Back Vowel Dynamics and Distinctions in Southern American English

Joseph A. Stanley  
Brigham Young University, joey_stanley@byu.edu

Margaret E. L. Renwick  
University of Georgia

Katherine Ireland Kuiper  
University of Georgia

Rachel M. Olen  
University of Georgia

Follow this and additional works at: https://scholarsarchive.byu.edu/facpub

Part of the Arts and Humanities Commons

Original Publication Citation

BYU ScholarsArchive Citation
Stanley, Joseph A.; Renwick, Margaret E. L.; Kuiper, Katherine Ireland; and Olen, Rachel M., "Back Vowel Dynamics and Distinctions in Southern American English" (2021). Faculty Publications. 6129. https://scholarsarchive.byu.edu/facpub/6129

This Conference Paper is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Back Vowel Dynamics and Distinctions in Southern American English

Joseph A. Stanley¹, Margaret E. L. Renwick², Katherine Ireland Kuiper², and Rachel M. Olsen²

Abstract
Southern American English is spoken in a large geographic region in the United States. Its characteristics include back-vowel fronting (e.g., in GOOSE, FOOT, and GOAT), which has been ongoing since the mid-nineteenth century; meanwhile, the low back vowels (in LOT and THOUGHT) have recently merged in some areas. We investigate these five vowels in the Digital Archive of Southern Speech, a legacy corpus of linguistic interviews with sixty-four speakers born 1886-1956. We extracted 89,367 vowel tokens and used generalized additive mixed-effects models to test for socially-driven changes to both their relative phonetic placements and the shapes of their formant trajectories. Our results reinforce previous descriptions of Southern vowels while contributing additional phonetic detail about their trajectories. GOOSE-fronting is a change in progress, with greatest fronting after coronal consonants. GOAT is quite dynamic; it lowers and fronts in apparent time. Generally, women have more fronted realizations than men. FOOT is largely monophthongal, and stable across time. LOT and THOUGHT are distinct and unmerged, occupying different regions of the vowel space. While their relative positions change across generations, all five vowels show a remarkable consistency in formant trajectory shapes across time. This study’s results reveal social and phonetic details about the back vowels of Southerners born in the late nineteenth and early twentieth centuries: GOOSE-fronting was well underway, GOAT-fronting was beginning, but FOOT remained backed, and the low back vowels were unmerged.

¹Brigham Young University, Provo, UT, USA
²University of Georgia, Athens, GA, USA

Corresponding Author:
Joseph A. Stanley, Department of Linguistics, Brigham Young University, Provo, UT, 84602, USA.
Email: joey_stanley@byu.edu
1. Introduction

Among varieties of English spoken in the United States, the dialect of the Southern states is distinct, and often stereotyped in popular culture for its characteristic pronunciation patterns and morphosyntax (e.g., pronouncing /aʊ/ so that fire is a near homophone of far, or the use of y’all and double modals like might could). A variety of recent studies have presented evidence that speakers in the US South, especially young speakers, are shifting away from “traditional” speech patterns (Prichard 2010; Dodsworth & Kohn 2012; Stanley forthcoming). Synchronic studies are crucial to document the current state of language and to glimpse incipient linguistic change. In this paper, however, we contribute to this body of work by beginning to fill in the diachronic backstory of Southern US speech. We draw from a large, newly available phonetic dataset taken from a corpus of legacy recordings gathered from eight contiguous states: Georgia, Florida, Alabama, Tennessee, Mississippi, Louisiana, Arkansas, and Texas. These provide a multi-generational perspective on the Southern vowel system, as its speakers approached the set of characteristic speech patterns captured by the Atlas of North American English as the Southern Vowel Shift (Labov, Yaeger & Steiner 1972; Labov, Ash & Boberg 2006). We distill acoustic measurements into snapshots for three generations of Southern men and women, highlighting the relative positions of vowels both within the vowel space and across apparent time. This illustrates changes in vowel pronunciation that took place among speakers born over a seventy-year period spanning 1886 to 1956.

This paper investigates the sociophonetics of back vowels in Southern US English, focusing on the relative phonetic placements and vowel-inherent dynamics of phonemes represented by the words GOOSE, FOOT, GOAT, THOUGHT, and LOT. In Standard American English, these words’ stressed vowels are typically transcribed as /u ʊ ʊʊ/, plus /ɔɑ/ (respectively) when the low back vowels are unmerged. Southern US English can diverge considerably from these qualities, as described in more detail in section 2 and in Thomas (2004). Our data come from the Digital Archive of Southern Speech (DASS; Kretzschmar et al. 2013), a collection of Linguistic Atlas interviews from eight southern states. DASS’ forty-eight European American speakers (twenty-four women) were born between 1886 and 1956, and contribute 290 hours of audio, whose extent and time depth allow us to test for sound patterns characteristic of Southern speech. The preservation of these recordings is significant because audio samples from talkers born in the nineteenth century are rare. Though interviews took place between 1970 and 1983, this corpus has only recently been transcribed (Kretzschmar et al. 2019), providing a fresh look at Southern US English in the early and mid-twentieth century.
Southern American English (SAE) is spoken in a large geographic region in the Southern United States, extending from Texas to Virginia. The *Atlas of North American English* (Labov, Ash & Boberg 2006) defines multiple dialect areas within this region, including the “Texas South” and “Inland South,” where Southern speech features are particularly concentrated; more localized varieties can also be found in areas such as Appalachia, Charleston, and New Orleans. Despite variation within SAE, certain features are characteristic of the entire region, like glide-deletion in the *PRICE* vowel, and the Southern Vowel Shift (Labov, Yaeger & Steiner 1972; Labov 1994; Labov, Ash & Boberg 2006), which includes the raising and diphthongization of the front lax vowels in *KIT* and *DRESS* while lowering the nuclei of tense vowels in *FLEECE* and *FACE*.

Back vowels have also undergone change in the American South, though they are not as socially marked as glide deletion in *PRICE* nor as conditioned by social factors like age, gender, and ethnicity as shifts among the front vowels (Renwick & Stanley 2020). One such change is Back Vowel Fronting (BVF), or the fronting of *GOOSE*, *GOAT*, and, to some extent, *FOOT*. BVF is widespread in the American South, though it is also found in many other dialects of North American English (Labov, Ash & Boberg 2006), so it is not uniquely Southern (Fridland 2012). Among the low vowels, *THOUGHT* and *LOT*, SAE is described as shifting these vowels and, especially in younger generations, merging them (Baranowski 2008).

One aspect of back vowels that is under-described, particularly for the US South, is their formant dynamics. Our research questions therefore focus on the formant trajectories of back vowels. We test whether their shape and relative position vary, across gender and across generational time, using DASS. This corpus is an ideal test bed for an investigation of back vowels because its speakers are a cross-section of the population through which sound changes spread within the South. Among vowels affected by BVF, we expect variable fronting across phonemes and also allophones in postcoronal contexts versus elsewhere, as described in more detail in section 2. Our analysis of the low back vowels *THOUGHT* and *LOT*, discussed further in section 2.1, evaluates the presence of merger. Where the social factors of gender and generation are concerned, we test whether women and men are at the same stage of ongoing back-vowel changes. We hypothesize that across generations, younger speakers have more advanced BVF, and we anticipate little evidence of a low back merger, except in the youngest generation.

Analysis of formant dynamics requires the consideration of multiple formant measurements per vowel token. We argue that this technique contributes to our understanding of phonetic and sociolinguistic variation in human speech, and we use it to illustrate diachronic change in Southern US vowel systems. We use measurements from five time points across each vowel token, and we apply generalized additive mixed-effects models (or GAMMs; Wood 2017a) to nearly 90,000 back-vowel tokens.

