DIFFERENCES IN FINAL /z/ REALIZATION IN SOUTHWEST AND NORTHERN VIRGINIA

RACHEL HARGRAVE
AMY SOUTHALL
ABBY WALKER
Virginia Polytechnic Institute and State University

ABSTRACT: Two apparently contradictory observations have been made about consonantal voicing in Southern U.S. English: compared to other U.S. varieties, Southern speakers produce more voicing on "voiced" stops, but they also "devoice" word-final /z/ at higher rates. In this article, regional differences in final /z/ realization within Virginia are investigated. Thirty-six students from Southwest and Northern Virginia were recorded completing tasks designed to elicit /z/-final tokens. Tokens were acoustically analyzed for duration and voicing; they were also automatically categorized as being [z] or [s] using an HTK forced aligner. At the surface level, the two approaches yield incompatible results: the single acoustic measures suggest that the Southwest Virginians produce more [z]-like /z/ tokens than the Northern Virginians, and the aligner finds that the Southern-identifying participants produce the most [s]-like tokens. However, both analyses converge on the importance of following environment: the Southwest Virginians are relatively least voiced prepausally and more voiced in other environments. These combined findings confirm previous work showing that Southern "voiced" consonants generally have more voicing than other regional U.S. varieties but also suggest that the dialect may exhibit greater phrase-final fortition. There are also differences within Southwest Virginian speakers based on differences in their rurality or in their orientation to the South.

KEYWORDS: laryngeal setting, fricative voicing, orientation to place, categorical coding

THE PURPOSE OF THE CURRENT STUDY is to find tangible evidence for the largely anecdotal observation that speakers of Southern U.S. English produce final /z/ differently than non-Southern speakers in a way that sounds more [s]-like. Specifically, we elicit and compare the acoustics of word-final /z/ tokens from speakers in Southwest (SWVA) and Northern (NOVA) Virginia and also look for more nuanced differences among Southwest Virginians based on their individual attitudes and identities.

While the process by which a /z/ would become more [s]-like is commonly referred to as devoicing (e.g., Saito 1994; Holmes 1996; José 2010; Bayley and Holland 2014), that description is controversial within Germanic pho-

American Speech, Vol. 97, No. 3, August 2022 DOI 10.1215/00031283-9308362

nology. The term *voicing* implies that a contrast derives from the presence or absence of periodicity in the wave form, a function of vocal fold stiffness. Researchers taking a laryngeal realism approach have argued that while many languages, like Spanish, have a true voicing contrast, English does not. Instead, the difference between the so-called voiced and voiceless stops in English is not [voice] but rather [spread glottis] (Iverson and Salmons 1995; Avery and Idsardi 2001), which creates increased airflow for aspirated phonemes.[1] Moreover, unlike in voicing languages, the fortis-lenis contrast does not pattern voiced-voiceless but rather aspirated-unaspirated, such that /z/ is lenis, not fortis. Such an account explains why English critically contrasts stops based on aspiration (Lisker and Abramson 1964), exhibits only variable periodicity on "voiced" consonants (e.g., Docherty 1992; Smith 1997), and also accounts for historical patterns of change (Salmons 2020).

Consistent with a [spread glottis] account, Smith (1997) finds that /s/ consistently has more airflow than /z/, and other work found that /s/ is typically longer in duration than /z/, either absolutely or relative to preceding vowel duration (e.g., Baum and Blumstein 1987; Crystal and House 1988; Docherty 1992; Stevens et al. 1992; Smith 1997; Purnell, Salmons, and Tepeli 2005). However, a laryngeal realism account still allows for voicing—the literal vibration of the vocal folds—to also consistently differ between aspirated and unaspirated consonants, as a secondary or enhancing cue to stop identity (Salmons 2020). For example, lenis, unaspirated stops (/b, d, g/) are often voiced, especially in assimilatory contexts, like between vowels, but fortis, aspirated stops (/p, t, k/) are not since the spread glottis inhibits voicing. Similarly, studies that have investigated periodicity in /z/ find that it can be marginally or inconsistently voiced, but critically, it is typically produced with more voicing than /s/ (e.g., Hughes and Halle 1956; Docherty 1992; Stevens et al. 1992; Smith 1997). Therefore, /z/ could become /s/-like through changes to a number of acoustic cues (e.g., Purnell, Salmons, and Tepeli 2005a), including, but not limited to, a decrease in glottal pulsing.

Interspeaker variation in word-final /z/ pronunciation has been attested in varieties of English around the world, including in England (Verhoeven, Hirson, and Basavaraj 2011), New Zealand (Holmes 1996), and North America (Smith 1997; Purnell, Salmons, and Tepeli 2005; Purnell 2005; Bayley and Holland 2014). In some cases, it appears that the phenomenon may be due to (historical) second-language transfer (Thompson 1977; Purnell, Salmons, and Tepeli 2005; Anderson and Davis 2013; Bayley and Holland 2014). Unsurprisingly, age is a factor in situations where phonation is changing in the community (Purnell et al. 2005), but gender (Holmes 1996; Verhoeven, Hirson, and Basavaraj 2011; Bayley and Holland 2014), ethnicity (Holmes 1996), and social class (Holmes 1996; see also Rohena-Madrazo 2015) have also been identified as predictors of /z/ realization. Importantly for our study, some researchers have also found effects of attitude and orientation on /z/ pronunciation. Thompson (1977) argues that /z/ devoicing in Mexican American males in Texas is less prevalent in speakers who thought that Spanish-accented speech would negatively impact job prospects, and Bayley and Holland (2014) found that the Texas Latinx in their study who oriented more toward their local community devoiced more than those who did not. Combined, these results show that fricative phonation can vary in socially meaningful ways (see also Docherty et al. 2011 and Grieser 2015 for examples of socially meaningful patterns in stop voicing).

Of most relevance to the current study, Hazen et al. (2015) informally noticed "/z/ devoicing" in their corpus of (mostly white) West Virginian speakers. They began acoustically investigating whether /z/ and /s/ were neutralizing in their speakers, but their results suggest that so far the phonological distinction is intact. Similarly, we had also noticed, in a listening experiment that asked participants to transcribe the recorded speech of White Southwest and Northern Virginians (Walker, Burlile, and Askew 2016), that the only transcription errors centered on final /z/ realization (i.e., typing <grace> when they heard *graze*) occurred when participants were listening to the Southwest Virginians and not the Northern Virginians, though such errors were infrequent.

This idea that Southerners may produce less voiced /z/ tokens contrasts with consistent data showing that U.S. Southern speakers produce more voicing during closure on /b, d, g/ than other U.S. regional varieties (Jacewicz, Fox, and Lyle 2009; Hunnicutt and Morris 2016; Walker 2020).[2] Given that the Southern Appalachian region in particular was largely settled by English speakers from the British Isles, with an arguably dominant influence from Scotch-Irish English (Montgomery 1997), this increased voicing could reflect influence from Scots, which has more stop voicing than other U.K. English dialects (Scobbie 2006; Docherty et al. 2011; Sonderegger et al. 2020). While this work has been almost exclusively done on stops—raising the possibility that fricatives might differ (cf. Iverson and Salmons 2003)—it is worth noting that producing the word *greasy* with a voiced fricative instead of a voiceless fricative has long been considered a shibboleth marker of Southern U.S. English (Kurath and McDavid 1961) and could be evidence of increased voicing in fricatives as well (though see Ching 1996).

In the study described below, we designed an experiment to elicit /z/-final tokens from participants from Virginia Tech who came from either Southwest Virginia or Northern Virginia. While in the same state, the areas are both culturally and linguistically distinct (Kurath and McDavid 1961; Labov, Ash, and Boberg 2006; University of Virginia Demographics Research Group 2022). The western border of the state includes the Appalachian Mountains and borders North Carolina, Tennessee, Kentucky, and West

Virginia. The state's northeastern border is culturally and economically intertwined with Washington, D.C., the political center of the United States. This makes comparisons between the two student populations a fruitful site for investigating Southern versus non-Southern features, or Appalachian versus non-Appalachian features.

The purpose of our investigation is to see whether we can establish a difference in the quality of final /z/ in these communities. Following our own observations, and those of Hazen et al. (2015), our primary hypothesis is that we would see less voicing on final /z/ from Southwest Virginians than from Northern Virginians, but given the arguments that Southern English produces more voicing on lenis consonants, we also entertain the idea that we might find more voicing on /z/ in the Southern English speakers. We approach this task with two separate analytical tools: automatic [s, z] coding using the Penn Forced Aligner (Yuan and Liberman 2008; Rosenfelder et al. 2011) and automatic acoustic analysis using Praat (Boersma and Weenink 2017). We are not seeking to argue about the phonemic status of /z/ as it relates to /s/—in other words, the experiment is not designed to investigate neutralization across the two phonemes (cf. Purnell, Salmons, and Tepeli 2005; Hazen et al. 2015)—and thus do not analyze /s/ realizations in our study. Rather, we seek converging evidence for differences in the quality of final /z/ across two different populations.

