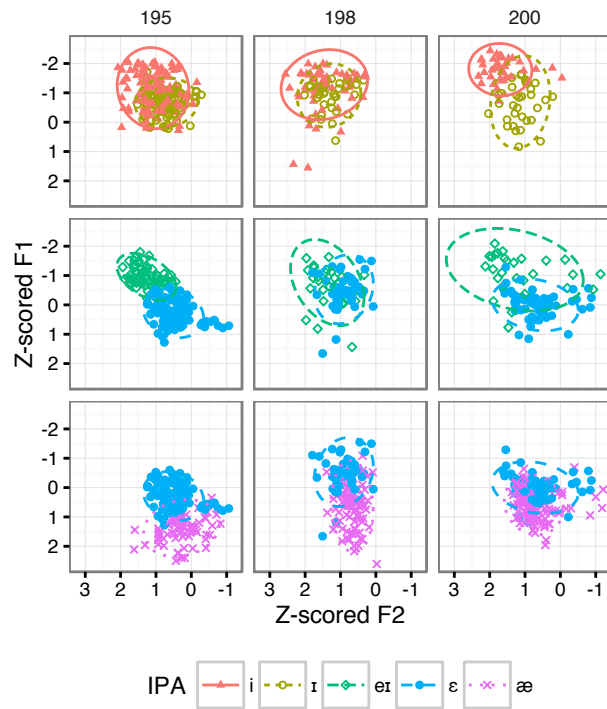
Pillai scores were grouped by speaker race to test for aspects of the SVS and AAVS. One-sided Wilcoxon rank sum tests were used, since the comparison groups have unequal numbers and a non-parametric test is needed. With a larger speaker sample, within-group tests of apparent time differences will become possible, using linear models to control or test for effects of gender and age alongside race.

Among EA and AA speakers, some overlap among adjacent pairs of front vowels is expected, within the SVS and AAVS. Greater overlap for EAs, seen in lower Pillai scores, would indicate raising and fronting of lax vowels, in conjunction with lowering and backing of tense vowels, leading toward a reversal of tense-lax pairs in the vowel space. For /i ɪ/, these values range from 0.132 (speaker 197A, M, Generation 1) to 0.720 (201A, F, Generation 2) among EA speakers, while the two AA speakers have similar values (0.614 and 0.690), indicating low to moderate overlap. For /i ɪ/, the effect of Race approaches significance: EA speakers have lower Pillai scores than AA speakers ($W = 14$, $p = 0.089$). Between /eɪ ε/, Pillai scores for EA speakers ranges from 0.030 (speaker 201A, for whom the distinction is not significant) to 0.823 (speaker 195, male, the oldest speaker); for the AA speakers, scores are again similar and moderate (0.661 and 0.580). For /eɪ ε/, the effect of race is not significant ($W = 11$, $p = 0.267$), but it is significant for /ε æ/ ($W = 1$, $p < 0.05$). EA speakers' range for /ε æ/ lies between 0.415 (201A) and 0.731 (195), while the two AA speakers have much lower scores (0.107 and 0.318). This trend suggests that the effects of raising are similar across races for the /eɪ ε/ pair, but AA speakers have greater /æ/ raising, and the overlap between high vowels is only slightly more advanced in EA than AA speakers.

Turning to back vowels, EA speakers have significantly lower Pillai scores than AA speakers for the /u ɪ/ pair ($W = 1$, $p < 0.05$), indicating greater fronting in that group. A similar

effect is marginal for /oo i/ ($W = 14, p = 0.089$); this is consistent with descriptions that /oo/-fronting is a newer phenomenon (Thomas, 2005), thus less advanced in this historical population. The scores for EA /u i/ range from the greatest fronting, a score of 0.245 (speaker 197, a female in Generation 1), to 0.815 (197A, M, Generation 1); the two AA speakers have the highest scores for that vowel pair (0.934, 0.947), indicating a back /u/. For /oo i/ Pillai scores are generally higher, with a range of 0.620 (197, F, Generation 1) to 0.875 (195, M, Generation 1), while the AA speakers' scores are still higher (0.881, 0.913). Finally, among AA speakers, fronting of low back /a/ is expected. While speaker 200, the older AA speaker, has one of the lowest Pillai scores for the /a æ/ pair (0.404), speaker 201 has a much higher score (0.791), and the effect of race is not significant ($W = 7, p = 0.444$). EA speakers' scores range from 0.355 (speaker 199, M, Generation 1) to 0.867 (speaker 202, F, Generation 2).

The findings for front vowels are captured in Fig. 3. Data are shown for speaker 195 (EA M, Generation 1), the oldest speaker in the corpus whose speech impressionistically shows the fewest Southern features, and speaker 198 (EA M, Generation 1), who impressionistically has the most Southern speech, in comparison to 200, the older AA speaker. Based on the evaluation of Pillai scores, it is clear that speaker 200 participates in the AAVS, while 195 and 198 participate, to different degrees, in the SVS. This is reiterated in Fig. 4, for the back vowels, which also illustrates their intraspeaker variability: data for older (197, F) and younger (201A, F) EA speakers are compared to 201, the younger AA speaker. Speaker 201A has moderate /i u/ overlap but a comparatively compact /oo/ distribution that overlaps little with /i/ in F2 space, while speaker 197 shows fronting of both /u oo/, and the distribution of her /u/ is more compact than its mid vowel counterpart. Speaker 201 has a relatively disperse /u/ which does not overlap with /i/; and her /oo/ is more compact and far back in the vowel space.

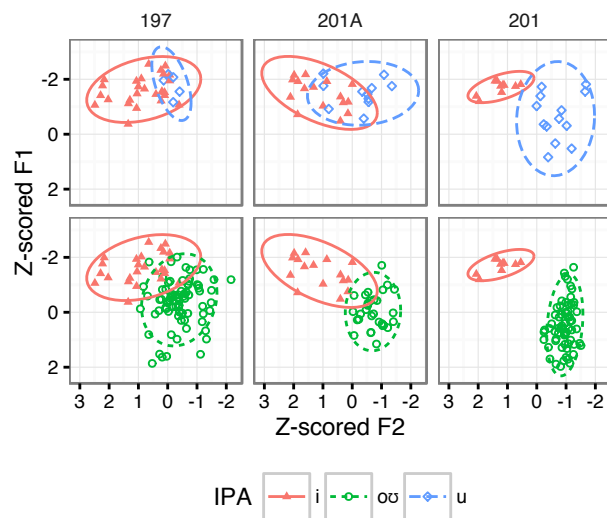


497

498 Figure 3. Illustration of normalized vowel spaces for front vowel pairs, across three speakers

499 (195: M, b. 1894, W; 198: M, b. 1986, W; 200: F, b. 1900, AA). Plots show /i ɪ/ (top), /eɪ ε/

500 (middle), /ε æ/ (bottom) (Color online).



501

Figure 4. Illustration of normalized vowel spaces for front-back vowel pairs, across three speakers (197: F, b. 1897, W; 201A: F, b. 1916, W; 201: F, b. 1954, AA). Plots show /i u/ (top) and /i oo/ (bottom) (Color online).

In summary: both Pillai scores and data visualization indicate that there is significantly more overlap, or less distance, among /i ɪ/, /u i/ for EA speakers, as well as marginally more overlap between /eɪ ε/, /ou i/, compared to AA speakers. This is consistent with participation in the SVS by Southern speakers, *modulo* individual differences in implementation. For AA speakers, the overlap among front vowels, but not back vowels, indicates their participation in the AAVS. Generally, there is less variation among the two AA speakers' Pillai scores than among those of EA speakers, suggesting a uniform implementation of the AAVS even across generational divides; nonetheless, a larger sample of both races, as well as a comparison to male AA speakers, is needed to fully evaluate trends in apparent time.

C. Dynamics of vowels

While many cross- and intra-dialect differences can be seen in F1/F2 plots, it is clear (as described in II.B) that the *dynamics* of vowels also differ. Examining how formants change over time reveals whether vowels are broadly monophthongal or diphthongal, but also can show more subtle distinctions, such as magnitude or rate of formant change across time (Fox and Jacewicz, 2009), or differences in relative timing of formant movements across speakers and dialects.