Before describing our dataset and methods further, we use the remainder of this introduction to review previous literature on back vowel fronting (section 2) and the low back merger in greater detail (section 2.1), especially with respect to SAE and the relevant vowels’ formant trajectories (section 2.2).
2. Background

Although most sociophonetic descriptions focus on the front vowels or PRICE (Standard American English /aɪ/), fronting of GOOSE and GOAT is likely older than some front-vowel shifts in Southern speech. Fridland (2012) describes back vowel changes documented in Civil War veterans’ data. In fact, demographic shifts resulting after the Civil War may have spurred the spread of GOOSE-fronting (Thomas 2004). BVF is now widespread across Southern speakers of different ethnicities and social classes (Fridland 2003; Fridland & Bartlett 2006), although age-related effects remain (e.g., Thomas 2004). For example, among speakers from Alabama, Feagin (2003) compares the more conservative system of a man born in 1881 against the vowel space of a woman born in 1957, whose back vowels exhibit a large degree of fronting. With regard to social class, BVF is led by high-status women and younger speakers in South Carolina (Baranowski 2008), but urban working-class speakers in Alabama (Feagin 2003). Phonological context is also relevant for Southern BVF: it is rare before /l/, hiatuses (going, do it), and nasals (Thomas 2004). Surveys of SAE from the Linguistic Atlas Project (Kurath & McDavid 1961) and the Telsur project (Labov, Ash & Boberg 2006) show that fronting of GOOSE is greater after coronal consonants. A phonetic split is found for both GOOSE and GOAT in Kansas City (Strelluf 2018) and Manchester, England (Baranowski 2017) where postcoronal tokens shift at different rates and as a result of different social factors from other allophones of these vowels.

Of the three higher back vowels, fronting of GOOSE is the oldest, most widespread, and most advanced. In the speech of “older” white Southerners the high back vowel can be realized as [ʉ̞ʊ̟] or [y̞ʉ̟], fronting further to [y̞y] among younger speakers (Thomas 2004:99), compared to the Standard American English transcription of [u]. As mentioned above, it began at least 150 years ago. By the time of data collection for the Linguistic Atlas of the Gulf States (LAGS; Kurath & McDavid 1961), GOOSE-fronting had begun across the entire South. Today, it continues to be a feature in all varieties of SAE (Labov, Ash & Boberg 2006), though it is less advanced among African American Southerners (Thomas 2001; Fridland 2003). The large geographic coverage may not necessarily be due to geographic diffusion alone: Fridland (2012) points out that physiologically-driven factors like overcrowding in the back portion of the vowel space may have caused GOOSE-fronting to develop independently and language-externally in multiple regions. Koops (2010) provides phonetic evidence of this independent development by showing two different realizations of fronted GOOSE in Houston: a monophthongal “southern” variant ([y]) and a back-gliding diphthongal “mainstream” variant ([ɨn]). These two variants also coincide with Thomas’ (2004) transcriptions of GOOSE with a central nucleus [u] or even a front [y]. Thanks to its robust nature in the American South, we expect to find GOOSE-fronting in our legacy data, but because it was a change in progress in the South during the early twentieth century, we expect some degree of variation in our corpus.

Though FOOT-fronting is typically considered part of BVF in Southern speech, descriptions of the vowel are less detailed than for GOOSE or GOAT. The degree to which it
fronts is usually correlated with goose- and goat-fronting (Thomas 2004), although Feagin (2003) finds that it chronologically comes between them in Alabama. The vowel is realized as [ʊ] in Standard American English, but among Southern white speakers it can vary between [ʊ] and [y] (Thomas 2004:96). Its shifted form is primarily found among younger speakers born well into the twentieth century, and thus it may have been the last of the three vowels to shift (Fridland 2001). Since foot-fronting is a recent, if not current, change in progress in the South, it is unlikely to be an active sound change in the dataset analyzed here.

In SAE, goat-fronting strikes a middle ground between goose and foot. Compared to Standard American English /ou/, the Southern diphthong can appear as [œu] or [eu] among older, White Southerners, and as [3y] or [æu] in younger speakers (Thomas 2004:99). It is a more recent development than goose-fronting, and Thomas (1989) finds its distribution was once limited to North Carolina, the Delmarva Peninsula, and the Georgia-Alabama border. Thomas (2004) suggests that goat-fronting may have originated in northeastern North Carolina in the late nineteenth century, and though it experienced relatively slow geographic diffusion initially, racial polarization occurring during the Civil Rights Movement helped it become a widespread feature among European Americans in the South. These ethnic divisions could explain why goat-fronting is not a prominent feature in Southern varieties of African American Language (Thomas 2007).

Among speakers in DASS, goat-fronting is expected to be less advanced than goose-fronting, since extreme forms such as [æu] are primarily found in young people (Thomas 2004). While goose fronting to [u] is described as a “majority feature” in LAGS, goat-fronting is only a “minority feature” (Kurath & McDavid 1961; Thomas 2001, 2007). However, the acoustics of these vowels’ relative placements have not yet been examined on a large scale. Examining the data from DASS will not only expand the geographic region in which this phenomenon has been studied, but it will also uncover a change in progress, evidenced by generational differences in the corpus.

2.1. Low Back Vowels in the US South

The merger of the low back vowels lot and thought (Standard American English /ɑ/ and /ɔ/, merging toward [ɑ]) has become one of the most widespread sound changes in North American English in the past century. It is characteristic of the western United States and Canada (Labov, Ash & Boberg 2006), though it has developed independently in other pockets like eastern Pennsylvania (Herold 1990) and can be found in numerous other regions of North America like western Pennsylvania and eastern New England (Labov, Ash & Boberg 2006). In younger generations of speakers in the South, the low back merger is reported in areas like rural Kentucky (Irons 2007), Tennessee (Fridland 2015), South Carolina (Baranowski 2013), Georgia (Stanley forthcoming), and Florida (Doernberger & Cerny 2008).

Although the merger of lot with thought is increasingly widespread in the US, we expect it to be comparatively rare in DASS speakers. Even in Southerners a generation
young than the youngest speakers in DASS, LOT is relatively backed but THOUGHT is high, resulting in a clear distinction between the two vowels in both F1 and F2 (Clopper, Pisoni & De Jong 2005). It is this higher realization of THOUGHT, particularly with its accompanying upglide, that has inhibited the merger (Thomas 2001; Labov, Ash & Boberg 2006). While LOT remains [ʌ] among white Southerners, THOUGHT is more diphthongal: [ɔo] or [ʌn] (Thomas 2004:96-98). In fact, Irons (2007) finds that, as THOUGHT’s offglide is lost, the distinction between the two vowels becomes untenably small, resulting in a merger. However, Fridland, Kendall, and Farrington (2014) find that there is an inverse correlation between spectral differences and durational differences and that duration keeps the vowels distinct even if their formant trajectories are similar. Because our data include speakers born well before the low back merger is reported to be widespread in the South, we expect to see clear differences in these two vowels in DASS, particularly in their formant dynamics.