We also designed our study so that we could explore differences among Southwest Virgina speakers. Linguists have known for a long time that speakers in the same dialect region are far from homogenous (i.e., Labov 1972). Given the findings of Bayley and Holland (2014) regarding the role of orientation to community in "/z/ devoicing" and the large body of work showing that orientation toward place is a key factor in dialectal variation (e.g., Labov 1972; Hazen 2002; Carmichael 2014), including speakers from Appalachia (Hazen and Hamilton 2008; Greene 2010; Reed 2016), we included questions in our study that measured participants' orientation to place. This way the study not only investigates a Southern versus non-Southern difference but also examines whether there is evidence that /z/ quality indexes Southern speakers' more individual attitudes or orientation to the South.

METHODOLOGY

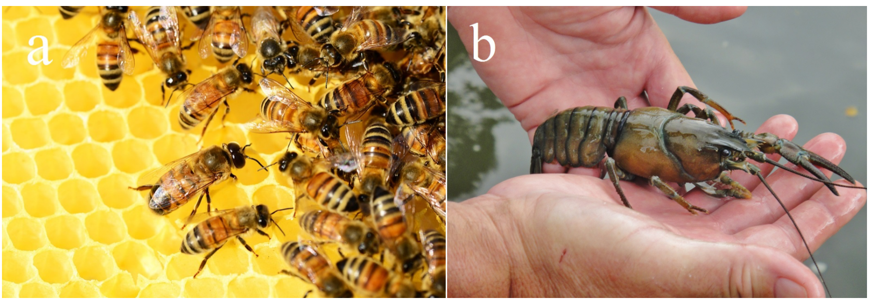
In order to investigate the differences in /z/ realization among Southern and Northern Virginians, we ran an experiment designed specifically to elicit final /z/ in both read and nonread speech. The session also included a questionnaire that asked about participants' orientation toward various aspects of place: rural versus urban, North versus South, and their hometowns.

EXPERIMENT DESIGN. We built the experiment in E-Prime 2.0 and recorded participants with an H-4 Zoom recorder (44.1KHz, 16 bit) attached to a Shure Beta 53 head-mounted microphone. The experiment consisted of four tasks, which participants all completed in the following order: picture-naming task, reading passages, questionnaire, and interview.

Picture-Naming Task. In the first part of the experiment, participants were shown 76 pictures in randomized order on the computer screen and were asked to say aloud what the picture represented. Fifty of the pictures were designed to elicit word-final /z/, like the picture of honey bees in figure 1a. The remaining 26 images either invoked Southern or Appalachian culture or were images that might elicit dialect-specific vocabulary, such as the crustacean in figure 1b, which most participants called either a crawfish, crawdad, or crayfish. The purpose of the filler images was threefold: (1) to detract from the variable of interest, final /z/; (2) to let participants think that our primary interest was in lexical variation; and (3) to maximize the chance of seeing a Southern feature in an experimental setting at a university, hoping that the Southern-related images would prime Southern U.S. English, especially from the Southwest Virginia participants (see Hay and Drager 2010; Walker 2019).

Picture-naming tasks have the benefit of largely eliciting set responses without asking participants to read text aloud, which can affect what variants they might use, especially if the variant under investigation is nonstandard (Labov 2001). In choosing the images, some pictures were intended to elicit monomorphemic final /z/ (i.e., *rose*, *pause*) as well as the inflectional, plural marker /-z/ (i.e., *bees*) since previous work suggested that "devoicing" rates may be influenced by differences in the morphemic status of /z/ (José 2010; Bayley and Holland 2014). All responses that contained a word-final /z/ were included in our analysis, whether it was the word we intended to elicit or not. The complete list and count of word types are given in appendix 1.

FIGURE 1
Sample Images Used in the Picture-Naming Task



Reading Passages. To analyze "/z/ devoicing" in read, connected speech, participants read two short passages (in appendix 2) that were thematically related to Southern culture, such as towns located in West Virginia and Virginia in passage 1 and mountains and trucks in passage 2. Again, the goal was to prime Southern identifiers on the South or on Southern culture, hoping that participants would be more likely to use Southern variants. Strategically placed words that ended in final /z/ tokens were placed throughout each passage, followed by a variety of phonetic environments.

Questionnaire. Still at the computer, each participant next completed a questionnaire that began with free response questions targeting basic demographic information: where they had lived; what they considered their hometown; their native language; other languages they speak; their age, gender, and ethnicity; whether they had speech or hearing problems; their major; linguistics experience; and critically for one of the ways we categorized participants in this study, the yes-no question "Do you identify as Southern?"

Participants were then presented with 83 first-person statements and were asked to indicate on a five-point Likert scale how much they agreed or did not agree with the statements. The statements targeted participants' general mood on the day ("I am in a good mood right now"), their orientation toward their hometown ("I have a lot of hometown pride"), their more general regional orientation ("I would like to live in the South for the rest of my life"), and their identification with Appalachia ("I am Appalachian"). The questions are provided in appendix 3.

Interview. For the last part of the experiment, participants completed a brief recorded interview with the experimenter (one of the first two authors), during which they were asked about their regional background and their experiences at college. The quantity of speech elicited using this method varied widely across participants, which influenced the decision to not analyze "/z/ devoicing" in the interviews. However, the interviews were used to verify where participants were actually from, since many of them responded on the online questionnaire with the closest large city to their smaller hometowns.

PARTICIPANTS. To recruit participants for this study, we placed flyers around Virginia Tech, offering a $10 voucher in compensation for participants' time. The flyers asked for participants from Appalachia, West Virginia, Southwest Virginia, and Northern Virginia. We chose to target students using the term *Appalachia* because Hazen and colleagues had framed their observations of "devoicing" in West Virginia as describing Appalachian English, and we wanted to include anyone who might identify with that label, from anywhere in Appalachia. However, although Southwest Virginia is considered part of

Central or South Central Appalachia (Appalachian Regional Council 2009), our experience is that students from Southwest Virginia often do not self-identify as Appalachian, and we were open to Appalachian-identity being a factor in /z/-realization. We included participants from Northern Virginia as a comparison to the Southern/Appalachian students, because, given that other studies have established that the majority of English speakers /z/-devoice a large amount of time (i.e., Smith 1997), it was important to establish that Southwest Virginians were not simply just "devoicing" but that they were devoicing MORE than Northern Virginians.

We ran a total of 39 participants, removing three from analysis they did not choose a hometown in Virginia. The map in figure 2 plots the 36 included participants' responses to the location they consider their hometown. The 13 participants to the left of the black line were categorized as coming from Southwest Virginia, and the 23 participants to the right were categorized as coming from Northern Virginia, with the majority of the latter group living in the greater D.C. region. Table 1 shows the age, gender, and ethnicity of participants based on these regional counts. While there is some ethnic diversity in the Northern Virginians, the group is predominantly White, and the Southwest Virginia group is entirely White.

We also created an alternative participant regional categorization based on whether participants self-identified as Southern. Specifically, in response to the short answer question "Do you identify as Southern?" 15 participants identified as Southern (11 female, 4 male), 3 as sorta Southern (they entered "sorta," "kinda," or "somewhat") (2 female, 1 male), and 18 did not

FIGURE 2
Destribution of Participants Based on Hometown and Southern-Identification (regions defined by the University of Virginia Demographics Research Group 2022)

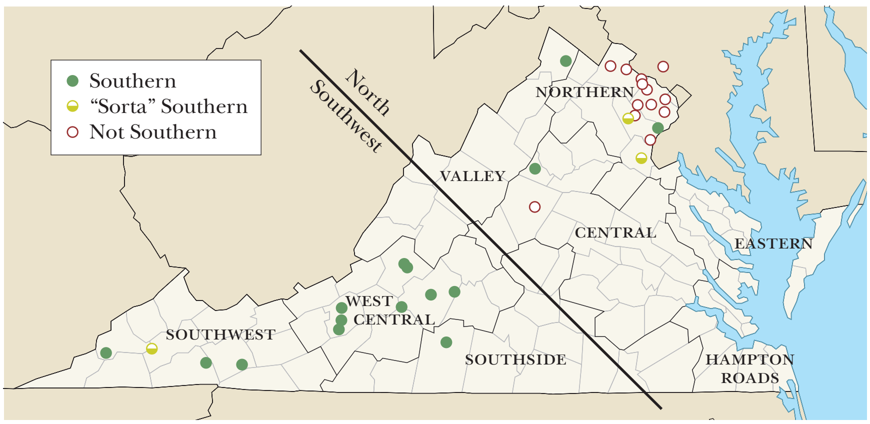


TABLE 1
Participant Demographic Information, by Hometown

	Southwest Virginia	*Northern Virginia*
Mean Age (range)	21 (19–18)	19 (18–21)
Gender	8 female, 5 male	20 female, 3 male
Ethnicity	13 White	14 White; 4 Asian; 3 multiracial; 1 Afghan; 1 not reported
Self-identified Southernness	12 yes; 1 sorta	18 no; 2 sorta; 3 yes
"I am Appalachian"	8 agree; 5 disagree	1 agree; 22 disagree
TOTAL	13	23

identify as Southern (15 female, 3 male). Because not all of the Southwest Virginian participants agreed with the statement "I am Appalachian" (unlike their unanimous Southern self-identification), we will refer to this group as Southern (not Appalachian) moving forward.