Below we evaluate formant movement in three types of vowels from the *LAGS* data: (a) the closing diphthongs /aɪ aʊ ɔɪ/, predicted to undergo glide weakening in Southern varieties; (b) the front vowels /i ɪ eɪ ε æ/, and (c) the back vowels /u oo ʌ/, which are expected to be more fronted

in EA than AA speech. We focus on F2 movements in diphthongs and back vowels, but discuss F1 movements for front vowels, given their potential for changing height in the SVS and AAVS.

1. Formant trajectories in LAGS area AK

Figures 5 – 7 illustrate the formant trajectories of different vowels, plotted by race and generation. The plots use Lobanov normalized formant values, and the smoothed trajectories were created with ggplot2 (Wickham, 2009) within R, by fitting a cubic polynomial function to the data. For all vowels, a greater degree of curvature corresponds to a more diphthongal realization. Figure 5 shows the closing diphthongs /aɪ aʊ ɔɪ/, with a separate trajectory for each following context (final, pre-voiced, pre-voiceless) as this can affect degree of glide weakening. Among EA speakers, both generations appear to retain some diphthongal realizations of /aɪ/, although in Generation 2 /aɪ/ is weakest in final position and strongest before a voiceless consonant. Conversely, although Generation 1 EA speakers maintain similar /aʊ/ trajectories across contexts, Generation 2 speakers have the most diphthongal /aʊ/ in final position. Across both generations, /ɔɪ/ is backer (lower F2) before voiceless consonants than in final or pre-voiced contexts. Among the AA speakers, generational differences separate speaker 200 (Generation 1) from speaker 201 (Generation 3). Both have diphthongal /aɪ/ in pre-voiceless position, consistent with previous descriptions of AA speech patterns. While both retain diphthongal /aʊ/, their trajectories are differently shaped; in Generation 1 the trajectories are nearly linear, while the younger speaker's are more curved. In /ɔɪ/, the greatest F2 change occurs in final position for both speakers, but all three trajectories are closer together than those of EA speakers, suggesting less of a contextual effect on that diphthong's realization.

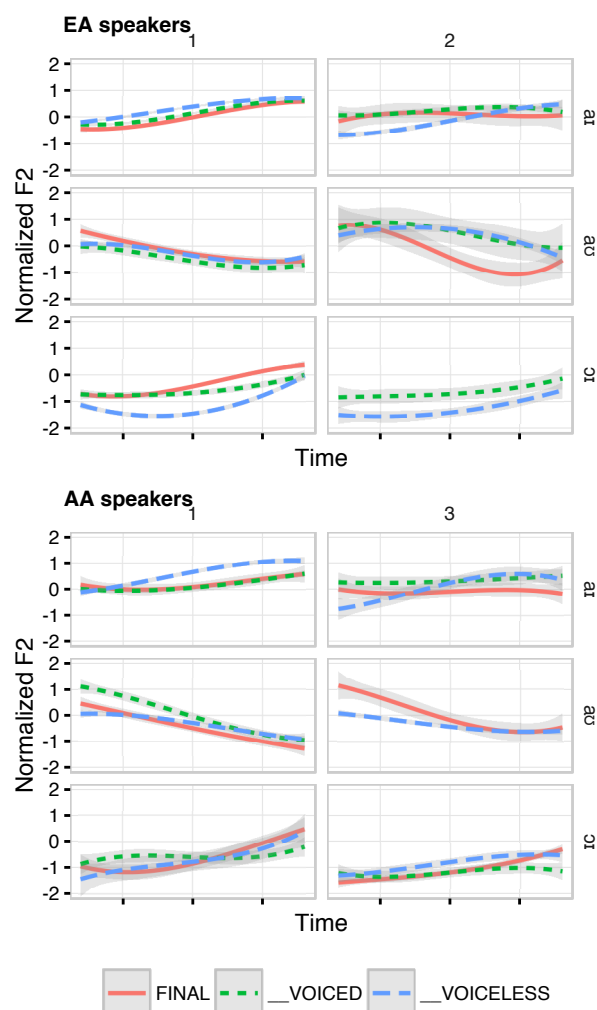
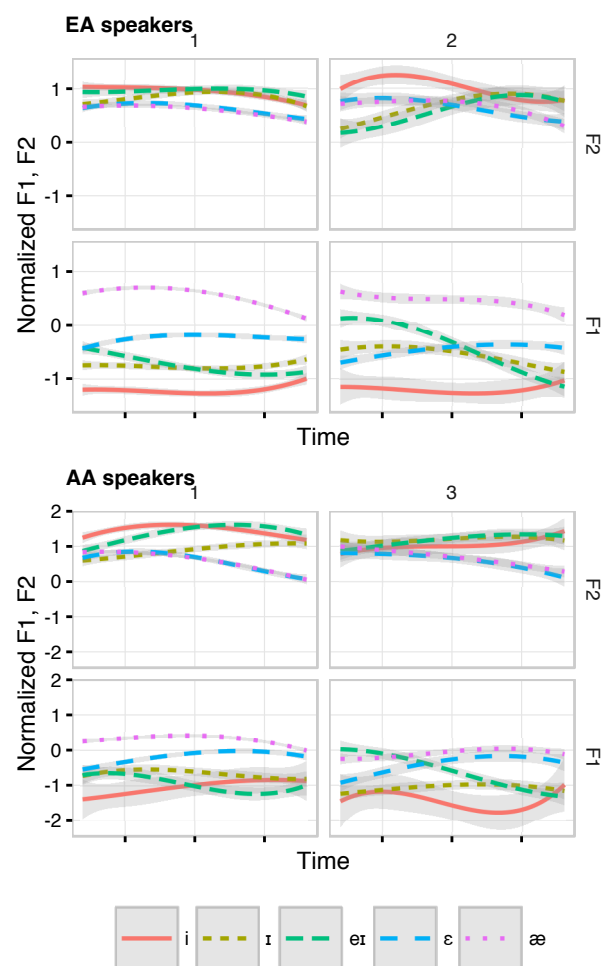


Figure 5: F2 trajectories of closing diphthongs across following contexts for EA speakers (above) vs. AA speakers (below), divided by Generation.

Figure 6 shows smoothed F1 and F2 trajectories of front vowels /i ɪ eɪ ɛ æ/. The tense vowels undergo lowering and backing in the SVS, while the lax vowels can show raising (F1) and fronting (F2) in the SVS and AAVS. Among EA speakers, Generation 1 appears the most conservative in F1: /i/ is the highest vowel, followed by /ɪ/, though /eɪ/'s trajectory ends nearly as high as /i/; /ɛ/ is lower than its tense counterpart, and /æ/ is the lowest front vowel. Among Generation 2 EA speakers, /eɪ/ and /ɛ/ may swap positions, early in their trajectories: /ɛ/ begins

relatively higher in the vowel space, while /eɪ/ begins lower, though its F1 drops (while F2 rises steeply), indicating a high offglide in the tense mid vowel. /æ/ appears raised in Generation 2. Among AA speakers, the small F1 difference between /æ/ and /ɛ/ suggests raising of the low vowel with respect to EA speakers. This is stronger for the Generation 3 speaker, who also has a lowered /eɪ/ compared to the onset of /ɛ/, with respect to the Generation 1 speaker. Lax vowels are described as more monophthongal for AAVS speakers than for SVS speakers, while they may be subject to breaking for SVS speakers; in fact, Koops (2014) uses a dynamic trajectory analysis to argue that Southern /æ/ may be an “incipient triphthong” in Houston Texas. In this data set, there do not appear to be obvious trajectory differences in lax vowels across race groups, but these may be revealed by statistical modeling.