2.2. The Role of Dynamics in Back Vowels

Although several Southern back vowels have been described as “generally diphthongal” (Wells 1982b:539), their vowel-inherent formant dynamics are rarely explored; we seek to fill this gap in the literature. Using data from Alberta, Michigan, and Texas, Nearey (2013) shows that the five vowels we analyze exhibit varying degrees of inherent formant movement, often influenced by surrounding consonants. GOAT and, to a lesser extent, GOOSE show higher, backer offglides, FOOT has a shorter centralized offglide, and the low vowels were more monophthongal but become lower over the course of their durations. In adult speakers from Houston, Texas, Assmann and Katz (2000) describe the same general patterns, though GOOSE was more monophthongal (and centralized), and the low vowels were more diphthongal with centralized offglides. The difference between these two descriptions shows that there is regional variation in back vowels’ formant dynamics, though more detailed phonetic descriptions are lacking.

The dynamic nature of Southern back vowels has been noted in the literature. Thomas (2004) describes the widespread fronting of the nucleus of GOOSE among White Southerners ([u, u]) with the offglide itself possibly also fronted ([u, uu]). In fact, Koops (2010) finds these two distinct realizations of GOOSE co-occurring in Houston Anglos. Both versions have a fronted nucleus, but while the “southern” variant also has a fronted offglide, resulting in relatively little spectral change along the vowel’s duration, the “mainstream” version is more diphthongal with its backed offglide (see also Hinrichs, Bohmann & Gorman 2013). These two forms are auditorily distinct and the critical difference between them is their formant dynamics, rather than the nucleus. Turning to the dynamics of GOAT, Thomas (2004) describes the most conservative forms of that vowel as monophthongal [oː]. However, virtually all speakers born in the last 100 years use a diphthongal variant, with regional variation in the quality of the nucleus and glide (e.g., [u, u, u, a, a]; Thomas & Coggshall 2014). Similarly, Thomas (2004) transcribes THOUGHT as diphthongal and upgliding ([ɔo, oʊ, oʊ, oʊ]) with [ɔʊ] being the more frequent form in contemporary speakers. Thomas’ (2004) thorough
account of vowel qualities among white Southern speakers does not mention any diphthongization in *foot* or *lot*.

While sociophonetic studies on front vowels in the South are plentiful (Fox & Jacewicz 2009; Fridland, Kendall & Farrington 2014; Renwick & Olsen 2017, among others), detailed acoustic analyses to support descriptions of back vowels are rare for many parts of the South. The most thorough descriptions come from the peripheries of the South, usually outside the region analyzed in this study. Thomas (2001) provides individual vowel plots for dozens of Texans and North Carolinians. These plots suggest that many speakers have fronted variants of *goose* and *foot* in all contexts (including prelaterally) and a lowered and diphthongal *goat*. Their low vowels are either merged, or *thought* is diphthongal with a raising trajectory while *lot* remains monophthongal.

The role of formant dynamics in the low-vowel distinction is explicitly examined in Missouri (Majors 2005) and Kentucky (Irons 2007), both of which are outside the LAGS area. Holt (2016) takes a different approach to analyze whether there is regional variation in formant dynamics within North Carolina. One analytic measure used was the spectral rate of change, which measures how quickly formants change and how that speed varies over the course of the duration of the vowel (Farrington, Kendall & Fridland 2018). Holt (2016) finds that European American speakers realize *goose* and *goat* with different spectral rates of change, with speakers in western North Carolina having a lower rate of change (meaning more monophthongal vowels) than speakers in eastern North Carolina. Furthermore, Holt (2018) extracted measurements at multiple time points from these speakers to analyze additional measurements like trajectory length, but a rigorous analysis of the formant dynamics was lost for the back vowels since the measurements from all timepoints were averaged together. Thus, while researchers have acknowledged vowel trajectories in back vowels of the American South, an in-depth analysis of them is lacking.

Recently, researchers focused on other varieties of English have adopted a more quantitative approach to back vowel trajectories. GAMMs (Wood 2017a) allow rigorous analysis of vowel trajectories by fitting nonlinear, smoothed lines to the data while testing for significant differences between speaker groups or other factors. Sóskuthy, Foulkes, Hughes, and Haddican (2018) analyze *goose* with and without preceding /j/ in the English Midlands to highlight subtle differences in trajectory shapes across phonological contexts and generational time. Warburton (2018) uses GAMMs to study mergers among *goat* and *thought* in the North of England. Finally, Strycharczuk and Scobbie (2016) use a related model, SS-ANOVA, to analyze the dynamic nature and contrast between allophones of *goose* in Southern British English. These techniques, which uncover more nuanced differences in formant trajectories than other analyses, are underutilized in studying vowel formant dynamics in North American English. One recent example is Onosson’s (2018) analysis, via GAMMs, of “abbreviatory patterns” of /au/ in Canadian English. Hualde et al. (2021) use GAMMs to support their analysis of the emerging marginal phonemic contrast between /ai/ and /au/ in Chicago English. While Renwick and Stanley (2020) use GAMMs to analyze formant dynamics of front vowels in the American South, to our knowledge, GAMMs have not previously been used as an analytic tool for back vowels in this region.
3. Data and Methods

3.1. The Digital Archive of Southern Speech

The complete DASS interviews were selected by the editor of the LAGS project, Lee Pederson, to constitute a representative sample across multiple social dimensions, a total of sixty-four speakers. The speakers interviewed for LAGS hailed from eight states: Georgia, Florida, Alabama, Tennessee, Mississippi, Arkansas, Louisiana, and Texas. Its European American (EA) speakers were categorized into three social “types”: Type 1, “Folk” speakers (less educated, few community ties); Type 2, “Common” speakers (at least some high school education and community ties); and Type 3, “Cultivated” speakers (some college or equivalent education, experience traveling outside the community, and strong social ties; Kretzschmar et al. 2013). Although African American (AA) speakers also varied along social dimensions such as social class, these speakers were summarily categorized in LAGS as a cohesive fourth “type.” For each of the sixteen regional sectors into which LAGS was divided (see Figure 1 of Pederson 1981 for a map of these sectors), one EA speaker of each “type” (i.e., three EA speakers) and one AA speaker are included in DASS. DASS speakers represent a wide range of ages at the time of interview (15-90 years, with a mean age of 60.3 years) and birth years ranging from 1886 to 1965.

Originally recorded on reel-to-reel and cassette tape, LAGS (and thus DASS) has been digitized and is maintained by the Linguistic Atlas Project (LAP; Kretzschmar 2011). A full transcription of the DASS audio was completed in 2019 (Kretzschmar et al. 2019; see Olsen, Olsen, Stanley, Renwick & Kretzschmar 2017 for a fuller description of this endeavor). During this process it became apparent that five of the original DASS speakers’ interviews were unintelligible due to poor audio quality, and replacement speakers were selected matched precisely for regional sector and social “type.” The revised DASS corpus consists of over 367 hours of fully transcribed audio. Duration of interviews ranged from roughly two to ten hours with the average interview lasting about five and a half hours. Interview length and content vary widely based on the talkativeness of the speaker, but each contains narrative portions of continuous speech as well as question/answer portions in which fieldworkers elicited specific lexical items. These portions are not treated separately.

Our analysis only includes EA speakers. Because of the small number of AA speakers in the corpus, they are not included in the analysis here (for other work that explicitly includes AA speakers in LAGS/DASS, see Renwick & Olsen 2017; Renwick & Stanley 2020). Table 1 summarizes the characteristics of the EA sample with respect to the social factors we model, namely gender and generation, and Figure 1 shows these speakers’ geographic locations, with shapes representing gender, and color representing generational cohort (see section 3.3 on how these are defined).