DATA ANALYSIS AND RESULTS

The 83 Likert scale questions from the questionnaire were grouped into categories by the authors (see appendix 3) and mean scores for these clusters were put through a factor analysis to see whether they reflected fewer latent variables (following Weatherholtz, Campbell-Kibler, and Jaeger 2014).[3] The output of the factor analysis motivated three latent variables, which we are calling RURAL-SOUTHERN (loading over ±0.5 for Have an Accent, Misunderstood, Southern, From Rural, Appalachian, Pro-Hometown, Southern Is Best, Future Rural, and negatively for Rich), BAD-MOOD-MISUNDERSTOOD (loading over ±0.5 for Misunderstood, and negatively for Good Mood and Mumbler), and PRO-SOUTH (loading over ±0.5 for Southern Is Best and Like Southern Accents). These three factors were tested (along with the remaining clusters that did not load over 0.5) as independent factors in the models.

Two methods were used to compare /z/ realization in the speech of the participants with the hope that they might provide converging evidence of a difference in Northern and Southwest Virginian speech. First, we acoustically analyzed final /z/ tokens in terms of two continuous measures: duration of the fricative and the amount of periodicity in the fricative. Second, we categorically coded for whether a token was an [s] or [z] using two methods—human coders and FAVE-align—though we focus in this article on the results of the aligner, as will be discussed later. For both continuous and categorical approaches, we only included tokens of /z/ in word final position. Function words were excluded from analysis in part because they were much more common in the passages than the picture-naming task.[4]

A total of 2,403 tokens are included in the analyses below. The original number was 2,432, but 29 were removed for having durations greater than 2.5 standard deviations (69 ms) above the original mean (153 ms).[5] Table 2 shows the distribution of tokens across following environments and tasks. In the picture-naming task, the majority of tokens were produced in isolation and were IP final, which is why the large majority of them have pauses following; in this way, task and following environment are largely conflated. Across tasks, the majority of /z/ were preceded by vowels or approximants (66%) and were in stressed syllables (74%).

For all analyses, we use mixed-effects linear or logistic models. Multiple linguistic variables were included as independent factors in the analysis: the morphological status of the /z/ (monomorphemic vs. inflectional), the preceding environment (vowel vs. obstruent/nasal vs. approximant), the stress of the syllable (stressed or unstressed), and the following environment (vowel vs. voiced consonant vs. voiceless consonant vs. pause). Social factors tested included speaker region (Southwest vs. Northern Virginia) and speaker self-identification (Southern vs. non-Southern).[6] For the models where we looked at Southwest Virginians only, we also explored their responses to the Likert scale questionnaire (their PRO-SOUTH scores, etc.). Finally, we tested social factors in interaction with the following environment and the task. Continuous measures were all centered, and categorical comparisons were all sum coded. Random factors typically include intercepts for participant and word with random slopes on the latter for social factors; though, sometimes this was not possible for convergence reasons. We started with full models and used a log-likelihood test to compare models in a step-down fashion; in the article, we present the best-fit models, which contain predictors that significantly improve on models that do not contain these predictors.

ACOUSTIC ANALYSES. We transcribed recordings and put them through Penn FAVE alignment (Rosenfelder et al. 2011; described in more detail below). For our analysis of fricative and periodicty duration, we started with the time-

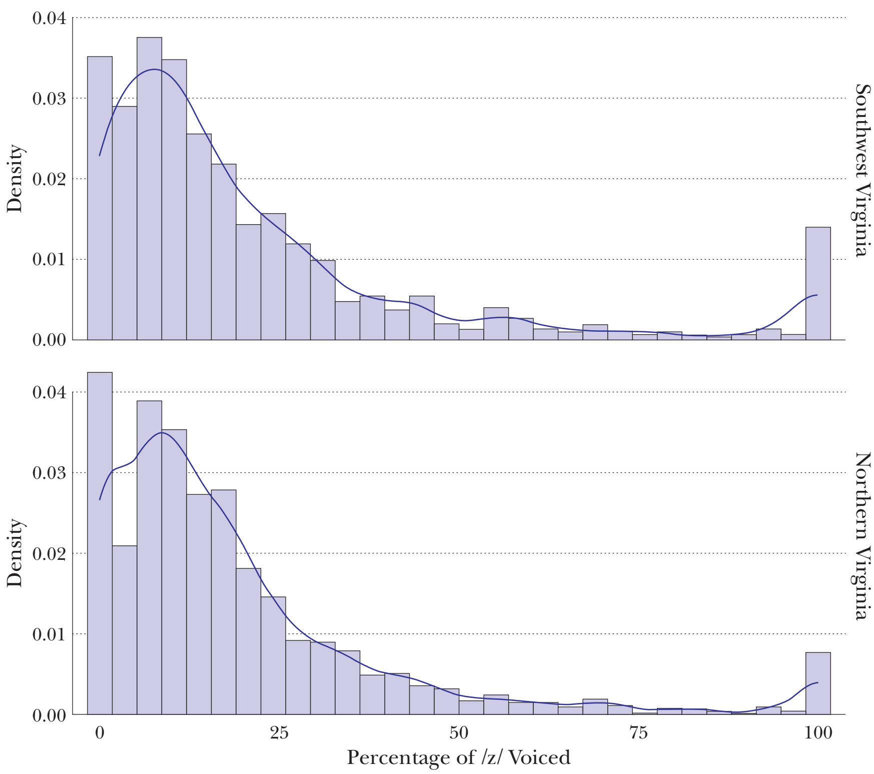
TABLE 2
Distribution of Tokens by Following Environment in Picture-Naming Task and Reading Passages

Following Environment	*Picture-Naming Task*	*Reading Passages*
Pause	1,446 (92.5%)	219 (26.1%)
Voiceless consonent	32 (2.0%)	170 (20.2%)
Voiced consonent	56 (3.6%)	122 (14.5%)
Vowel	29 (1.9%)	329 (39.2%)
TOTAL	1,563	840

aligned output of the forced aligner as our base and hand-corrected [s] and [z] boundaries by visually inspecting the start-end of frication-aperiodicity in the wave form as a guide. We then extracted duration and voicing measures. For the voicing measures, we set the window size to 0.001 seconds and told Praat to look for periodicity between 75–300 Hz, then extracted voice reports (VR) for each fricative to calculate a PERCENTAGE VOICED score.[7] Gradoville (2011) and Eager (2015) report that this measure correlates well with impressionistic and manual coding of voicing, respectively, and other studies have used this measure to capture periodicity in consonants (e.g., Davidson 2016).

Voicing Analysis. Figure 3 shows the distribution of voicing percentages across dialectal background. The mean PERCENTAGE VOICED for the Southwest Virginians was 22%; for the Northern Virginians, it was 19%. Following Davidson (2016), we categorized tokens with less than 10% voicing as devoiced,

FIGURE 3
Distribution of PERCENTAGE VOICED Scores by Speaker Dialect Region



tokens with greater than 90% voicing as fully voiced, and tokens in between as partially voiced. In both dialects, the majority of tokens (SWVA 55%; NOVA 57%) were partially voiced with very few fully voiced tokens (SWVA 8%; NOVA 3%). There were almost no fully voiced tokens following a pause for either group (<1%), which is consistent with previous findings (Docherty 1992; Smith 1997).

We initially did an analysis looking for a categorical difference in voicing across the dialects, contrasting devoiced tokens with tokens with some voicing (i.e., partial and fully voiced). Speaker region (or speaker self-identification as Southern) did not significantly improve the model on its own or in interactions. Therefore, there is no statistical evidence that one region produces more fully devoiced tokens than the other region.