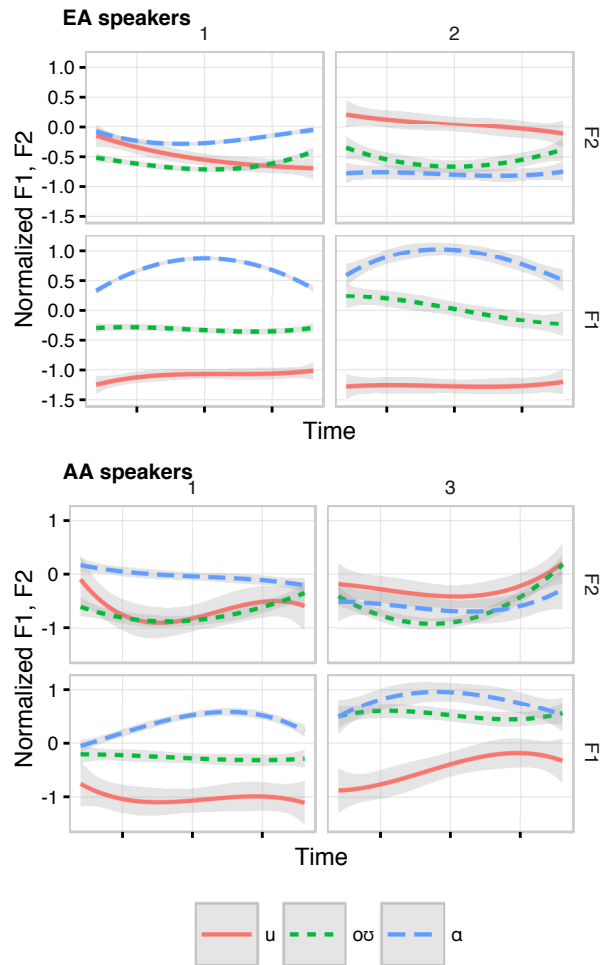
Figure 7 includes trajectories of F1 and F2 for the back vowels /u ʊ ɑ/. For SVS speakers, /u ʊ/ are expected to front; based on patterns in Pillai scores, we expect more fronting of /u/ than /ʊ/. For AAVS speakers, /ɑ/ is expected to be fronted rather than the non-low vowels, though in the Pillai scores above the strongest case for this was in the Generation 1 speaker only. These differences should appear primarily in F2. For EA speakers, an age difference is apparent in Fig. 7: while Generation 1 speakers’ back vowels have similar F2 values to one another, /u/ for Generation 2 has a much higher F2, indicating a fronted vowel, while both /ʊ, ɑ/ retain lower F2 values. The two AA speakers also show different patterns: the Generation 1 speaker has a relatively fronted /ɑ/, indicated by a higher F2, while /u, ʊ/ are farther back; the younger speaker by contrast has a small amount of /u/-fronting, but both /ʊ, ɑ/ have low F2 values.



576

577 Figure 6: F1 and F2 trajectories of front vowels for EA speakers (above) vs. AA speakers

578 (below), divided by generation.



579

580 Figure 7: F1 and F2 trajectories of back vowels for EA speakers (above) vs. AA speakers

581 (below), divided by generation.

582

2. Functional Data Analysis

583 Here, Functional Data Analysis is used to quantify and compare vowel formant trajectories over

584 time. A cubic polynomial was fitted to each token in the data set, producing four coefficients per

585 token. The intercept (a_0) indicates the overall height of the formant, which for F2 corresponds to586 frontness, and for F1 to height, inversely. The linear term (a_1), or slope, is positive when a

587 formant rises over time (e.g. F2 of diphthongal /aɪ/) and negative when it falls (e.g. F2 in

588 diphthongal /aʊ/). The quadratic term (a_2) corresponds to how parabolic the curve is, increasing

when the curve is sharper; and the cubic term (a_3) is expected to increase in absolute value when the formant trajectory is more S-shaped, with greater curvature, as applied in e.g. Risdal & Kohn (2014). A positive cubic term indicates that the final tail of the trajectory is increasing in value, while a negative cubic term implies a downward final curve.

Unlike previous applications of FDA to dialect analysis, which focus on a single coefficient, we test for contextual and sociolinguistic effects on all four terms, a technique that indicates which aspects of a dynamic trajectory vary across vowels, contexts, and groups. We model three groups of vowels predicted to shift differently across social groups, using coefficients of F2 for diphthongs and back vowels, whose shifts are primarily in backness, while F1 coefficients are used to model front vowels, which shift primarily in height for the SVS and AAVS. We hypothesize that among the diphthongs /aɪ aʊ ɔɪ/, the degree of glide weakening may vary across generations (especially among EA speakers) or races; this would be indicated by smaller coefficients for one group. However, a larger factor may be phonological context, as glide weakening applies less before voiceless consonants than elsewhere. Among front vowels /i ɪ eɪ ɛ æ/, given differences in the SVS and AAVS, we expect the intercept and linear coefficients of tense and lax vowels to differ across races; low /æ/ may have different coefficients across races, because of greater raising in the SVS. However, especially due to the unbalanced distribution of speakers across generations, age differences may outweigh or override race differences alone. Finally among back vowels /u oʊ ɑ/ we test F2 coefficients for evidence of fronting; as the SVS exhibits stronger /u, oʊ/ fronting than the AAVS, those vowels' intercept and linear terms may be higher for EA than AA speakers, while the same coefficients may be higher for AA speakers' /ɑ/ due to its reported fronting in the AAVS. Generational effects within races can also be expected (cf. the difference in /u/ across EA generations in Fig. 7).

612 These hypotheses were tested by fitting linear mixed-effects models to the cubic
 613 polynomial coefficients resulting from FDA. For each of the three vowel types (diphthongs,
 614 front, back) a different model was selected, and it was fitted to each coefficient individually for a
 615 total of twelve models. Models were selected on the basis of our hypotheses, and were run in R
 616 using the lme4 and lmerTest packages (Bates and Maechler, 2009, p. 4; Kuznetsova et al., 2013),
 617 for model construction and calculation of *p*-values, respectively. The dependent variables were
 618 the intercept, linear, quadratic, and cubic coefficients resulting from FDA of normalized F1 or F2
 619 values. Treatment coding was used for all categorical predictors, as this method works well for
 620 unbalanced datasets, and interaction terms retain transparent interpretability (Tagliamonte and
 621 Baayen, 2012). Random slopes and intercepts were fitted for speaker, for the factor Vowel Type;
 622 while the maximal random effects structure would involve a greater number of random effects
 623 (cf. Barr et al., 2013), larger models than those presented would not converge. Similarly, because
 624 the number of lexical items sampled for each Vowel Type was small, a random effect of Word
 625 would be highly collinear with Vowel Type and thus was not included (cf. Coleman et al., 2016).
 626 Once all fixed effects had been added, two- and three-way interactions were tested, and retained
 627 if they were significant and the model converged. For diphthongs, the fixed effects were Vowel,
 628 Race, Generation, Position, log of vowel duration (following Risdal and Kohn, 2014; henceforth
 629 log(dur)), Gender; plus a two-way interaction of Vowel:Race, and a three-way interaction of
 630 Vowel:Position:log(dur). For front vowels, fixed effects were Tense (vowels were Tense or Lax),
 631 Height (High, Mid, or Low; note that Tense and Height together replace the factor Vowel), Race,
 632 Generation, Gender, Position, log(dur), plus an interaction of Height:Race. The models for back
 633 vowel coefficients included fixed effects Vowel, Race, Generation, Gender, Position, log(dur),
 634 and an interaction of Vowel:Generation. In the interest of space we do not present the full results

of all twelve models; for each of the three vowel sets we summarize (Tables IV, V, VI) the β values (Estimates) and p -values of each fixed effect and interaction, for each FDA coefficient. Probabilities < 0.05 , calculated with the lmerTest package, are bolded.