We do not consider subregional variation within DASS, that is, the relationship between the state or sector a speaker comes from and their degree of participation in
back vowel changes. At present, the size of our speaker sample is insufficient to address variation due to geography, since adding a factor like STATE would introduce eight additional subdivisions of the dataset. In our work with DASS, we have found that consistent geographically-driven patterns are scarce and can be overridden by variation from other sources. However, some work on DASS has successfully illustrated subregional variation. For example, Olsen, Olsen, and Renwick (2018) show that /au/ monophthongization varies systematically across the six sub-regions into which DASS is divided (Coast, Delta, Highlands, Piedmont, Piney Woods, and Plains; see Pederson, McDaniel & Adams [1986] for more details). Using GIS mapping techniques of local spatial autocorrelation analysis, Jones and Renwick (2020, forthcoming) have presented evidence that there are feature-specific pockets of greater and lesser adherence to Southern speech features in DASS (see also Renwick & Olsen 2016; Olsen, Olsen, Stanley, Renwick & Kretzschmar 2017; Renwick & Stanley 2017; Stanley, Kretzschmar, Renwick, Olsen & Olsen 2017).

<table>
<thead>
<tr>
<th>Table 1. Number of DASS Speakers Analyzed by Gender and Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lost Generation</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Men</td>
</tr>
</tbody>
</table>
3.2. Data Processing

The DASS interviews were manually transcribed at the utterance level by trained student workers. The transcriptions were subsequently checked multiple times for typos. For the purposes of this study, passages that overlapped with interviewer speech or were tagged as unclear were excluded. Interviews consisted of both narrative speech and portions featuring elicitation of lexical items. We acknowledge that there may be stylistic variation depending upon the kind of speech featured, but we did not include this as a factor in the current analysis. After manual transcription, the corpus was force-aligned using an in-house installation of the Montreal Forced Aligner (McAuliffe, Socolof, Mihue, Wagner & Sonderegger 2017). Due to the size of this corpus, it was not possible to manually correct all alignments, but systematically listening to roughly 10,000 tokens of individual words shows that approximately 86 percent of transcribed words are correctly aligned with the corresponding audio. This is a higher accuracy rate than has been found in other automatically aligned corpora of natural speech, such as the Audio British National Corpus, which found an average word boundary accuracy rate of approximately 66 percent (Coleman, Renwick & Temple 2016; Renwick & Cassidy 2015).

Together with audio files, the Praat TextGrids (Boersma & Weenink 2018) resulting from forced alignment were processed using a local installation of FAVE (Rosenfelder et al. 2014) to measure formants at five timepoints for each token, at 20 percent, 35 percent, 50 percent, 65 percent, and 80 percent of vowel duration. Including data from AA speakers, 1,979,244 total vowel tokens were measured, leaving 1,480,519 for the EA speakers analyzed here. Formant values were converted into Barks (Zwicker 1961), using the formula presented in Traunmüller (1990), to transform them from a linear to a logarithmic scale. Vowels that came from highly frequent stopwords, did not have primary (lexical) stress, or preceded liquids were removed. Only GOOSE, FOOT, GOAT, THOUGHT, and LOT, as coded by The CMU pronouncing dictionary (Lenzo 2013), were retained in this dataset, leaving 102,229 vowel tokens for analysis. We acknowledge that relying upon classifications listed in The CMU pronouncing dictionary may introduce some inaccuracies in the results, particularly for the low vowels, since regional or individual variation in lexical specification may occur for certain words.

Because back vowels tend to be more fronted after coronal sounds (e.g., /t/), the lexical sets of GOOSE, FOOT, and GOAT were each split into two allophones based on whether the preceding sound was a coronal consonant. The labels we assign for the postcoronal allophones are TOOT, TOOK, and TOE, respectively, and for the elsewhere allophones, they are BOOT, PUT, and BOAT. We retain the use of the labels GOOSE, FOOT, and GOAT when referring to the vowel class as a whole rather than specific context-dependent allophones of those vowels.

Data from legacy audio, like the DASS recordings, is expected to produce outliers as a result of incorrect measurements, but correcting these manually is impractical with a large dataset. Consequently, we have filtered the data based on each measurement’s Mahalanobis distance (Mahalanobis 1936; Labov, Rosenfelder & Fruehwald 2013), to automatically remove potential outliers. This technique measures how far each token is in the F1-F2 space from the average F1-F2 measurements for that vowel. It is effective in filtering out tokens that have extreme (and likely incorrect) formant measurements due to formant tracking.
errors since it accounts for F1 and F2 simultaneously (for additional details, see Renwick & Ladd 2016; Renwick & Stanley 2020; Stanley 2020). Filtering was carried out independently for each speaker, for each allophone (that is, TOOT and BOOT were treated separately), and for each time point along the vowels’ durations (i.e., distances were calculated separately for data from the 20 percent versus. 35 percent measurements, etc.). Measurements with a high Mahalanobis distance, which we define as greater than the 95th percentile based on a chi-squared distribution, were excluded from analysis. After this Mahalanobis filter was applied, excluding 14,014 individual measurements, we removed any vowel token that did not have a complete set of ten measurements remaining (that is, F1 and F2 at all five time points), excluding 11,436 tokens. After these filters were applied, 89,367 vowel tokens remained for analysis, distributed across the vowels and allophones according to Table 2.

3.3. Treatment of Speaker Metadata

DASS includes metadata for a variety of factors, as described in section 3.1. Of interest to the present study is speakers’ age. For the purposes of this study, speakers were grouped by generational cohort, coded as a three-level categorical variable. As shown in Figure 2, speakers born between 1886 and 1900 were coded as the Lost Generation ($N = 17$). Those born between 1901 and 1927 are the G.I. Generation ($N = 19$). Typically, the Silent Generation would include those born between 1928 and 1942 and the Boomer Generation would be those born after 1942, but since there were only seven speakers from the Silent Generation and five from the Boomer Generation, these two were combined into a single Silent/Boomer Generation ($N = 12$) to create a third group approximately equal in size as the other two. These bins, which take their names and age ranges from Strauss and Howe’s

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOOSE</td>
<td>15,872</td>
</tr>
<tr>
<td>BOOT</td>
<td>2,763</td>
</tr>
<tr>
<td>TOOT</td>
<td>13,109</td>
</tr>
<tr>
<td>FOOT</td>
<td>8,767</td>
</tr>
<tr>
<td>PUT</td>
<td>7,002</td>
</tr>
<tr>
<td>TOOK</td>
<td>1,765</td>
</tr>
<tr>
<td>GOAT</td>
<td>33,615</td>
</tr>
<tr>
<td>TOE</td>
<td>6,938</td>
</tr>
<tr>
<td>BOAT</td>
<td>26,677</td>
</tr>
<tr>
<td>THOUGHT</td>
<td>12,639</td>
</tr>
<tr>
<td>LOT</td>
<td>18,474</td>
</tr>
<tr>
<td>TOTAL</td>
<td>89,367</td>
</tr>
</tbody>
</table>

Note: Bold indicates overall vowel classes, which are combined to calculate the total, while non-bold rows show counts for postcoronal versus elsewhere allophones.
work on generational theory, represent cohorts that broadly share similar experiences and worldviews. We expect speech patterns within generations to pattern similarly.

3.4. Modeling

Our dataset was subjected to statistical modeling in order to test whether vowel realizations changed significantly across generations, genders, and allophonic contexts. Since the speakers in DASS represent a stratified but random sample of US Southerners, a mixed-effects approach is necessary to argue that any findings could generalize beyond this study. In order to model formant values as trajectories, or curves unfolding in time, we adopt GAMMs (Wood 2017a). This type of model is useful for trajectory data as it can account for vowel-inherent formant dynamics (i.e., nonlinear curves), while simultaneously modeling the effect of demographic factors like generation and gender and their effect on the position and shape of the trajectory. A schematic of the R code for models in this paper is provided, and interpreted, in Figure 3. This code corresponds to “full models” discussed in the paper.