We then subset the data to only include tokens with some voicing (i.e., partially and fully voiced tokens) so that we could investigate whether there were any differences in how much voicing there was in voiced tokens. The best-fit model of this approach (table 3) includes a main effect of preceding environment (more voicing following vowels/approximants vs. obstruent/nasals) and an interaction between following environment and speaker region. There is more voicing when voiced elements follow, and Southwest Virginians generally voice more than Northern Virginians, but this effect is strongest when voiced consonants follow and muted when a pause follows (figure 4).

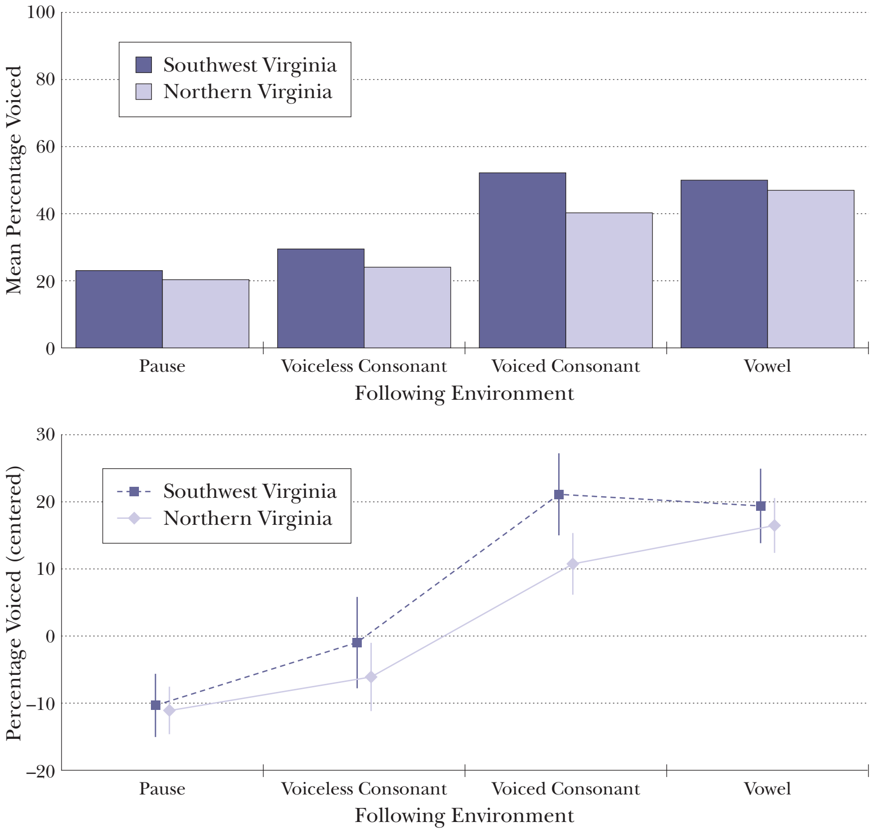
An arguably corresponding pattern is seen when we subset the data and look for an interaction between PERCENTAGE VOICED and individual differ-

TABLE 3
Model Coefficients for Best-Fit lmer for PERCENTAGE VOICED in Partially (10–90%) + Fully (>90%) Voiced /z/ Tokens

	Estimate	*Std. Error*	*t-Value*
(Intercept)	3.7	1.5	2.4
SouthernMAP = SWVA	2.4	1.5	1.6
Following = Pause	–15.6	0.84	–18.6
Following = Vowel	13.0	1.05	12.4
Following = Voiced consonant	11.0	1.21	9.1
PrecedingBINARY = Obstruent/nasal	–2.3	0.60	–3.84
SouthernMAP × Following = Pause	–2.0	0.83	–2.42
SouthernMAP × Following = Vowel	–0.9	1.05	–0.90
SouthernMAP × Following = Voiced C	2.8	1.21	2.30

NOTE: 36 Speakers; 1,451 tokens. All categorical factors were sum coded. Model included a random intercept for speaker.

FIGURE 4
Mean PERCENTAGE VOICED in Partially (10–90%) + Fully (>90%) Voiced /z/ Tokens by Following Environment and Speaker Region



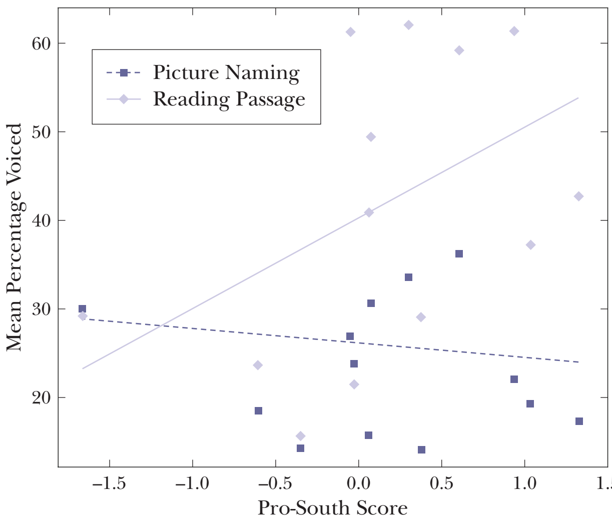
ences among Southwest Virginians. Specifically, the best-fit model (table 4) finds an interaction between pro-South scores and task: more pro-South Southwest Virginians produce more voicing than less pro-South Southwest Virginians in the reading passages but not in the picture-naming task.[8] In figure 5, each speaker is represented by a square (for the picture naming task) and a diamond (for the reading passage), and this effect appears to be driven by the fact that for the three speakers with pro-South scores under 0, they have very little voicing differences across the two tasks. In contrast, speakers with pro-South scores above 0 have substantially more voicing on /z/ in the reading passage versus the picture naming task. We interpret this interaction as similar to the interaction we saw in the larger model between speaker region and following environment, since tokens followed by pauses and tokens produced in the picture-naming task are more likely to be IP

TABLE 4
Model Coefficients for Best-Fit lmer for PERCENTAGE VOICED in Partially (10–90%) + Fully (>90%) Voiced /z/ Tokens for Southwest Virginia Speakers Only

	Estimate	*Std. Error*	*t-Value*
(Intercept)	3.7	2.91	1.27
ProSouth Score	5.0	3.67	1.35
Following = Pause	–16.7	1.97	–8.51
Following = Vowel	13.2	2.03	6.49
Following = Voiced consonant	12.2	2.20	5.53
Morpheme = Mono	2.27	1.05	2.17
PrecedingBINARY = Obstruent/nasal	–2.32	1.11	–2.09
Task = PictureNaming	0.4	1.37	0.27
ProSouth × Task = PictureNaming	–5.71	1.27	4.49

NOTE: 13 speakers; 517 tokens. All categorical factors were sum coded, and PRO-SOUTH score was centered. Model included a random intercept for speaker.

FIGURE 5
Mean PERCENTAGE VOICED in Partially (10–90%) + Fully (>90%) Voiced /z/ Tokens by Task and PRO-SOUTH Score for Southwest Virginia Speakers Only



final. That is, both findings suggest that Southwest Virginians and more pro-South Southwest Virginians produce more voicing but such interspeaker differences disappear phrase-finally.

This model also replicates the larger model in terms of main effects of following and preceding environment. It additionally includes an effect of morphological status, such that inflectional /z/ is produced with less voicing, which matches previous work looking at /z/-realization and morphemic status (José 2010; Bayley and Holland 2014; cf. Saito 1994).

Duration Analysis. Because the relative duration of a fricative and the preceding vowel appears to be critical for distinguishing /s/ and /z/ in production (e.g., Purnell, Salmons, and Tepeli 2005; Purnell et al. 2005) and perception (e.g., Raphael 1972; Soli, 1982), we made RELATIVE DURATION the dependent variable in our analysis, that is, the duration of the fricative divided by the duration of the preceding vowel. Since some of our /z/ tokens followed consonants, thus complicating any calculation of vowel duration for our duration analysis, we analyzed only the subset of tokens that were directly preceded by a vowel or the approximants /l/ and /r/, in which case we calculated preceding vowel duration by adding the vowel and approximant durations together (see Smith 1997). We also excluded tokens with a relative duration of greater than 2 (losing 15 tokens). These exclusions left 1,550 tokens for duration analysis with a mean duration of 158 ms and a mean relative duration of 71%.

Table 5 shows the mean relative durations of /z/ by speaker region and following environment, as well as the mean preceding vowel and /z/ duration. The Southwest Virginians have slightly longer vowel and fricative duration than the Northern Virginians but a slightly shorter relative /z/ duration. However, none of these differences are significant in *t*-tests or mixed-effects regression modeling, even if we limit our analysis to prepausal tokens.