We interpret each fixed effect with regard to the coefficients it affects significantly. For diphthongs (Table IV), Vowel affects the linear term (a_1): the slope of /aʊ/ is lower than that of default /aɪ/, indicating the drop in F2 that occurs in /aʊ/. Coefficients for /ɔɪ/ are not significantly different from those of /aɪ/. Race affects the intercept of diphthongs as well: EA speakers have a lower intercept ($p < 0.01$), indicating lower, less diphthongal realizations of /aɪ/. Generation has no significant effect on diphthong coefficients. Position does have some significant effects: before voiceless consonants, slope is higher ($p < 0.05$), and the quadratic term is lower in the same context ($p < 0.001$), indicating a more convex, i.e. diphthongal and rising, trajectory for /aɪ/. Duration has an effect on the quadratic term: as duration increases ($p < 0.01$), so does the quadratic term, suggesting a more concave (less rising) /aɪ/ trajectory. Gender also affects that term: males have a higher quadratic coefficient ($p < 0.05$) than females, suggesting a rising trajectory. However, these fixed factors cannot be fully interpreted outside the context of interactions, several of which are significant. The interaction between Vowel and Race is not significant, but that between Vowel and log(dur) is, indicating that the effect of duration on formant trajectories is not the same across all three diphthongs. The intercept of /ɔɪ/ increases with duration ($p < 0.05$), suggesting higher F2 or greater diphthongization overall for that vowel. Log(dur) also interacts with Position: in pre-voiceless contexts, a duration increase corresponds to a decrease in quadratic term (a more convex trajectory). The significant interaction between Vowel and Position affects both /aʊ/ and /ɔɪ/ in pre-voiceless contexts. The slope of /aʊ/ is more negative there ($p < 0.05$), suggesting a greater drop in F2 and more diphthongal realization. The

intercept of pre-voiceless /ɔɪ/ is lower ($p < 0.01$), but its slope is significantly reduced ($p < 0.05$) and its quadratic term is strongly positive ($\beta = 4.501$, $p < 0.001$), suggesting that the best polynomial fit to /ɔɪ/ in this position curves steeply upward. Finally the three-way interaction between IPA, Position, and Duration is significant, indicating that the effects of Position and Duration on F2 trajectories vary across vowels. Significant effects were found on the slope of /aʊ/ before voiced and voiceless consonants ($p < 0.05$), meaning that the presence of any coda, in combination with an increase in duration, negatively affected the slope of the diphthong (suggesting a more diphthongal /aʊ/). Significant effects on /ɔɪ/ occurred in pre-voiceless position with an increase in duration, for both the intercept, which decreased under those conditions ($p < 0.05$), indicating lower F2 overall, while the quadratic term increased ($p < 0.001$), suggesting a steeper upward curve to the formant's trajectory in that position.

Table IV. Diphthong mixed model results. Default levels are: Vowel = /aɪ/, Race = AA, Generation = 1, Gender = F, Position = Final.

Predictor	a ₀		a ₁		a ₂		a ₃	
	β	p	β	p	β	p	β	p
(Intercept)	-0.157	0.488	1.3	< 0.05	1.011	< 0.01	-0.04	0.876
Vowel = /aʊ/	-0.276	0.58	-2.256	< 0.01	0.116	0.846	0.267	0.548
Vowel = /ɔɪ/	-0.099	0.753	0.37	0.486	-0.487	0.216	0.308	0.338
Race = EA	-0.266	< 0.01	-0.106	0.808	-0.091	0.561	0.04	0.756
Generation = 2	-0.106	0.254	0.04	0.913	0.075	0.525	-0.14	0.227
Generation = 3	-0.205	0.06	-0.193	0.627	-0.154	0.201	-0.185	0.123
+Voice	-0.152	0.548	0.237	0.607	-0.044	0.895	-0.097	0.732
-Voice	0.42	0.134	1.182	< 0.05	-1.413	< 0.001	-0.075	0.811
log(dur)	-0.262	0.071	0.359	0.175	0.622	< 0.01	0.009	0.955
Gender = M	0.026	0.743	0.088	0.792	0.211	< 0.05	-0.13	0.212
/aʊ/ : Race = EA	0.476	0.209	0.264	0.431	-0.231	0.513	0.101	0.413
/ɔɪ/ : Race = EA	0.406	0.069	-0.024	0.919	-0.062	0.75	-0.021	0.869
/aʊ/ : log(dur)	0.128	0.63	0.017	0.972	-0.112	0.746	0.084	0.773
/ɔɪ/ : log(dur)	0.472	< 0.05	0.02	0.951	-0.378	0.114	0.25	0.221
+Voice : log(dur)	-0.151	0.362	0.152	0.614	-0.054	0.803	-0.048	0.794
-Voice : log(dur)	0.137	0.417	0.392	0.203	-0.724	< 0.01	-0.065	0.731
/aʊ/ : +Voice	0.684	0.266	-2.169	0.052	-0.23	0.775	0.265	0.699
/ɔɪ/ : +Voice	0.041	0.91	-0.222	0.737	-0.057	0.904	-0.053	0.895
/aʊ/ : -Voice	-0.485	0.35	-2.167	< 0.05	0.529	0.435	0.725	0.203

Predictor	a ₀		a ₁		a ₂		a ₃	
	β	p	β	p	β	p	β	p
/ɔɪ/ : -Voice	-2.41	< 0.001	-2.964	< 0.05	4.501	< 0.001	-0.757	0.318
/aʊ/ : +Voice : log(dur)	0.478	0.287	-1.6	< 0.05	0.167	0.777	0.143	0.774
/ɔɪ/ : +Voice : log(dur)	0.196	0.422	0.253	0.569	-0.059	0.854	-0.24	0.377
/aʊ/ : -Voice : log(dur)	-0.213	0.522	-1.438	< 0.05	0.313	0.471	0.403	0.269
/ɔɪ/ : -Voice : log(dur)	-0.946	< 0.05	-1.428	0.062	2.262	< 0.001	-0.609	0.195

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

For front vowels, the Estimates in Table V indicate a higher F1 – thus a lower vowel – as they increase. Vowel tenseness significantly affects the intercept and slope of F1 coefficients: specifically, lax vowels have higher intercepts ($p < 0.001$) than tense vowels, consistent with their being lower in the vowel space. Additionally the slope term is significantly higher for lax vowels ($p < 0.05$), suggesting a more rising F1 trajectory. Vowel height also affects F1. (It was not possible to test for an interaction between Tense and Height, because /æ/ was included as a lax vowel but has no tense counterpart.) Mid vowels have significantly higher intercepts than default /i/ ($p < 0.01$), and low vowels' intercepts are still higher ($p < 0.01$). Additionally, the slope of low vowels is negative ($p < 0.05$), consistent with the illustration in Fig. 6 showing a concave F1 trajectory for /æ/ compared to a flatter (slightly positive) one for /i/. Race has a significant effect on front vowels: higher intercepts are found for EA speakers ($p < 0.05$). This indicates that AA speakers have higher tense front vowels (less lowering, as suggested by the Pillai score comparisons in IV.B.1) than EA speakers. Additionally, the quadratic term is affected by race ($p < 0.05$); EA speakers have a more convex (positive) trajectory for front vowels than AA speakers. Turning to age, both Generation 2 and 3 differ significantly from Generation 1. Generation 2, including two EA females, has significantly lower intercepts ($p < 0.05$) and quadratic terms ($p < 0.05$) than Generation 1, suggesting greater vowel raising accompanied by a change in trajectory shape. Generation 3, the young AA female, has a higher F1 slope, which may indicate a greater degree of dynamic F1 change over the timecourse of her

productions. Gender has significant negative effects on both the intercept ($p < 0.001$) and slope ($p < 0.05$) of front vowels. The male speakers, who are also in Generation 1, are thus indicated to have higher front vowels with a more convex F1 trajectory than the AA speaker representing the default levels (cf. the Generation 1 speakers in Fig. 6). Position also has significant effects: before voiced consonants, the linear and quadratic terms are more negative (both $p < 0.001$); before voiceless consonants, the intercept drops with respect to final position ($p < 0.01$). While following context is not known to be a major conditioning factor for front-vowel shifting in Southern EA or AA speech, there are context-dependent F1 differences at least in the set of words examined here. Similarly, $\log(\text{dur})$ significantly affects slope ($p < 0.001$) and the quadratic term ($p < 0.01$), raising both, indicating that vowels' F1 trajectories become lower in the vowel space and more convex as duration increases. For /i ɪ/, this may indicate centralization towards schwa at the end of a vowel, particularly for word-final /i/. Finally, an interaction between Height and Race was tested, but no significant effects were found.