We built four separate GAMMs for back vowels in DASS: one each for GOOSE, FOOT, and GOAT, plus a fourth for LOT and THOUGHT together. The same model specification was applied to each. Bark-transformed formant measurements of both F1 and F2 were used as the dependent variable in each GAMM. The models’ independent variables included FORMANT and ALLOPHONE (or, in the case of the low back model, VOWEL), since all data for each vowel was pooled together (see Gahl & Baayen 2019; Faraway 2006:214–218; Wieling et al. 2016). Social predictors included GENDER and GENERATION. To allow for each combination of GENDER and GENERATION to have its own best fit curve for each formant and allophone, these four variables (FORMANT, ALLOPHONE, GENDER, GENERATION) were combined into a single four-way interaction term. The reference level of this interaction term was for F1 of the elsewhere allophone, or LOT for the low back model, corresponding to a woman in the G.I. Generation. The interaction was included as a parametric effect and as a smooth in the model. Smooth terms, within a GAMM, capture non-linear effects. The smooth included PERCENT to indicate the normalized time point from which each

Figure 2. Speaker Birth Years and Generations by Gender

(1991)
measurement came. The smooth used four knots, a parameter to functionally subdivide modeled curves; four is the maximum available when five measurement points are used. Additionally, \textsc{log duration} was included as a parametric effect in all models.

On top of these fixed effects, a random effects structure was incorporated into the GAMM to account for multiple sources of random error in the data. \textsc{Speaker} was included as a random intercept and slope, interacting with \textsc{formant} and \textsc{allophone}, allowing the model to adjust the position of the best fit curves independently for each speaker. \textsc{Word} was included as a random intercept, interacting with \textsc{formant} and \textsc{allophone}, which also allows for some word-level adjustments within each vowel, independently for F1 and F2. Furthermore, \textsc{previous segment} and \textsc{following segment} were both included as random smooths, interacting with \textsc{formant} and \textsc{allophone}, which allow the model to make by-segment adjustments to the overall shape of the predicted trajectories. Random slopes per \textsc{word} and random smooths for both \textsc{speaker} and \textsc{word} were desired but impossible due to constraints on computational power. These random effects account for much of the idiosyncratic effects and speaker- and word-level phonological correlations that are present in the data but are not the focus of this study (see also Gahl & Baayen 2019; Renwick & Stanley 2020). Their inclusion allows the analysis and comparison of vowel trajectory variation above and beyond the variation found in individual speakers, words, and neighboring segments. They also prevent overfitting to the data and allow the model to be more generalizable to the population at large.

These full models were fitted using the \texttt{bam()} function within the \texttt{mgcv} package in R (Wood 2017b). The output for these four models can be found in the Supplemental Materials (posted online). As the models are complex and lengthy, we analyze them chiefly through visualizations of predicted measurements, extracted from each model for each allophone, gender, and generation. Additionally, due to the challenges in

<table>
<thead>
<tr>
<th>Model Specification in R</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 \texttt{ngcv::bam(}</td>
<td>1 Function call.</td>
</tr>
<tr>
<td>2 \texttt{bark_raw ~}</td>
<td>2 Dependent variable: Back-transformed, raw values.</td>
</tr>
<tr>
<td>3 \texttt{formant_allophone_gender_generation +}</td>
<td>3 Fit different smooths for each combination of formant, allophone, gender, and generation.</td>
</tr>
<tr>
<td>4 \texttt{log_duration + formant_allophone_gender_generation +}</td>
<td>4 Control for duration, including interaction term.</td>
</tr>
<tr>
<td>5 \texttt{s(previous_phone, percent, by = formant_allophone, bs=&quot;fs&quot;, x=x=&quot;cr&quot;, m=1, k=4) +}</td>
<td>5 Random smooths for preceding and following phone, each interacting with formant and allophone.</td>
</tr>
<tr>
<td>6 \texttt{s(speaker, formant, allophone, bs=&quot;re&quot;, k=4) +}</td>
<td>6 Random intercept for slope and speaker, each interacting with formant and allophone.</td>
</tr>
<tr>
<td>7 \texttt{s(word, formant, allophone, bs=&quot;re&quot;, k=4)}, \texttt{s(word, formant, allophone, bs=&quot;re&quot;, k=4)}</td>
<td>7 Random intercept for word, by formant and allophone.</td>
</tr>
<tr>
<td>8 \texttt{data = dass_subset)}</td>
<td>8 Data specification (varies by vowel).</td>
</tr>
</tbody>
</table>

Figure 3. Model Specification in R (Left), with Interpretation of Code (Right)
directly interpreting GAMM summaries, the significance of each factor was tested via model comparison, following Renwick and Stanley (2020). Corresponding to each vowel’s full GAMM, we built a series of models that were equivalent except that one predictor (gender, generation, or allophone) was eliminated, or dropped, from the model. We also created a baseline model, in which gender, generation, and allophone were not incorporated in a single interaction term with formant, but were each crossed separately with them, and treated as parametric predictors only. All model specifications (five per vowel) can be found in the Supplemental Materials (posted online). Using the compareML() command within the itsadug R package (van Rij et al. 2017), we evaluated the effectiveness of the full model versus each impoverished/baseline model. For each of the four vowels (or vowel pairs) tested, we expect that this comparison will indicate a statistically significant advantage for the full model. Such an outcome justifies the increased complexity of that model, in this case indicating significance of the predictors gender, generation, or allophone, and justifying the use of smooth terms.

4. Results

We model back vowel trajectories in DASS to understand how they are influenced by the social factors of speaker age and gender, and we also discuss differences in the realization of positional allophones. Our focus is on visualizations of model output, but first we present the robustness of that output via model comparisons. For each of the four vowels treated by GAMM, the full model (including social and phonological predictors) accounts for more than 91 percent of the variance in the data, indicating a good model fit. The adjusted $R^2$ values, which indicate model fit, appear alongside the results of model comparisons in Table 3. For each vowel modeled, columns 2–5 of Table 3 indicate the difference in log-likelihood $\chi^2$ for each model comparison (full < nested). Note that degrees of freedom (final row; full > nested) vary only across model structures, not across vowel types. Adjusted $R^2$ values for full models appear in the sixth column. In all cases, the full model shows significant improvement ($p < 0.001$) over the baseline. This indicates that there is considerable nonlinearity in the data and justifies our use of GAMMs to construct complex models. Full models also show

<table>
<thead>
<tr>
<th>Vowel modeled</th>
<th>Baseline</th>
<th>Drop gender</th>
<th>Drop generation</th>
<th>Drop vowel</th>
<th>Adj. $R^2$ full</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOOSE (BOOT, TOOT)</td>
<td>388.641</td>
<td>233.721</td>
<td>150.891</td>
<td>1064.949</td>
<td>0.959</td>
</tr>
<tr>
<td>FOOT (PUT, TOOK)</td>
<td>70.816</td>
<td>145.652</td>
<td>55.090</td>
<td>861.329</td>
<td>0.971</td>
</tr>
<tr>
<td>GOAT (BOAT, TOE)</td>
<td>323.155</td>
<td>169.085</td>
<td>100.427</td>
<td>1925.676</td>
<td>0.910</td>
</tr>
<tr>
<td>LOWBACK (LOT, THOUGHT)</td>
<td>503.226</td>
<td>260.060</td>
<td>388.253</td>
<td>7295.406</td>
<td>0.909</td>
</tr>
<tr>
<td>Difference in df</td>
<td>63</td>
<td>48</td>
<td>64</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>
improvement over models lacking individual social or phonological factors, which is evidence of statistically significant variation in trajectory characteristics across genders, generations, and phonological contexts.