To investigate whether there might be differences among the Southwest Virginians depending on individual differences within the group, we

TABLE 5
Mean Duration Measures of /z/ by Speaker Region and Following Environment

Following Environment	*/z/ (ms)*		*Preceding Vowel (ms)*		*Relative Duration*	
	Mean	*sd*	*Mean*	*sd*	*Mean*	*sd*
Southwest Virginia						
Pause	191.30	(47.15)	282.80	(92.03)	75%	(31%)
Voiceless consonant	98.17	(46.67)	162.38	(80.59)	73%	(41%)
Voiced consonant	76.12	(20.78)	195.20	(99.71)	50%	(26%)
Vowel	77.81	(27.28)	160.50	(70.62)	56%	(25%)
Northern Virginia						
Pause	185.77	(43.55)	272.00	(90.60)	76%	(30%)
Voiceless consonant	92.44	(30.48)	143.82	(57.86)	73%	(32%)
Voiced consonant	75.60	(21.13)	185.65	(90.06)	53%	(32%)
Vowel	69.74	(22.28)	142.80	(73.79)	59%	(30%)

TABLE 6
Model Coefficients for Best-Fit lmer for (Centered) RELATIVE DURATION for Southwest Virginia Speakers Only

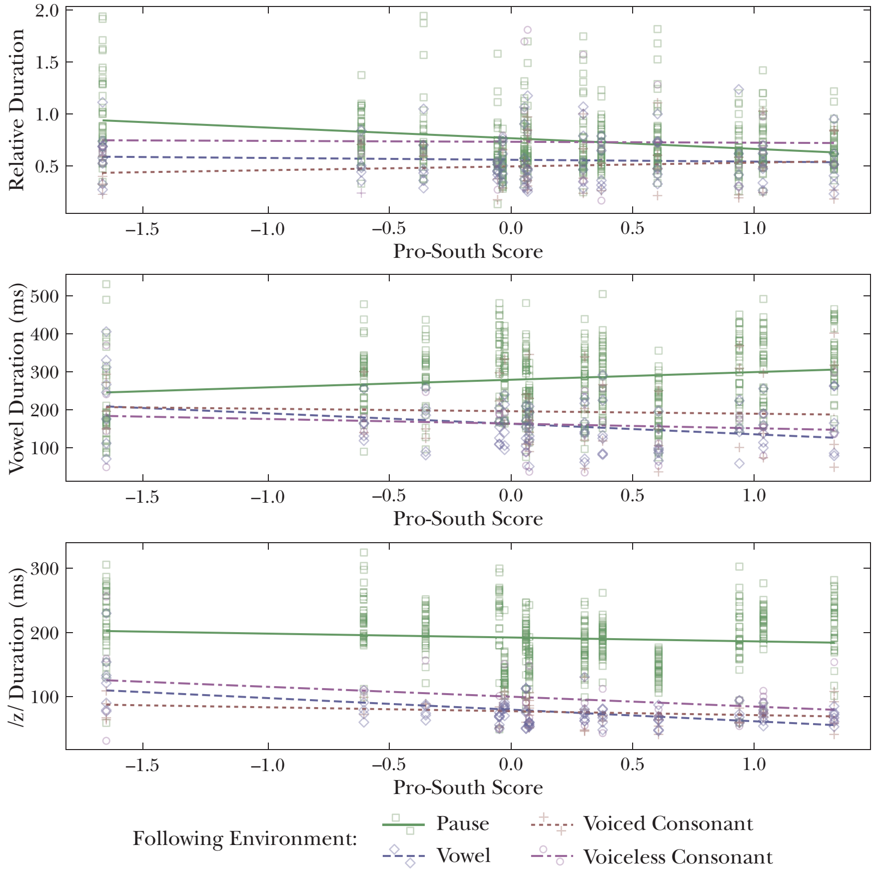
	Estimate	*Std. Error*	*t–Value*
(Intercept)	0.01447	0.03312	0.437
ProSouth Score	–0.02409	0.03162	–0.762
Following = Pause	0.14045	0.02203	6.375
Following = Vowel	–0.03405	0.03834	–0.888
Following = Voiced consonant	–0.09051	0.03934	–2.301
Stress = yes	–0.19305	0.02175	–8.874
ProSouth × Following = Pause	–0.07327	0.02163	–3.388
ProSouth × Following = Vowel	–0.02569	0.03136	–0.819
ProSouth × Following = Voiced C	0.07393	0.03687	2.005

NOTE: 13 speakers; 547 token. All categorical factors were sum coded, and PRO-SOUTH score was centered. Model included a random intercept for speaker.

created a subset of the duration data limited to Southwest Virginia group. The best-fit model of RELATIVE DURATION (table 6) shows that fricatives are relatively shorter on stressed syllables (presumably reflecting increased duration of a vowel), significantly longer prepausally, and shorter before voiced consonants. While there is not a main effect of Southern identity attributes, there is an interaction of a participants' PRO-SOUTH score and following environment: participants with higher scores do not have longer relative durations before pauses and have longer relative durations before voiced consonants—essentially, differences based on following environment are neutralized (see figure 6a). Based on the literature, this would be interpreted as having more canonically produced [z]; though, since the effect is being primarily driven by lengthening of preceding vowel duration (compare figure 6b and 6c), it is unclear whether it is fundamentally about the fricative or the vowel duration.

In summary, the duration results suggest that more pro-South Southwest Virginians produce more [z]-like /z/s than less pro-South Southwest Virginians, but there are other interpretations of the necessarily relative durational measures. There was no evidence (ambiguous or not) in differences in /z/ duration between the Southwest and Northern Virginia groups. Taken together, this could suggest limited difference in /z/ realization across the speaker groups, but it may be more indicative of the limited role that duration plays in the perception of [s]-like /z/s (Baum and Blumstein 1987; Purnell, Raimy, and Salmons 2009)

FIGURE 6
Duration of /z/ and Preceding Vowel by Following Environment for Southwest Virginia Speakers Only



NOTE: RELATIVE DURATION (*top row*) is calculated by dividing the duration of the fricative (*bottom row*) by the duration of the preceding vowel (or duration of vowel + /l/ or /r/) (*middle row*).

CATEGORICAL ANALYSIS. In addition to our acoustic analyses, we also categorically coded each token as being either [s] or [z], primarily because it is clear from the literature that multiple cues are used to distinguish laryngeal pairs (e.g., Lisker 1984), and we wanted to see what a more holistic approach would reveal that our analyses of voicing and duration might miss. Initially, the first two authors auditorily coded the data (which has the added bonus of demonstrating how audible any acoustic differences are). However, they produced only 75% agreement ($\kappa = 0.49$, considered "moderate agreement";

Landis and Koch 1977, 165), and subjectively, we lacked confidence in our choices and that the task was inherently difficult.[9] This difficulty is in itself interesting and raises questions about how perceptible variation on final /z/ is in these communities. But ultimately the difficulty in auditory coding motivated us to try automatically coding the data using FAVE-align (Rosenfelder et al. 2011).

FAVE-align automatically segments a sound file at the phonemic level and is based on the Penn Phonetics Lab Aligner developed by Yuan and Liberman (2008). The standard aligner was trained on phonemically transcribed speech from the Supreme Court Justices of the United States. We used a version of the program that was locally installed on our lab's hard drive. To use the aligner, one inputs a sound file and a corresponding orthographic transcript of the recording that has been time-aligned into chunks no longer than 3 minutes (in our data, it was divided roughly every 30 seconds). The aligner looks up the words in the Carnegie Mellon University (CMU) Pronouncing Dictionary and returns a TextGrid segmenting the sound file at the phoneme level. Part of the built-in functionality of FAVE-align is that words can have more than one entry in the phonemic dictionary, since homographs are spelled the same but pronounced differently. When faced with such words in the transcription, the forced aligner chooses a pronunciation based on which dictionary entry better matches the spectral information in the sound wave according to its training data.

While the primary purpose of FAVE-align is automatic alignment for acoustic analysis, a few researchers have demonstrated how you can use the homographic flexibility in the dictionary to use FAVE-align to automatically code categorical variables. Yuan and Liberman (2011) showed that this could be done in coding (ING) variation by explicitly training the aligner on (ING) variation. More recently Walker and Sedarous (2015) and Bailey (2016) showed that the aligner could categorically code variables without training, finding relatively high aligner-human agreement rates for the BATH-TRAP split, rhoticity, *th*-fronting, and *h*-dropping, in both British and American English.

To use the aligner to code for "/z/ devoicing" in our study, we simply added extra entries for target words that only differed by the final phoneme: the alternative entry ended in /s/ instead of /z/. For example, the default entry for the word *keys* in the dictionary is K IY1 Z, and we added another entry pronounced K IY1 S. This process gave the aligner a choice: did it think the part of the sound wave it was segmenting better matched Z or S based on its training data?[10] It is important to note that while we were motivated to explore using a forced aligner after difficulties in auditory coding, we do not assume that FAVE-align is telling us anything about how people would hear these variants necessarily.[11] FAVE-align and auditory coding share that they are holistic and categorical descriptions of /z/ realization, but the aligner is still an acoustic measure.