Table V. Front vowel mixed model results. Default levels are: Vowel = /i/, Tense = Tense, Height = High, Race = AA, Generation = 1, Position = Final, Gender = F.

Predictor	a ₀		a ₁		a ₂		a ₃	
	β	p	β	p	β	p	β	p
(Intercept)	-1.195	< 0.001	0.821	< 0.01	0.44	< 0.05	0.321	< 0.05
Tense = Lax	0.398	< 0.001	0.292	< 0.05	-0.116	0.056	-0.015	0.696
Height = Mid	0.463	< 0.01	-0.011	0.94	-0.062	0.619	-0.111	0.277
Height = Low	0.861	< 0.01	-0.384	< 0.05	0.038	0.786	-0.165	0.074
Race = EA	0.45	< 0.05	0.26	0.195	0.415	< 0.05	-0.072	0.487
Generation = 2	-0.235	< 0.05	-0.243	0.093	-0.277	< 0.05	-0.054	0.419
Generation = 3	-0.027	0.818	0.316	< 0.05	-0.016	0.907	-0.067	0.391
Gender = M	-0.394	< 0.001	-0.352	< 0.05	-0.183	0.094	-0.023	0.685
+Voice	0.04	0.441	-0.4	< 0.001	-0.25	< 0.001	-0.077	0.159
-Voice	-0.163	< 0.01	-0.159	0.065	-0.095	0.139	-0.014	0.784
$\log(\text{dur})$	-0.02	0.581	0.247	< 0.001	0.148	< 0.01	0.049	0.217
Height = Mid : Race = EA	-0.073	0.639	-0.101	0.536	-0.117	0.41	0.09	0.421
Height = Low : Race = EA	0.417	0.103	-0.231	0.217	-0.303	0.098	0.12	0.222

For back vowels (Table VI), a positive Estimate indicates an increase in F2 and vowel frontness. Coefficients change significantly depending on Vowel type: /oʊ/ has a lower F2 intercept ($p < 0.01$) and higher linear term ($p < 0.001$) than the default /u/, indicating that it has a backer realization than the high vowel, and a more steeply rising slope (especially for AA speaker 200, represented by the default levels). The slope of /a/ is also higher than that of /u/ ($p < 0.05$). Race does not significantly affect any F2 coefficients of back vowels, indicating that with the measures of vowel positions quantified by FDA, the simple comparison of EA vs. AA speakers does not reveal major trends. However, Generation 2 does have a significantly higher ($p < 0.05$) back vowel intercept. This suggests that the younger, female EA speakers indeed have greater fronting than the older (AA) generation; this is borne out further in the interactions below. Gender, like race, does not significantly affect the coefficients. Position does, however: before voiced ($p < 0.001$) and voiceless ($p < 0.01$) consonants, intercept coefficients decline, indicating slightly backer vowels in that position. The cubic coefficient also increases in pre-voiceless position ($p < 0.05$). Log(dur) has a small in magnitude but significant ($p < 0.05$) effect on the intercept, which decreases, indicating a backer vowel, as duration increases. Finally, the interaction between Vowel and Generation is significant for two F2 coefficients. Generation 2's intercept coefficients are significantly lower ($p < 0.05$) for both /oʊ/ and /a/ than for /u/, which confirms that among younger, EA speakers, fronting is greater for the high vowel than for the mid or low vowels.

Table VI. Back vowel mixed model results. Default levels are: Vowel = /u/, Race = AA, Generation = 1, Gender = F, Position = Final.

	a ₀		a ₁		a ₂		a ₃	
Predictor	β	p	β	p	β	p	β	p
(Intercept)	-0.227	0.138	-0.417	0.12	0.433	< 0.05	-0.027	0.838

	a ₀		a ₁		a ₂		a ₃	
Predictor	β	p	β	p	β	p	β	p
Vowel = /ou/	-0.409	< 0.01	0.619	< 0.001	0.011	0.93	0.097	0.333
Vowel = /a/	0.011	0.952	0.52	< 0.05	-0.106	0.466	-0.059	0.554
Race = EA	0.121	0.411	0.119	0.622	0.029	0.777	0.057	0.552
Generation = 2	0.477	< 0.05	0.131	0.721	-0.276	0.244	0.022	0.904
Generation = 3	0.312	0.215	0.807	0.096	0.17	0.554	0.21	0.381
Gender = M	-0.227	0.089	0.061	0.764	-0.038	0.676	0	0.998
+Voice	-0.559	< 0.001	0.01	0.907	-0.026	0.641	0.065	0.166
-Voice	-0.151	< 0.01	-0.097	0.306	0.011	0.859	0.134	< 0.05
log(dur)	-0.094	< 0.05	0.092	0.239	0.076	0.152	0.062	0.16
Vowel = /ou/ : Generation = 2	-0.504	< 0.05	-0.246	0.361	0.3	0.267	-0.094	0.62
Vowel = /a/ : Generation = 2	-0.979	< 0.05	-0.143	0.737	0.118	0.678	0.073	0.704
Vowel = /ou/ : Generation = 3	-0.345	0.227	-0.309	0.356	0.363	0.294	-0.169	0.489
Vowel = /a/ : Generation = 3	-0.459	0.376	-0.622	0.307	-0.065	0.869	0.004	0.988

731

732 In summary, mixed-effects modeling demonstrates the effects of social variables and

733 phonological context, on vowel dynamics. Diphthongs' F2 coefficients vary by Race and

734 Gender, but not Generation: there is evidence of greater /aɪ/ monophthongization by female EA

735 speakers than by EA males or AA females. Front vowels' F1 coefficients vary by Race,

736 Generation, and Gender: higher tense front vowels are found for AA than EA speakers, though

737 this effect differs across generations; Generation 2 speakers have higher vowels with a different

738 trajectory shape, and male speakers have higher front vowels overall. Back vowels' F2

739 coefficients vary not by Race or Gender, but by Generation: the younger (EA) speakers in

740 Generation 2 have fronter back vowels, particularly /u/. In addition, all three sets of models

741 found significant effects of syllable structure (Position) and vowel duration, as well as the

742 phonological properties of the vowels themselves (Height, Tenseness, Backness). In diphthongs,

743 evidence of the SVS is most visible in long, open syllables: increases in slope coefficients before

744 voiceless consonants indicate greater diphthongization in that position, and duration has a similar

745 effect. Among front vowels, trajectory changes associated with longer durations suggest

746 increased centralization, particularly in /i ɪ/. Conversely, however, back vowels are backer as

747 their duration increases.

748

749 **V. Discussion**

750 Acoustic analysis provides evidence of both the Southern Vowel Shift and African American
 751 Vowel Shift in speakers from southeastern Georgia. These patterns are uncovered by comparing
 752 Pillai scores and formant trajectories across the two groups, and linking them to specific
 753 hypotheses about Southern speech. Despite the small number of speakers involved in this study,
 754 it is possible to show effects of Race, Gender, and Generation by sampling a large number of
 755 tokens per speaker.

756 **A. Evidence for the Southern Vowel Shift**

757 Analysis of speech from *LAGS* reveals that European American speakers show evidence of the
 758 SVS, as predicted for this region of the South; but some speakers participate more fully than
 759 others. Pillai scores show that some EA speakers had moderate to high overlap among /i, ɪ/ in
 760 particular, and significantly lower distances between /i, u/ than did AA speakers, respectively
 761 indicating switching of the high front vowels and fronting of /u/. All EA speakers' /i u/ Pillai
 762 scores are smaller than their /i oo/ score, supporting the claim that /u/-fronting was relatively
 763 more advanced. Additionally, trends in Pillai scores suggest that /ɛ/ is fronted and raised for EA
 764 speakers.