In the Supplemental Materials (posted online), we include model summaries for all models, as well as the output of all model comparisons. The full models’ output shows, in almost all cases, that F1 terms are not significantly different with respect to the reference level (F1 of *boot, put, boat, or lot* from a woman of the G.I. Generation), while F2 terms are. Coefficient values are much higher for F2 than for F1, which is expected given the large differences between those formants’ typical Bark values. Since GAMMs can incorporate complex interaction terms and nonlinear effects, the results are best summarized as a visualization in the F1-F2 vowel space. Figure 4 provides such a visualization and illustrates several important findings uncovered by these GAMMs. Each panel of the figure represents speakers from one gender, and each arrow-tipped curve shows the predicted trajectory of one allophone as spoken by a particular generation. The arrow represents predicted formant measurements at 80 percent of vowel duration, while the unadorned end of each line corresponds to 20 percent. By visually comparing trajectories of an individual allophone across generations, we see how its shape and position changed across time.

4.1. Description of Formant Trajectories

We first describe the general direction and shape of vowel trajectories of the elsewhere allophones of the higher back vowels, and of the low vowels. First, all back vowels display various degrees of spectral change. The tense vowels *boot* and *boat* begin with a more centralized onset and end with a more peripheral glide. The women realized *boot* as somewhat monophthongal since the overall change in F2 from the onset to the glide is relatively short (0.29 Barks). However, for the men, it is more dynamic (0.61 Barks). For both genders, the nucleus is lower than the offglide, a pattern found in other phonetic studies in the South (Koops 2010). Where vowel trajectory direction is concerned, *boot* and *boat* both raise and back in the vowel space, though this is more apparent in *boat*. Between *boot* and *boat* is *put*, which displays a different pattern. Unlike the tense vowels, this lax vowel starts in a more backed position and progresses to a more centralized point in the F1/F2 space. It is ingliding, unlike the other vowels in this study. The length of *put*’s trajectory is more consistent across all groups (standard deviation = 0.07 Barks) than *boot* (standard deviation = 0.22 Barks) and primarily changes along the front-back dimension. Among the women, *put*’s trajectory gets slightly longer as its onset backs over three generations, though this change is relatively small.

The low vowels, *lot* and *thought*, exhibit a third pattern. They become more peripheral over their time courses, but in slightly different directions: *lot* primarily lowers, while *thought* primarily backs, though this varies by gender. Most groups realize *lot* with a centralized offglide. The trajectory of both low vowels is U-shaped, meaning their F1 measurements peak at some point near the midpoint of the vowels’ durations. With an average trajectory length of 0.73 Barks, *thought* is more diphthongal than *lot*, which
averages 0.52 Barks. It is evident that LOT and THOUGHT occupy different portions of the vowel space and show different patterns in their formant trajectories. We therefore conclude that the two are distinct vowels in this sample, with little evidence of merger between them.

4.2. Back Vowels Shift at Different Rates

While the previous section addressed the overall shapes of formant curves, this section addresses the rate of change in the relative positions of vowels in the F1/F2 space. Figure 4 shows that the relative positions of back vowels in DASS change in apparent time. With BOOT, we see that women fronted this vowel a small amount between the Lost and the G.I. Generations, and a larger amount after the G.I. Generation. The Silent/Boomer Generation (in both genders) lowered BOOT as well. The oldest generation of men realize BOOT only slightly fronter than PUT, while all other groups exhibit virtually no overlap in the F2 between BOOT and PUT. Judging by the distance between PUT, which exhibits relatively little change in apparent time in this sample, and BOOT, the women appear to have a more fronted BOOT than the men do.

Unlike BOOT and BOAT, which do show change in apparent time, PUT remains stable across generations. While the model comparisons showed that the addition of GENERATION as a predictor did significantly improve the full model’s fit, we presume much of this improvement comes from the postcoronal allophone, TOOK, which shows more apparent change in time (see section 4.3).

Turning to BOAT, Figure 4 shows that all three generations of women had different realizations of the vowel. The Lost Generation of women had the highest vowel while
the G.I. and Silent/Boomer Generations used a lower variant. The G.I. group also had the most diphthongal realization with an offglide somewhat further back than the other groups, though this is not retained in the youngest generation. Therefore, the bulk of the change in boat lowering by women occurred between the Lost and the G.I. Generations, or roughly around 1900.

DASS men also lowered boat in apparent time, but unlike the women, there is clearer evidence of BVF. The Lost and G.I. Generations of men used a higher and backer variant and Silent/Boomer speakers used a lower and fronter variant. Therefore,
the bulk of the change in boat fronting and lowering by men occurred between the G.I. and the Silent/Boomer Generations, or roughly around 1924.

Among low back vowels, patterns are less clear. Both genders show evidence of lowering in apparent time, in tandem with boat-lowering. The Lost Generation of women used higher variants of lot and thought, while the G.I. and Silent/Boomer Generations realized these vowels lower in the vowel space. Among men, the lowered forms are only evident in the Silent/Boomer Generation. In addition to lowering, younger speakers also backed these vowels by a small degree.

Regarding the lowering of boat, lot, and thought, these shifts do not appear to be the start of a robust change. The fact that the younger two generations of women had similar F1 measurements suggests that this change had already reached completion by the Silent/Boomer Generation (as early as 1924).

While Figure 4 allows easy intergenerational comparisons, its two panels obscure the relative placements of allophones within each generation. This alternative view is provided in Figure 5, which plots predicted vowel trajectories for each gender and generation separately. Unlike Figures 4 and 6, which plot women and men with gender-specific scales to highlight the details of vowel trajectories within each group, all sub-plots of Figure 5 are deliberately plotted on the same scale to highlight the difference in relative positions of back-vowel systems across genders and generations. The placements of front vowels fleece and trap, based on averages from a single formant measurement point, are provided for reference. Figure 5 provides trajectories for both allophones in this study, showing that toot and boot were distinct for all groups, with the postcoronal allophone placed slightly higher in the vowel space than other realizations. The postcoronal allophone took is more fronted than the elsewhere allophone put, which is an expected consequence of coarticulation; those two allophones of foot converge upon the same location in F1/F2 space.

4.3. Postcoronal Allophones

We have not yet discussed the postcoronal allophones, toot, took, and toe, outside of Figure 5; Figure 6 highlights these allophones, and for reference, the corresponding trajectories (boot, foot, and boat) from Figure 4 are shown as thinner, transparent lines with open arrow heads. While we attempted to account for most allophonic behavior through the random effect for word, previous segment, and following segment, we felt it prudent to include postcoronal allophones in each model because of their potentially large expected differences (postcoronal allophones of back vowels tend to be fronter than non-postcoronal allophones). We did this by placing them in the large GAMM interaction term to test for change and variation between allophones. For example, as mentioned in section 2.1, Strelluf (2018) finds that the postcoronal versus elsewhere allophones of goose and goat exhibit different patterns across time and are evaluated differently by Kansas Citians,
suggesting that there is potential for sociolinguistic differences between these pairs of allophones in regions with BVF. A model that collapses these allophones into a single category would obscure potential systematic differences. The results of model comparison (Table 3) confirm that allophones are significantly different for all the non-low back vowels.