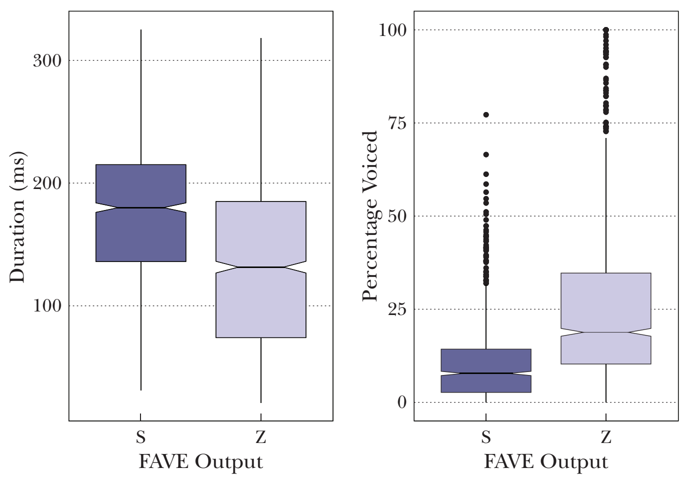
FAVE-Align Results. The FAVE-align coded 50% of the /z/ tokens in the picture-naming task as S and 31% in the reading passages; it also coded more S for prepausal tokens than prevoicing tokens (table 7). This reading passage versus picture naming pattern is similar to the main effects observed in the acoustic measures; in fact, in a simple logistic regression model predicting Z versus S coding by FAVE-align, both duration and voicing were strong predictors of category assignment (see figure 7). These facts demonstrate that the aligner is not randomly assigning codes to tokens.

Critically, we were interested in whether there were differences in S versus Z categorization depending on the speaker's regional background. Table 7 shows percentage S categorization depending on following environ-

table 7
Percentages of Tokens Coded as S by FAVE-Align by Speaker Region and Following Environment

Following Environment	*Southwest Virginia*	*Northern Virginia*
Pause	60.5%	48.6%
Voiceless consonant	45.3%	41.2%
Vowel	21.5%	16.0%
Voiced consonant	11.7%	16.8%

figure 7
Relationship between FAVE-Align Code (S or Z) of /z/ Tokens and Duration of Token and Percentage Voiced of Token



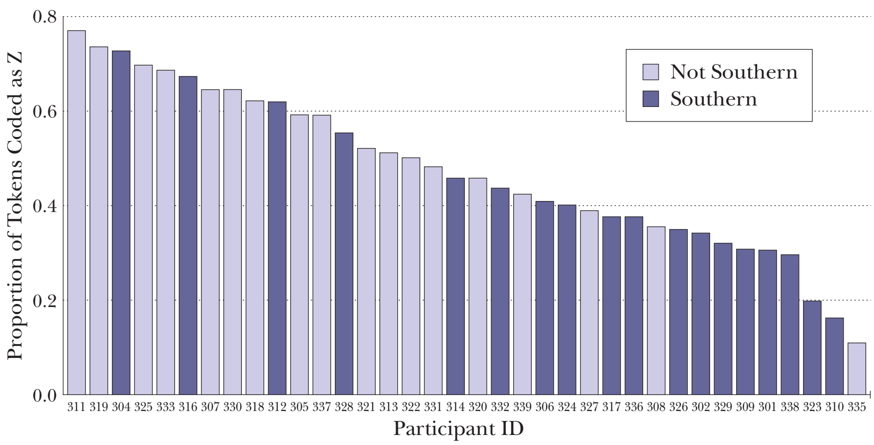
ment by speaker region and suggests that Southwest Virginians produce more [s] tokens than Northern Virginians but only substantially more when pauses follow. While including participant region as a predictor does not significantly improve a glmer model, if we instead look at the impact of Southern self-identification (responses to the question "Do you identify as Southern?"), there is a significant interaction between following environment and Southernness (SouthernID in table 8)—Southerners produce more [s] tokens than non-Southerners. Figure 8 visually demonstrates this effect by

TABLE 8
Model Coefficients for Best-Fit lmer for FAVE-Align Code (S or Z)

	Estimate	*Std. Error*	*z-Value*	*p-Value*
(Intercept)	1.01	.13	7.84	< .001
Morpheme = Mono	.41	.07	6.24	< .001
PrecedingBINARY = Obstruent/nasal	–.52	.05	–10	< .001
SouthernID	–.09	.13	–0.7	.49
Following = Pause	–1.06	.09	–11.6	< .001
Following = Vowel	.83	.13	6.29	< .001
Following = Voiced C	.87	.17	5.04	< .001
SouthernID × Following = Pause	–.23	.09	–2.6	.01
SouthernID × Following = Vowel	–.20	.13	–1.6	.12
SouthernID × Following = Voiced C	.29	.17	1.7	.09

NOTE: 36 speakers; 2,403 token. All categorical factors were sum coded. Model included a random intercept for speaker.

FIGURE 8
Proportion of Prepausal Tokens Coded as Z by FAVE-Align
by Speaker Self-Identification as Southern



ordering all speakers by proportion of /z/ tokens prepausally—Southern-identifying participants dominate the tail end of the graph.

Moreover, if we look for variation just within participants who come from Southwest Virginia, there is a significant interaction between following environment and the RURAL-SOUTHERN score (table 9). The higher a speaker's score (more rural their hometown), the more S that FAVEassigned to their tokens in pre-pausal environments (figure 9). In other environments, RURAL-SOUTHERN score does not impact rates of Z/S, and trends that there is less S in speakers with higher RURAL-SOUTHERN scores, especially before-voiceless consonants.

DISCUSSION

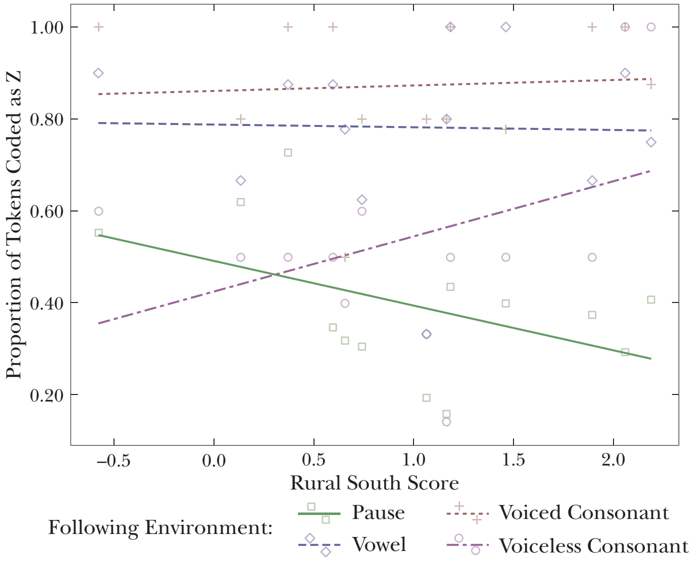
In the introduction, we outlined anecdotal evidence suggesting that Southerners may be "devoicing" final /z/, which appears to contradict other descriptions of Southern American English as having relatively more consonantal voicing than other U.S. varieties. Our study was designed to compare word-final /z/ realizations in the speech of speakers from Northern and Southwest Virginia, and we find evidence for both claims, depending on (1) the analysis we use, (2) whether we look across—or within—a dialect group, and (3) the interaction with the following environment or task.

TABLE 9
Model Coefficients for Best-Fit lmer for FAVE-Align Code (S or Z) for Southwest Virginia Speakers Only

	Estimate	*Std. Error*	*z-Value*	*p-Value*
(Intercept)	1.01	.13	7.84	<.001
RuralSouthern	0.03	.28	0.09	.93
Following = Pause	–1.21	.15	–7.91	<.001
Following = Vowel	0.68	.21	3.19	.001
Following = Voiced C	1.12	.30	3.73	<.001
Morpheme = Mono	0.52	.12	–4.49	<.001
PrecedingBINARY = Obstruent/Nasal	–0.40	.19	–2.34	<.001
RuralSouthern × Following = Pause	–0.46	.19	–2.34	.02
RuralSouthern × Following = Vowel	–0.10	.29	–0.34	.73
RuralSouthern × Following = Voiced C	0.04	.40	0.11	.92

NOTE: 13 speakers; 850 token. All categorical factors were sum coded. Model included a random intercept for speaker.