765 However, it is important to note that not all speakers showed these features, and in fact *no*
 766 speaker exhibits all features of the SVS. Wide intradialectal variation occurs even within this
 767 small geographic area, and wide individual variation was observed as well. For example, in
 768 terms of Pillai scores, speaker 198 (male, Generation 1) scores quite low for the /i ɪ/ distinction,
 769 but has a moderate score for /eɪ ɛ/. Conversely, 201A (female, Generation 2) has the highest
 770 score for /i ɪ/, but the lowest score for both /eɪ ɛ/ and /ɛ æ/ among EA speakers.

Previous literature predicts that speakers in this region monophthongize /aɪ/ word-finally and before voiced consonants, and the results of mixed-effects modeling suggest positional differences in /aɪ/. However, considerable interspeaker variation also occurs, as illustrated in Fig. 8 below (cf. Renwick and Olsen, 2016, Fig. 3). Speaker 195 (male, age 80), does not monophthongize in any context, while 198 (male, age 76) monophthongizes throughout – a pattern not predicted to occur in this region of the South. A younger female speaker (202), however, shows the expected pattern.

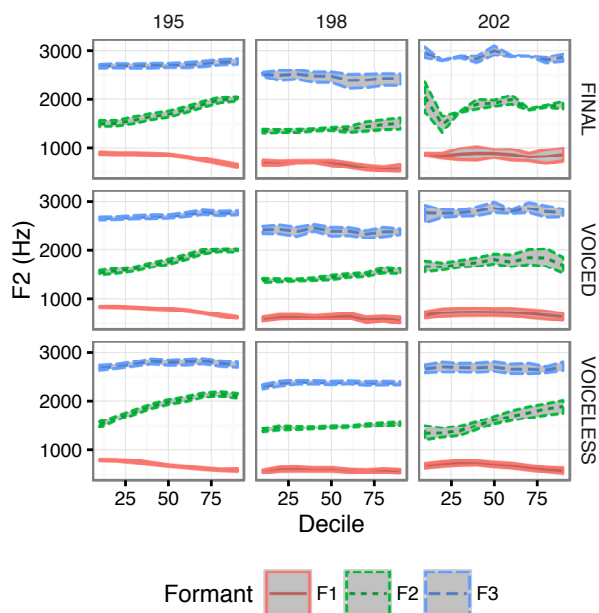


Figure 8: Mean formant trajectories of /aɪ/ across phonetic environments for EA speakers 195, 198, 202 (confidence band = 1SD) (Color online).

Monophthongization of /aɪ/ is claimed to trigger the SVS (Labov et al., 2006). It is a highly salient, stereotypical feature of Southern speech; however, some speakers participate in the SVS without weakening the glide of /aɪ/, while others weaken /aɪ/ without other aspects of

the SVS. Speaker 195, for instance, does not show /aɪ/ monophthongization, but his low Pillai score for /i ɪ/ demonstrates the predicted SVS pattern for the high front vowels. Conversely, while 198 does produce monophthongized /aɪ/, and strong overlap of /i ɪ/ and /e ε/, he does not show evidence of /æ/ raising. This speaker also does not show particularly fronted /oo, u/. Given the diverse collection of SVS features that each speaker displays, one cannot argue for a consistent shift, wherein one vowel movement uniformly triggers another's movement.

B. Evidence for the African American Vowel Shift

The two African American speakers sampled in this study were born nearly 50 years apart, but both hail from Glynn County; both appear to participate in the AAVS, and are consistent with one another. Their Pillai scores are similar, and indicate moderate raising, for /i ɪ/ (0.691, 0.614) and /eɪ ε/ (0.661, 0.580); they have the lowest scores for /ε æ/, indicating stronger /æ/-raising than in EA speech; and neither shows /u, oo/ fronting. Some distinction is apparent in the /ε æ/ scores (0.318, 0.107), where the younger speaker produces a higher /æ/. The two diverge most in the /ɑ æ/ (0.404, 0.791) comparison, indicating differences in LOT-fronting; the older speaker has a fronter /ɑ/. Beyond midpoint data, the speakers' vowel trajectories are quite similar. As in the SVS, the vowel /ɪ/ becomes higher and more fronted over its trajectory (Fig. 6), though it does not reach normalized values of /i/, indicating less raising and fronting for /ɪ/ in the AAVS. Among diphthongs, /aɪ/ is more monophthongal for AA speakers, a result consistent with recent sociolinguistic surveys of Atlanta (Prichard, 2010). The data analyzed here provide evidence that the AAVS was established in Southeast Georgia at the time of recording.

In addition to the variation described across EA speakers, the AA speakers exemplify widespread variation within the speech of individuals. In data for other vowels not analyzed here, like /ɔ/, there is evidence of upgliding, such that tokens of *call* are variably realized as [kol]. This

is particularly strong for AA speaker 200, for whom more than 100 tokens of this word were sampled. This phenomenon, part of the Back Upglide Shift (Labov et al., 2006), is an example of intra-speaker variation that remains for further study. EA speakers are variable as well; for example as seen in Fig. 1, fronting of back vowels among EA speakers is not uniform: some tokens of /u/ are centralized, while others have low F2 values.

A separate question that is beyond our present scope is how other social factors, including socioeconomic status and education, may influence speakers' vowel productions, for example /aɪ/ glide weakening as argued by Labov et al. (2006). Because we discuss only nine speakers, all the levels of these factors were not balanced (for example, there were no black males, and levels of SES and education were not distributed fully across races and genders). An expansion of this study into more *LAGS* speaker areas from will generate a more balanced sample, allowing inclusion of more social factors.

C. Reconciling variable participation with sociolinguistic variation

The variable implementation of SVS features among Southern EA speakers has been found in many sociolinguistic studies, providing evidence for synchronic and diachronic variation. In Charleston, South Carolina, English is losing dialectal features without converging towards the more general Southern dialect, and the change seems to be led by upper-class rather than middle-class speakers, which is typically unexpected (Baranowski, 2008). This indicates that transmission of language change is complex, and that convergence toward the norm is not guaranteed. In Raleigh, North Carolina, a large influx of non-Southern speakers is causing the area to rapidly shift *away* from the SVS in apparent time (Dodsworth and Kohn, 2012); the influence of individual speakers' social networks on their participation in language change highlights the need for detailed consideration of social factors and relationships among speakers.

Dodsworth and Kohn additionally predict – but do not find in their sample – high interspeaker variation if language shift is occurring. Although not a quantified focus of the present paper, linguistic atlas interviews may provide the detailed coverage necessary to detect such variation. Models based on urban speakers, however, are not necessarily applicable to atlas data, which is gathered from speakers who mostly do not know one another and often live in rural settings, where less uniformity and focusing of dialect is expected (Kerswill and Trudgill, 2005; Kerswill and Williams, 2000).

Capturing real trends in regional speech is complex because speakers who are closely connected, in a single community or even a family, may not speak alike. Studies of Southern speakers from Tennessee (Fridland et al., 2013, 2014; Fridland and Kendall, 2012; Kendall and Fridland, 2010, 2012) find considerable differences in the implementation of the SVS. Kendall and Fridland (2010) classify speakers as “shifting” or “non-shifting” based on their participation in the SVS. Among a subset of seven siblings in their production study, not all show the SVS in their speech (Fridland and Kendall, 2012; Kendall and Fridland, 2012). The diversity of shift implementation seen in our results, in the vowel spaces of people from a small geographic area, is consistent with these findings, although the data come from different eras and regions of the South.

VI. Conclusion

Previous work on pronunciation patterns in linguistic survey data has focused on narrow phonetic transcriptions, but this paper presents a method of large-scale analysis for a portion of the *Linguistic Atlas of the Gulf States*. As a corpus, *LAGS* contains thousands of hours of unanalyzed speech data that will contribute greatly to phonetic understanding of Southern U.S.

speech. By documenting variation in speech from a time period for which detailed acoustic analysis is lacking, we have highlighted the value of historical corpora to test and improve dialectal speech models.