We did find a small amount of sociolinguistic conditioning in TOOT, TOOK, and TOE. Phonetically these were fronter than BOOT, FOOT, and BOAT, as expected. Among the women, there is little change in TOOT’s F2 in apparent time, suggesting that the fronting shift was complete at the time of data collection. BOOT fronts and approaches TOOT such that among the Silent/Boomer women, the F2 difference between TOOT and BOOT was very small, indicating that BOOT fronting had neared completion. The trajectory of TOOT shows that most of the change is in F2, with the vowel generally backing over its time course, and relatively little change in F1.

TOOK was quite monophthongal for all groups. Across both genders, there was a small amount of backing in apparent time. For the men, there was a small amount of movement, and though its offset was close to PUT’s offset, its onset was lower and fronter. The older generations of men have a more dynamic TOOK vowel, suggesting that some degree of TOOK-monophthongization has occurred among younger speakers. The position of TOOK across generations suggests that TOOK has retracted in the vowel space over apparent time.

Finally, the difference between TOE and BOAT was relatively small. Like its elsewhere counterpart, TOE is diphthongal and exhibits the same general shape of backing and raising in the vowel space. TOE lowers in apparent time alongside BOAT. It also exhibits the same complex shifts in F2 as was seen in BOAT: the G.I. Generation women backed the vowel first and then the Silent/Boomer Generation women fronted its offset.

Figure 6. Predicted Formant Trajectories of Postcoronal Allophones, Based on GAMM Analysis of Back Vowels Spoken by European Americans in DASS
while the older two generations of men had similar realizations, only to have the Silent/Boomer Generation men front it. Overall, the postcoronal allophone is fronter, though the difference between TOE and BOAT gets smaller as BOAT fronts in apparent time.

The sociolinguistic differences represented by TOOT, TOOK, and TOE are consistent and potentially important. Additional research on the dynamics of these postcoronal allophones is required to determine whether these findings—TOOT being fronter than BOOT, TOOK being more monophthongal than FOOT, and TOE being most similar to BOAT—generalize to other Southerners in this era, and how they correspond to production and perception in the present day.

4.4. Stability in Formant Curves

Finally, we describe how the trajectories themselves change in apparent time. A finding among all five vowels is that while their relative positions change across generations, there is remarkable consistency in the shapes of the curves across apparent time. For example, BOOT exhibits a consistent diphthongal pattern, with the offset backer and higher than the onset. Each generation has slight differences, but the general underlying shape is similar. This is especially true when comparing the three generations within each gender: men’s BOOT trajectories are all more dynamic than women’s are. The same tendency can be found among all five vowels where the trajectory length and shape remain approximately the same across time, suggesting that changes in these vowels are primarily shifts in position.

It is tempting to think that this stability in formant curves is an artifact of modeling, and that a more flexible statistical model is required. However, recall that in the model specification, each combination of gender, generation, vowel, and formant was fit independently; the best fit formant curves were not constrained by a particular location in the F1-F2 space nor were they forced to conform to a particular shape by gender or vowel. In other words, the predicted values for the Lost Generation women’s BOAT were not tethered to any other generation, gender, or allophone, and the same holds for all other levels of the four-way interaction term. Thus, the similarities among formant trajectories predicted for each vowel are not due to any bias or constraints in the model.

5. Discussion

This study analyzed formant trajectories from five back vowels in Southern American English, a speech variety with characteristically diphthongal vowels. Our data came from Linguistic Atlas interviews conducted as early as 1970 with generally older, rural individuals from across the American South, offering a unique look into an earlier stage of SAE. Though various methods of analyzing formant trajectories have been employed in previous studies, we modeled the trajectories directly by fitting the data with generalized additive mixed-effects models (or GAMMs). The results confirm the presence of back vowel fronting in the region, as well as the lack of the low back merger, and shed light on the relative timing of these changes. Though the vowels’
positions change in apparent time, there is relatively little change in their trajectory shapes.

This data confirms earlier findings that GOOSE began fronting before both allophones of GOAT did. Figures 4 and 5 suggest that GOOSE-fronting was a rigorous change in the South in the years when DASS speakers were acquiring language (approximately 1886-1964). We argue that women were on the forefront of this change because the position of BOOT (relative to PUT) was fronter than in men and because it approached the stable TOOT vowel. Meanwhile, there is relatively little evidence of BOAT-fronting among these speakers. The only indication of this shift in this data is that the youngest speakers had a fronter realization than did the G.I. Generation.

Like GOOSE, GOAT (which includes BOAT and TOE) was not stable over time: over the course of two generations—first among the G.I. Generation women and then among the Silent/Boomer Generation men—GOAT lowered in the vowel space. This confirms observations that GOAT lowered after 1900, possibly as a result of demographic shifts after the Civil War (Thomas 2004). However, while Thomas describes the nucleus of GOAT as the element that lowers (Thomas 2007; Thomas & Coggshall 2014), our data suggests that both the nucleus and the glide lower since there was not an appreciable change in the trajectory shape.

In our data, it appears that fronting of GOOSE and GOAT precedes the fronting of FOOT, in line with Fridland (2001). Additionally, our results suggest that, on the whole, FOOT changes little over time. Thus FOOT-fronting is not correlated with either GOOSE- or GOAT-fronting, as suggested by Thomas (2004). We note that in places like Alabama, the order of BVF proceeds from higher to lower, such that FOOT begins fronting after GOOSE but before GOAT (Feagin 2003). Our analysis does not evaluate potential subregional differences; we leave that for a more targeted investigation. Other recent work has subjected F2 data from DASS to clustering techniques and GIS mapping. The results suggest no significant regional patterning in the degree of GOOSE fronting (Jones & Renwick 2020). However, DASS speakers along the Florida Gulf Coast have a significantly fronter GOAT vowel than their neighbors; meanwhile, speakers in northwestern Georgia, Atlanta, and eastern Alabama exhibit GOAT and FOOT vowels that are significantly backed compared to others nearby (Jones & Renwick forthcoming).

The trajectories and positions of the low back vowels appear to coincide with what is described in the South. It does not come as a surprise that the two vowels are distinct here, since mergers of LOT and THOUGHT are reported primarily for Southern speakers younger than those in DASS (Irons 2007; Baranowski 2008; Stanley forthcoming). The analysis of trajectory data shows distinct trajectory shapes, supporting Thomas’ (2001) suggestion (based on data contemporaneous with our own) that the dynamic nature of these vowels kept them distinct. Furthermore, though we do not have data on lip rounding, the acoustic data for THOUGHT appears to match Thomas’ (2004) transcription of the vowel as [an]. As we reported in section 2.1, the low back merger has become widespread in the South, so we do not anticipate that one sub-region was shifting more rapidly than another. This speculation is supported by other, non-dynamic analyses of DASS data, showing that speakers throughout the eight represented states generally
maintain a LOT and THOUGHT distinction, with very little evidence of geographic patterning (Jones & Renwick 2020).

Our analysis reinforces existing transcriptions of Southern vowels while contributing additional phonetic detail about their trajectories. There is change in apparent time for BOOT, TOOK, BOAT, TOE, and THOUGHT, and stability in TOOT, PUT, and LOT. When comparing the relative positions of the shifting vowels to the stable vowels, which we assume occupy the same approximate position in perceptual space across genders, women appear to have led the changes.