FIGURE 9
Proportion of /z/ Coded with Z by FAVE-Align by Following Environment and RURAL-SOUTHERN Score for Southwest Virginian Speakers Only



Specifically, in terms of voicing during the fricative, we find evidence that Southwest Virginia participants produce more voicing than Northern Virginian participants (table 3, figure 4) and that among the Southwest Virginians more pro-South participants produce more voicing than those who were less pro-South (table 4, figure 5); though, the difference disappears when pauses follow the /z/. We find a similar effect of PRO-SOUTH score on relative duration of /z/, such that we find relatively shorter /z/ (less consistent with hearing it as [s]) in more pro-South Southwest Virginians (table 6, figure 6). However, FAVE-align (table 8, figure 8) categorizes more tokens as S when they come from Southern-identifying speakers (driven by prepausal tokens), and we find a similar within-group effect of RURAL-SOUTHERN score (table 9, figure 9). While the two types of analyses initially appear to be at odds—our acoustic measures say Southwest Virginians voice more, automatic categorical coding says Southern-identifying speakers voice less—there is a fundamental parallelism in the interactions: Southerners voice least in prepausal environments (significantly so in the FAVE-align analysis) and voice more in other environments (significantly so on the acoustic measures and in the FAVE-align analysis).

Our nonprepausal results add to the growing literature suggesting that there is more voicing in Southern American English compared to other U.S. varieties (Jacewicz, Fox and Lyle 2009; Walker 2020). Critically, we believe this is the first study to show that this is not just true in stop consonants but appears to be true for fricatives, as well. There is some debate about whether this increased voicing is phonological (Hunnicutt and Morris 2016) or phonetic (Salmons 2020), and we think that our results—showing largely small differences across speakers and variation within speakers—probably support the latter interpretation.

The second finding that Southerners produce more [s]-like tokens in prepausal environments is mostly supported by the results using FAVE-align. As discussed earlier, this is a relatively novel way to code data, and given its contrasting results with the acoustic measures, it could be tempting to dismiss it. There are two reasons why we do not do so. First, /s/ and /z/ categorization by the FAVE-align actually correlate with the acoustic measures in the expected directions, and the relationship between /s/ and /z/ categorization and other factors (following environment, morphological status, etc.) match our acoustic analyses and the existing literature. Second, it is not clear that duration and glottal pulsing are in fact the most useful cues to /s/ and /z/ identity (e.g., Baum and Blumstein 1987; Smith 1997), and even in our own data, we can see that /z/ tokens are mostly produced with little to no glottal pulsing (figure 3). This already suggests that other cues better distinguish these sounds, and a forced aligner, which uses a more holistic measure of the spectrum, may actually be better at capturing that than the two acoustic measures we explicitly explored.

We think it is meaningful that the FAVE-align is predicting that our Southern-identifying participants produce more [s]-like tokens than other speakers only in prepausal positions. The large majority of prepausal tokens were IP final, and this is a phonetically special position. Indeed, Blevins (2006: 141) argues that devoicing must always start at the right edge, over larger domains, before it can be observed in smaller units (i.e., phrase finally before word finally, word finally before syllable finally). This pattern is consistent with work looking at /z/ realization in other dialects of English (e.g, Smith 1997), and Iverson and Salmons (2011, 1637) argue that the pattern should hold whether the neutralization is due to feature loss (i.e., the loss of [+voice]) or feature addition (i.e., the addition of [+spread glottis]). On that note, our stimuli were often single word IPs, and impressionistically, tokens that we heard as clearly "devoiced" (hearing [kis] instead of [kiz] for *keys*) did not seem reduced but rather strengthened. Therefore, consistent with a laryngeal realism view where /z/ could merge with /s/ by adding [spread] (not losing [voice]), we interpret what is happening as word/phrase-final fortition (Purnell, Salmons, and Tepeli 2005).

The differing accounts of Southern consonantal voicing then may largely reflect two independent processes—greater voicing in general but more fortition phrase finally than some U.S. English dialects. This finding highlights a number of potential directions for future work on laryngeal settings in U.S. Southern English, including, but not limited to, the history and progression of either phenomenon and an investigation of fricatives word initially and stops phrase finally.[12] Moreover, while our sample of Southwest Virginians was entirely White, racial/ethnic variation in word-initial voicing (Ryalls, Zipprer, and Baldauff 1997; Herd 2020) and phrase-final fortition (see Grieser 2015, 85–89, for a review) highlight the need to consider more diverse Southern populations in any diachronic and synchronic investigations.

Relatedly, we note that we find differences across our speakers not just in terms of their regional background, but critically on their attitudes, identity, and rural-urbanness. Specifically, we saw an impact of Southern participants' PRO-SOUTH score on voicing and duration and of a participant's identity (Southern vs. non-Southern) and RURAL-SOUTHERN score on the FAVE-align's coding of S versus Z tokens. A number of studies have found that the connection a speaker has with their home place or adopted home place is directly related to their rates of using the linguistic features of those places (see Labov 1972; Hazen 2002; Carmichael 2014; Reed 2016). Panregional variation then enables speakers to mark place identity in ways that allow for social and subregional mobility. Our study appears to support these patterns and shows the ways in which realizing fricatives specifically can be a marker of place-related attitudes and identities (Thompson 1977; Bayley and Holland 2014). The RURAL-SOUTHERN effect also highlights the issues with coding for regional dialect based on broad boundaries where critical differences between rural and urban experiences are lost; it is worth emphasizing that it was the RURAL-SOUTHERN factor that included statements about accentedness. That is, the FAVE-align is coding more S in people from Southwest Virginia who report having growing up around Southern accents.

What strikes us as especially interesting about these social, and possibly stylistic (see Walker 2020), patterns in /z/ realization is that the variation not only appears to be below the level of consciousness but that it was difficult for us to reliably hear. Clearly, more work on the perception of laryngeal variation is warranted—both in terms of discrimination/identification tasks (i.e., Can listeners reliably distinguish different variants?) and in speaker perception tasks (i.e., What do more [s]-/[z]-like variants socially mean?). If, as we suspect, speakers are producing differences that it is difficult even for them to hear, then one possibility is that the acoustic differences are the by-product of a fundamentally articulatory target (see Pratt 2018).

CONCLUSION

Both automated categorical coding and automated analysis of voicing suggest that there are differences in the ways that Southwest and Northern Virginians produce their word-final /z/. We argue that taken together the results suggest that generally Southwest Virginians speakers voice their /z/ fricatives more than Northern Virginians speakers, consistent with work looking at Southern stops, but in prepausal positions, Southwest Virginians are producing more [s]-like /z/ at greater rates than Northern Virginians. Looking within the Southwest Virginian group, we find individual differences based on attitudes to the South, rurality of hometown, and whether they identify as Southern or not, highlighting how assigning speakers "dialects" based on broad regions misses the nuances in the relationship between pronunciation and place.

APPENDIX 1

Words Analyzed and Counts across Tasks

Word	*RP*	*Pics*	*Word*	*RP*	*Pics*	*Word*	*RP*	*Pics*
airplanes	0	19	charles's	70	0	fries	0	37
almonds	0	1	charlie's	2	0	games	0	38
animals	0	1	cheese	0	72	gauze	0	1
arranges	0	1	cheeseburgers	74	0	gears	73	0
avengers	0	27	cheeses	0	4	girls	0	2
babies	0	17	chickens	0	27	glasses	0	35
bagels	0	31	chromosomes	0	11	grandfather's	37	0
balls	0	20	claus	0	41	hands	74	0
banjos	0	6	clothes	0	33	hens	0	3
beers	0	1	clovers	0	1	hose	0	40
bees	0	36	colors	0	2	james	108	0
benches	0	2	countries	0	2	jars	0	36
beverages	0	1	cows	0	10	jeans	0	1
blouses	0	2	crayons	0	2	keys	0	35
boards	0	1	cushions	0	1	kids	0	1
booze	0	1	days	72	0	knives	37	35
bottles	0	5	diamonds	0	1	lambs	0	6
bugs	0	2	dogs	0	35	lawyers	0	1
buns	0	12	dresses	0	30	leaves	0	35
calves	0	9	drivers	1	0	letters	0	28
candies	0	1	ears	0	34	loafers	0	2
cans	0	1	eighties	0	1	marvel's	0	2
caramels	0	1	faces	0	1	mcdonalds	0	1
cards	0	35	fathers	2	0	microbes	0	1
cells	0	2	files	1	20	missiles	0	1
chairs	0	34	flags	0	37	mountains	36	37
charles	36	0	folders	0	13	mrs	0	18

Word	*RP*	*Pics*	*Word*	*RP*	*Pics*	*Word*	*RP*	*Pics*
news	38	0	puppies	0	1	sperrys	0	1
nose	0	37	pyjamas	0	5	stools	0	1
numbers	0	37	relatives	74	0	straws	0	32
oxfords	0	3	roads	36	0	sunglasses	0	2
pajamas	0	23	rolls	74	45	superheroes	0	3
pause	0	35	roosters	0	8	suppose	0	4
pews	0	20	rose	74	36	ties	0	30
pickles	0	1	runs	0	1	towels	38	36
pillows	0	32	seeds	0	37	trees	0	30
pins	0	23	shoes	0	31	twins	0	22
pjs	0	2	skills	0	1	use	74	0
planes	0	6	skittles	0	2	waterfalls	0	1
polys	0	3	sneakers	0	1	wolves	0	31
potatoes	0	32	spades	0	1	TOTAL	1,031	1,659

APPENDIX 2
Words Analyzed and Counts across Tasks

READING PASSAGE 1

Last Friday, Rose and Charles made cheeseburgers and cinnamon rolls for their visiting relatives from Parkersburg, West Virginia. It was a long drive to Newport News, Virginia, so their relatives were really hungry when they arrived. After washing their hands in Rose and Charles's bathroom and drying them on their decorative his and hers towels, they immediately sat down at the table with their knives and forks already in their hands. They quickly ate their food but complained that there was no syrup for their cinnamon rolls.