The results of acoustic analysis, including Pillai scores, Functional Data Analysis, and mixed effects models reveal the importance of delving beyond impressionistic investigation. Impressionistic analysis deals with one token at a time, while taking an average reduces a variable cloud to a single point. Techniques like Pillai scores take variation into account to produce an estimate of distinction that simplifies group comparisons. Similarly, while close phonetic transcription can capture general changes in vowel quality, FDA decomposes the trajectory into more dimensions that can be linked to specific predictors. By considering objective results that take variation into account, and by considering variation not only across speakers but also within single speakers, linguistic models of dialect will become increasingly representative of natural speech.

VII. Acknowledgments

The authors thank the Linguistic Atlas Project at the University of Georgia for providing the data analyzed here. We are grateful for valuable feedback from Lauren Hall-Lew, Associate Editor Cynthia Clopper, one anonymous reviewer, and the audience at the 170th Meeting of the Acoustical Society of America's Special Session on Advancing Methods for Dialect Variation. We thank Joseph A. Stanley for technical assistance. This work was partially supported by a Presidential Graduate Fellowship from the University of Georgia.

VIII. Notes

¹ Speaker 200 showed relics of Gullah in her speech (Pederson, 1981). She was a cook and maid on St. Simons Island, and was the descendant of slaves brought there before the U.S. Civil War.

² Speaker 199A's recordings were analyzed, but are too poor in audio quality to include in the quantitative studies here. His demographic information is as follows: Ware County; age 50 (born 1929, Generation 2, recorded 1979); European American, male, middle class, 1h 55min of audio, 2,031 words.

³ Birth year is reconstructed based on age in year of recording.

⁴ We follow Podesva et al. (2015) in testing the degree of fronting of *both* /u, ou/ by comparing them to /i/, rather than comparing /ou/ to a vowel of comparable height, like /e/. In our data, despite evidence of /i/-lowering and other height “switches,” /i/ remains the highest and frontest vowel for all speakers and thus is a good reference point for cross-speaker comparisons.

IX. References

Adank, P., Smits, R., and Van Hout, R. (2004). “A comparison of vowel normalization procedures for language variation research,” *J. Acoust. Soc. Am.*, **116**, 3099–3107.

Amengual, M., and Chamorro, P. (2015). “The effects of language dominance in the perception and production of the Galician mid vowel contrasts,” *Phonetica*, **72**, 207–236.

doi:10.1159/000439406

Baranowski, M. (2008). “The fronting of the back upgliding vowels in Charleston, South Carolina,” *Lang. Var. Change*, **20**, 527–551. doi:10.1017/S0954394508000136

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). “Random effects structure for confirmatory hypothesis testing: Keep it maximal,” *J. Mem. Lang.*, **68**, 255–278.

Bates, D., and Maechler, M. (2009). *lme4: Linear mixed-effects models using S4 classes*.

- 901 Bigi, B., and Hirst, D. (2012). "SPeech Phonetization Alignment and Syllabification (SPPAS): a
 902 tool for the automatic analysis of speech prosody," Speech Prosody, Tongji University
 903 Press, Shanghai (China), 19–22. Retrieved from <http://www.sppas.org/>
- 904 Boersma, P., and Weenink, D. (2015). *Praat: Doing phonetics by computer [Computer*
 905 *program]*, Version 5.4.08. Retrieved from <http://www.praat.org>
- 906 Clopper, C. G., Pisoni, D. B., and De Jong, K. (2005). "Acoustic characteristics of the vowel
 907 systems of six regional varieties of American English," J. Acoust. Soc. Am., **118**, 1661–
 908 1676.
- 909 Coleman, J., Baghai-Ravary, L., Pybus, J., and Grau, S. (2012). *Audio BNC: the audio edition of*
 910 *the Spoken British National Corpus*. Retrieved from
 911 <http://www.phon.ox.ac.uk/AudioBNC>
- 912 Coleman, J. S., Renwick, M. E. L., and Temple, R. (2016). "Probabilistic phonetic
 913 underspecification in nasal place assimilation," Phonology [in press].
- 914 Dodsworth, R., and Kohn, M. (2012). "Urban rejection of the vernacular: The SVS undone,"
 915 Lang. Var. Change, **24**, 221–245. doi:10.1017/S0954394512000105
- 916 Dorrill, G. (2003). "Sounding Southern: A look at the phonology of English in the South," In S.
 917 J. Nagle and S. L. Sanders (Eds.), *Engl. South. U. S., Studies in English language*,
 918 Cambridge University Press, Cambridge, pp. 119–126.
- 919 Durian, D. (2012). *A New Perspective on Vowel Variation Across the 19th and 20th Centuries in*
 920 *Columbus, OH* (Ph.D. Dissertation), The Ohio State University.
- 921 Fabricius, A. H. (2007). "Variation and change in the TRAP and STRUT vowels of RP: a real
 922 time comparison of five acoustic data sets," J. Int. Phon. Assoc., **37**, 293–320.
 923 doi:10.1017/S002510030700312X

- 924 Feagin, C. (2003). "Vowel shifting in the southern states," In S. J. Nagle and S. L. Sanders
 925 (Eds.), *Engl. South. U. S., Studies in English language*, Cambridge University Press,
 926 Cambridge, pp. 126–140.
- 927 Fox, R. A., and Jacewicz, E. (2009). "Cross-dialectal variation in formant dynamics of American
 928 English vowels," *J. Acoust. Soc. Am.*, **126**, 2603–2618. doi:10.1121/1.3212921
- 929 Fridland, V., and Kendall, T. (2012). "Exploring the relationship between production and
 930 perception in the mid front vowels of US English," *Lingua*, **122**, 779–793.
 931 doi:10.1016/j.lingua.2011.12.007
- 932 Fridland, V., Kendall, T., and Farrington, C. (2013). "The role of duration in regional US vowel
 933 shifts," *Proc. Meet. Acoust.*, **19**, 060296. doi:10.1121/1.4799047
- 934 Fridland, V., Kendall, Tyler, and Farrington, Charlie (2014). "Durational and spectral differences
 935 in American English vowels: Dialect variation within and across regions," *J. Acoust. Soc.*
 936 *Am.*, **136**, 341–349.
- 937 Grabe, E., Kochanski, G., and Coleman, J. (2003). "Quantitative modelling of intonational
 938 variation," *Proc. SASRTLM Speech Anal. Recognit. Technol. Linguist. Med.*, 45–57.
- 939 Grabe, E., Kochanski, G., and Coleman, J. (2007). "Connecting Intonation Labels to
 940 Mathematical Descriptions of Fundamental Frequency," *Lang. Speech*, **50**, 281–310.
 941 doi:10.1177/00238309070500030101
- 942 Haddican, B., Foulkes, P., Hughes, V., and Richards, H. (2013). "Interaction of social and
 943 linguistic constraints on two vowel changes in northern England," *Lang. Var. Change*,
 944 **25**, 371–403. doi:10.1017/S0954394513000197
- 945 Hall-Lew, L. (2010). "Improved representation of variance in measures of vowel merger," *Proc.*
 946 *Meet. Acoust.*, **9**, 060002. doi:10.1121/1.3460625