The stability in the formant curves’ shapes in apparent time supports Koops’ (2010) proposed development of GOOSE-fronting in the South. As mentioned previously, Koops (2010) finds two distinct realizations of GOOSE in Houston Anglos: a fronted “southern” version that has relatively little spectral change ([y]), and a “mainstream” version with a fronted nucleus and a backer offglide ([iʌ]). Labov (2008) suggests that these two variants have a shared history, with an intermediate diphthongal stage [iɯ]. At that point, there was a split, and one variant lost its lip rounding while the other fronted the offglide to re-form a monophthong. Koops (2010) instead proposes that these two renditions of GOOSE were independent developments. The mainstream variant began as [u] and, over time, fronted only the nucleus while the glide remained in place, resulting in an increasingly diphthongal GOOSE with a greater drop in F2 over the vowel’s duration. The Southern variant also began as [u], but the nucleus and glide fronted in tandem, retaining the monophthongal quality over time. The stability in formant trajectories that we find in apparent time supports Koops’ (2010) proposed development of “southern” GOOSE, specifically by finding no evidence of the development and subsequent loss of an offglide.

6. Conclusion

Legacy audio corpora such as DASS allow researchers to test hypotheses about linguistic change with a greater time depth than is typically available from sociolinguistic fieldwork. Using a large new dataset, we have provided an analysis of back vowels’ formant dynamics in DASS, uncovering phonetic detail about their shape and change in apparent time. The vowels in GOOSE and GOAT both front over generational time, and the high back vowel is more advanced in this process, confirming its relatively earlier actuation. Phonological context, specifically the presence of a pre-vocalic coronal consonant, conditions fronting to some extent. The FOOT vowel is relatively more stable but does undergo mild backing in these materials. We have also shown clear distinctions between the low back vowels, suggesting that the merger of LOT and THOUGHT was not common when these speakers were acquiring language. All five vowels we study are shown to vary across time, progressing toward late twentieth-century descriptions that are more familiar. Our findings provide phonetic validation of existing results and support analyses from smaller datasets, with a larger corpus of forty-eight speakers and 290 hours of audio. This work relies on an automated workflow and complements, but does not supplant, careful exploration by ear and eye with a goal of detailed phonetic transcription or individual speaker analysis.
Although our dataset is a rich resource on SAE, we acknowledge that this may not fully capture the range of interspeaker variation, given that it includes only a subset of the population at the time it was collected. DASS speakers include several generations (born between 1886 and 1956) and a variety of different social dimensions (see section 3.1), providing a snapshot of the diversity present in SAE. While we stress the importance of vowel dynamics, we acknowledge that we work under the assumption that vowel-inherent spectral change is perceptually relevant and sociolinguistically meaningful. Some work has been done to support this assumption in the Southern US (e.g., Fridland & Kendall 2012; Farrington, Kendall & Fridland 2018; Gunter, Vaughn & Kendall 2020; Jacewicz & Fox 2020), though we leave such perceptual work on DASS for the future.

Finally, our results fill in several gaps in a description of SAE. First, to our knowledge, this study is the first analysis to focus on the dynamics of back vowels in the US South collectively, especially in the area covered by DASS. The second gap we fill is temporal: this study has uncovered detail about back vowel formant dynamics in Southerners born in the late nineteenth and early twentieth centuries. By filling in our knowledge of the phonetic past, we can better understand the phonetic present, and anticipate the phonetic future.

Acknowledgments
The authors thank Co-Editors Alexandra D’Arcy and Peter Grund and two anonymous reviewers for generous feedback on this manuscript. We are grateful to Keiko Bridwell for assistance with verifying accuracy of the forced alignment process. The analysis also benefited from helpful comments during the 93rd Annual Meeting of the Linguistic Society of America and from training via the Analysing Curves workshop hosted at the University of Manchester in 2018. DASS is provided by the Linguistic Atlas Project at the University of Georgia.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by NSF BCS Grant No. 1625680.

ORCID iDs
Joseph A. Stanley https://orcid.org/0000-0002-9185-0048
Margaret E. L. Renwick https://orcid.org/0000-0002-8759-1717

Supplemental Material
Supplemental material for this article is available online.
Notes

1. Throughout this paper we refer to vowels using Wells’ Lexical Sets (Wells 1982a) in order to unambiguously refer to vowel sounds without assuming phonetic quality or phonological representation of those sounds.

2. Demographic shifts are reported to be the catalyst for language change in other areas like Oklahoma (Bailey, Wikle, Tillery & Sand 1996), eastern Pennsylmania (Herold 1990), New England (Johnson 2010), and Washington (Stanley 2018). In the South, Thomas (2004) describes how the development of new industries (related to textiles, tobacco, timber, coal, and steel) led many White workers to flock to mill towns, spreading linguistic features that were once local to the whole region. Similarly, Bailey, Wikle, Tillery, and Sand (1996) describe how, in Oklahoma after World War II, many rural workers migrated towards urban areas resulting in a loss of traditional Southern features. While this second major demographic shift occurred during the lifetime of most of the speakers in DASS, many of these speakers lived in rural areas so such shifts may not have had as large an effect as it did on urban dwellers.

3. A reviewer points out that syllable structure may be an important factor in GOAT-fronting. Specifically, word-final or pre-pausal tokens of GOAT (go, no, so, though) front at a slower rate than other allophones in Northern England (Watt & Tillotson 2001; Jansen 2019) and pre-hiatus position (going, go out, so is) inhibits fronting in the American South (Thomas 2004). For simplicity in this study and to not overcomplicate an already complex statistical model, we do not include these potential factors. Tokens that contain GOAT in syllable-final or pre-hiatus positions were classified into allophones depending on their previous segment (e.g., go, snow, and knowing are BOAT and grow, show, and throwing are TOE).

4. This model does not include a parameter for autocorrelation, which can result when adjacent measurements are highly correlated with one another. Following Renwick and Stanley (2020), who argue that autocorrelation across measurement points does not affect GAMMs of DASS data, we do not construct an autocorrelation analysis here (see Sóskuthy [2017] for further discussion).

5. Trajectory length was found by sampling the model at eleven equidistant points, calculating the Euclidean distance between consecutive pairs of measurements, and summing the lengths of the ten inter-point segments. This was done for each combination of vowel, gender, and generation, and then averaged.

6. We acknowledge that our methods do not explicitly control for the physiological effects of aging on vowel acoustics (Reubold, Harrington & Kleber 2010; Reubold & Harrington 2015). Disentangling and interpreting the effects of aging and generational language change is challenging (Fruehwald 2017). We are encouraged, however, that our acoustic results match descriptions based on phonetic analysis by ear, such as Thomas and Coggshall’s (2014) finding that GOAT lowers over time.

Corpora and Software References


**References**


**Author Biographies**

**Joseph A. Stanley** is Assistant Professor in the Department of Linguistics at Brigham Young University. His research currently focuses on change in vowel formant dynamics in Southern and Western varieties of American English. He also works on quantitative methods, statistics, and data visualization in sociophonetic research.

**Margaret E. L. Renwick** is Associate Professor of Linguistics at the University of Georgia. She studies vowels and consonants, in English and the Romance languages, to understand how phonological contrasts are implemented phonetically. Her research projects are rooted in laboratory phonology, an approach which integrates experimental methods with more abstract linguistic representations of sound structure.

**Katherine Ireland Kuiper** is a PhD Candidate in Linguistics at the University of Georgia. She currently works as a Research Assistant for the Linguistic Atlas Project (LAP) and the Digital Humanities Lab; her research focuses on language variation, primarily in different contexts and genres of English over time.

**Rachel M. Olsen** is a PhD Candidate in Linguistics at the University of Georgia. In addition to exploring the spectral qualities of vowels in Southern speech, her work also examines prosodic cues to social identity and emotional affect.