- 947 Hay, J., Warren, P., and Drager, K. (2006). "Factors influencing speech perception in the context
948 of a merger-in-progress," J. Phon., **34**, 458–484.
- 949 Kendall, T., and Fridland, V. (2010). "Mapping Production and Perception in Regional Vowel
950 Shifts," Univ. Pa. Work. Pap. Linguist., **16**, 103–112.
- 951 Kendall, T., and Fridland, V. (2012). "Variation in perception and production of mid front
952 vowels in the US Southern Vowel Shift," J. Phon., **40**, 289–306.
953 doi:10.1016/j.wocn.2011.12.002
- 954 Kerswill, P., and Trudgill, P. (2005). "The birth of new dialects," In P. Auer, F. Hinskens, and P.
955 Kerswill (Eds.), *Dialect Change Converging Divergence Eur. Lang.*, Cambridge University
956 Press, Cambridge, pp. 196–220.
- 957 Kerswill, P., and Williams, A. (2000). "Creating a new town koine: Children and language
958 change in Milton Keynes," Lang. Soc., **29**, 65–115.
- 959 Kochanski, G., Grabe, E., Coleman, J., and Rosner, B. (2005). "Loudness predicts prominence:
960 Fundamental frequency lends little," J. Acoust. Soc. Am., **118**, 1038–1054.
961 doi:10.1121/1.1923349
- 962 Koenig, L. L., Lucero, J. C., and Perlman, E. (2008). "Speech production variability in fricatives
963 of children and adults: Results of functional data analysis," J. Acoust. Soc. Am., **124**,
964 3158–3170.
- 965 Koops, C. (2014). "Iconization and the Timing of Southern Vowels: A Case Study of /æ/," Univ.
966 Pa. Work. Pap. Linguist., **20**, 81–90.
- 967 Kretzschmar, W. A. J. (2011). *Linguistic Atlas Project*, Linguist. Atlas Proj. Retrieved from
968 <http://www.lap.uga.edu/>

- 969 Kretzschmar, W. A. J. (2015). "African American Voices in Atlanta," In S. Lanehart (Ed.), Oxf.
 970 Handb. Afr. Am. Lang., Oxford University Press, Oxford, pp. 219–235.
- 971 Kretzschmar, W. A., Jr., McDavid, V., Lerud, T. K., and Johnson, E. (Eds.) (1994). *Handbook of*
 972 *the Linguistic Atlas of the Middle and South Atlantic States*, University of Chicago Press,
 973 Chicago, 454 pages.
- 974 Kurath, H., and McDavid, R. I. (1961). *The Pronunciation of English in the Atlantic States:*
 975 *Based upon the Collections of the Linguistic Atlas of the Eastern United States*,
 976 University of Michigan Press, 182 pages.
- 977 Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2013). *lmerTest: Tests for random*
 978 *and fixed effects for linear mixed effect models (lmer objects of lme4 package)*.
- 979 Labov, W. (1991). "The three dialects of English," In P. Eckert (Ed.), *New Ways Anal. Sound*
 980 *Change*, Academic Press, New York, Vol. 5, pp. 1–44.
- 981 Labov, W., Ash, S., and Boberg, C. (2006). *The Atlas of North American English: Phonetics,*
 982 *phonology, and sound change: A multimedia reference tool*, Mouton de Gruyter, Berlin;
 983 New York, 318 pages.
- 984 Labov, W., Rosenfelder, I., and Fruehwald, J. (2013). "One hundred years of sound change in
 985 Philadelphia: Linear incrementation, reversal, and reanalysis," *Language*, **89**, 30–65.
- 986 Labov, W., Yaeger, M., and Steiner, R. (1972). *A quantitative study of sound change in progress,*
 987 *US Regional Survey*, Vol. 1.
- 988 Lobanov, B. M. (1971). "Classification of Russian vowels spoken by different speakers," *J.*
 989 *Acoust. Soc. Am.*, **49**, 606–608.

- 990 MacLagan, M., and Hay, J. (2007). "Getting fed up with our feet: Contrast maintenance and the
 991 New Zealand English front vowel shift," *Lang. Var. Change*, **19**, 1–25.
 992 doi:10.1017/S0954394507070020
- 993 McCarthy, C. (2010). "The Northern Cities Shift in Real Time: Evidence from Chicago," *Univ.*
 994 *Pa. Work. Pap. Linguist.*, **15**, 101–110.
- 995 Montgomery, M. B., and Nunnally, T. E. (1998). *From the Gulf States and Beyond: The Legacy*
 996 *of Lee Pederson and LAGS*, University of Alabama Press, Tuscaloosa, Alabama, 376
 997 pages.
- 998 Mooshammer, C. (2010). "Acoustic and laryngographic measures of the laryngeal reflexes of
 999 linguistic prominence and vocal effort in German," *J. Acoust. Soc. Am.*, **127**, 1047–1058.
- 1000 Pederson, L. (1981). *Protocol of a field record for the linguistic atlas of the Gulf States*,
 1001 University Microfilms International, Ann Arbor, MI.
- 1002 Pederson, L., McDaniel, S. L., and Adams, C. M. (Eds.) (1986). *Linguistic Atlas of the Gulf*
 1003 *States*, University of Georgia Press, Athens, Georgia, Vols. 1-7.
- 1004 Podesva, R. J., D'Onofrio, A., Van Hofwegen, J., and Kim, S. K. (2015). "Country ideology and
 1005 the California Vowel Shift," *Lang. Var. Change*, **27**, 157–186.
 1006 doi:10.1017/S095439451500006X
- 1007 Prichard, H. (2010). "Linguistic Variation and Change in Atlanta, Georgia," *Univ. Pa. Work.*
 1008 *Pap. Linguist.*, **16**, 140–149.
- 1009 R Core Team (2000). *R Language Definition*, Available from CRAN sites. Retrieved from
 1010 <ftp://155.232.191.133/cran/doc/manuals/r-devel/R-lang.pdf>
- 1011 Ramsay, J., Hooker, G., and Graves, S. (2009). *Functional data analysis with R and MATLAB*,
 1012 Springer, New York, 207 pages.

- 1013 Reed, P. (2014). "Inter- and intra- generational /ai/ monophthongization and Southern
 1014 Appalachian identity," *South. J. Linguist.*, **38**, 154–194.
- 1015 Renwick, M. E. L., Baghai-Ravary, L., Temple, R., and Coleman, J. (2014). "(t,d) Deletion in
 1016 Everyday Speech," Tokyo, Japan. Presented at the LabPhon 14, the 14th Conference on
 1017 Laboratory Phonology.
- 1018 Renwick, M. E. L., and Olsen, R. M. (2016). "Voices of coastal Georgia," *Proc. Meet. Acoust.*,
 1019 **25**, 060004. doi:10.1121/2.0000176
- 1020 Risdal, M., and Kohn, M. (2014). "Ethnolectal and generational differences in vowel trajectories:
 1021 Evidence from African American English and the Southern Vowel System," *Univ. Pa.*
 1022 *Work. Pap. Linguist.*, **20**, 139–148.
- 1023 Schützler, O. (2015). *A sociophonetic approach to Scottish Standard English*, Varieties of
 1024 English around the world: volume G53, John Benjamins Publishing Company,
 1025 Amsterdam; Philadelphia, 179 pages.
- 1026 Tagliamonte, S. A., and Baayen, R. H. (2012). "Models, forests, and trees of York English:
 1027 Was/were variation as a case study for statistical practice," *Lang. Var. Change*, **24**, 135–
 1028 178. doi:10.1017/S0954394512000129
- 1029 Thomas, E. R. (2001). *An acoustic analysis of vowel variation in New World English*,
 1030 Publication of the American Dialect Society, Duke University Press, Durham, NC, Vol.
 1031 85, 230 pages.
- 1032 Thomas, E. R. (2003). "Secrets Revealed by Southern Vowel Shifting," *Am. Speech*, **78**, 150–
 1033 170.
- 1034 Thomas, E. R. (2005). "Rural white Southern accents," In E. W. Schneider (Ed.), *Var. Engl.*
 1035 *Am.*, Mouton de Gruyter, Berlin, pp. 87–114.

- 1036 Thomas, E. R. (2007). "Phonological and Phonetic Characteristics of African American
1037 Vernacular English," *Lang. Linguist. Compass*, **1**, 450–475. doi:10.1111/j.1749-
1038 818X.2007.00029.x
- 1039 Thomas, E. R. (2010). "A Longitudinal Analysis of the Durability of the Northern-Midland
1040 Dialect Boundary in Ohio," *Am. Speech*, **85**, 375–430. doi:10.1215/00031283-2010-022
- 1041 Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*, Springer Science & Business
1042 Media, Dordrecht; New York, 212 pages.
- 1043