READING PASSAGE 2

Many pickup trucks use stick shifts. James was desperate to get his grandfather's old pickup truck, so he had to learn how to drive stick. It didn't go well at first, as James kept grinding the gears. To teach him how to use the clutch, his father had him drive up the mountains and shift through the winding roads until he knew how to shift gears smoothly. It took days and days of practice, but soon his father decided he was ready to take the truck, and James was very excited.

APPENDIX 3
Likert Scale Questions

Participants were asked to indicate the degree to which they agreed with the following statements participants on a five-point Likert scale. While all sentences were presented in random order, prior to factor analysis we sorted them into subgroups of like-questions and gave participants mean scores for each subgroup based on their responses. These means were then fed into factor analysis. The results of the factor analysis motivated further clustering into the following latent variables:

RURAL-SOUTHERN: Have an Accent, Misunderstood, Southern, From Rural, Appalachian, Pro-Hometown, Southern Is Best, Future Rural; negatively for Rich
BAD-MOOD-MISUNDERSTOOD: Misunderstood; negatively for Good Mood, Mumbler
PRO-SOUTH: Southern Is Best, Like Southern Accents

From Rural
+ I grew up in a rural community.
− I grew up in a suburb.
− I grew up in a city.

Future Rural
+ I'd like to live in the country.
− I definitely want to live in a big city.

Appalachia
+ I am Appalachian.
+ I grew up in Appalachia.
+ I sound like I come from Appalachia.
+ A lot of my friends are Appalachian.

Southern Is Best
− I like to vacation in the North.
− I hate to vacation in the South.
+ I would like to live in the South for the rest of my life.
+ Southern food is my favourite food.
+ I would hate to live in the North for the rest of my life.
+ I like Southern culture.

Southern
− Not many of my friends come from the South.
+ I talk to people with Southern accents pretty regularly.
+ I am Southern.
− I am Northern.
+ I have a lot of friends who come from the South.
− I don't really know much about Southern culture.

Rich
− My family is working class.
+ My family is middle class.
+ Most of my highschool friends went to college.
− There is not a lot of opportunity for social mobility in my hometown.

Staying in Hometown
− I don't want to live in my hometown after graduation.
+ I'd be excited to get a job in my hometown after college.
+ Living near where I grow up is a priority to me.

Pro-Hometown
− I do not like my hometown.
+ I have a lot of hometown pride.
+ I go home every chance I get.
+ I like spending breaks back home.
− I couldn't wait to leave home for college.
+ People in my hometown are friendly.
+ I feel like I belong to my hometown.
− I never felt like I fit in back home.
+ I love where I grew up.
+ The place I grew up is a great place to raise kids.

Home Contact
+ I go home every break.
+ My family expects me to come home every break.

Family Key
+ I'd rather be geographically close to family than have a great job.
− I'd rather have an amazing job than be geographically close to family.

Good Mood
+ I was in a good mood before starting this experiment.
+ I am in a good mood right now.
+ Today is going well.
− I'm having a bad day.

− I am in a bad mood right now.

Mumbler
+ I mumble a lot.
+ I think I'm a lazy speaker.
− I speak really clearly.

Misunderstood
+ People often ask me to repeat myself.
+ Other Americans sometimes have trouble understanding me.
+ People often misunderstand me.

Switcher
+ I switch how I speak depending on who I'm talking to.
+ I switch how I speak depending on the context.

Linguistic Insecurity
+ I don't like the accent I grew up speaking.
+ I've consciously changed my accent to stop being teased.
+ I make a conscious effort to make sure people don't make fun of me for my accent.
+ I consciously changed my accent since starting college.
+ I make a conscious effort to make my speech intelligible.
− People tell me they like the way I speak.
+ People tease me about my accent.
− I like my accent.

Have an Accent
− I grew up speaking Standard American English.
+ People often comment on my accent.
+ People can tell where I'm from based on how I talk.
+ People in my family have strong accents.
+ I have close friends with strong accents.
+ People in my hometown have strong accents.
+ I have an accent.
+ I have a strong accent.
− I don't have a Southern accent.

Like Southern Accents
+ I like Southern accents.
− People with Southern accents sound stupid.
+ People with Southern accents sound friendly.

Prescriptive
+ There are right and wrong ways to speak.
+ I hate it when people make grammatical errors.
+ I correct people when they make grammatical errors.
− Misspellings don't bother me.
+ Schools should do a better job of teaching people proper English.

There's a difference between Southern and Appalachian accents.

I have trouble understanding people with accents.

I'm tired today.

NOTES

1. This means, in terms of IPA symbols, that it would be more accurate to use /s/ and /s^h/ as opposed to /z/ and /s/, respectively, throughout our article. However, we stick with /z/ and /s/ based on their familiarity to most linguists.
2. Initial stop voicing also appears to differ across ethnic varieties of U.S. English: it has been found that Black speakers produce more negative lag voicing than White speakers in Florida (Ryalls, Zipprer, and Baldauff 1997) and Mississippi (Herd 2020).

3. The 83 Likert scale questions were not put through factor analysis in their raw form because some responses were perfectly correlated.
4. Except for *his* and *hers*, which were in the phrase *his and hers towels*. Some other *his* tokens also remain in the analysis.
5. These outliers were not measurement errors; on closer inspection, the most dramatic of them were due to audibly elongated productions that appeared to signal hesitancy.
6. The three participants who self-identified as "sorta" Southern were included in the Southern identifying group; though, we find similar results if we put them in the Northern identifying group.
7. The script was based on one by Shigeto Kawahara. The settings were informed by Eager (2015).
8. We found a very similar, significant effect on the initial categorical analysis (voiced vs. devoiced) when we just looked at the Southerners: PRO-SOUTH Southerners produced fewer tokens in the picture-naming task than in the reading passages, relative to the mean.
9. While we suspect that the term "/z/-devoicing" across studies includes qualitatively quite different realizations than the phenomenon described here, it's also worth noting that many studies using aural coding of [z] versus [s] used one coder, employed a combination of aural and visual (spectrogram) information, or do not specify how coding was done (e.g., Holmes 1996; Purnell, Raimy, and Salmons 2009, 343; Bayley and Holland 2014, 391). Therefore, it is not entirely clear how irregular the relatively low agreement rate is here. For the record, one coder was from Northern Virginia, and the other coder had spent formative years in Missouri, Nebraska, Connecticut, and North Carolina.
10. Using a relatively novel, automated tool does not come without caveats, and previous studies using FAVE-align have highlighted potential pitfalls. First, the status of the variable under investigation in the training data is critical (Walker and Sedarous 2015), and for our study, if the Supreme Court Justices are "devoicing" /z/ themselves, then the aligner will be less likely to code devoiced tokens as /s/ over /z/, since it would already associate /z/ with devoiced productions. Thus, it is wise not to interpret the results as a reflection of absolute rates of "/z/ devoicing" but instead as representing "/z/ devoicing" relative to the language usage of the Supreme Court Justices. Bailey (2016) found that as speech rate increased, the aligner's agreement with human coders decreased, but the human coders agreement with each other did not worsen. In terms of our study then, the forced aligner may be less reliable in the passages compared to the picture-naming task because people are speaking at a faster rate.
11. Of the two human coders, the most-agreeing was in 67% agreement with FAVE-align ($\kappa = 0.353$, "fair agreement"). Both human coders had more Z codes than FAVE-align.
12. In regard to final-stop realization, Wolfram and Christian (1976, 63) report anecdotal evidence that Appalachian English speakers in Mercer and Monroe Counties in West Virginia, which bornders Southwest Virgina, glottalize or "devoice" /b, d, g/ word finally on unstressed syllables (e.g., the /d/ in *salad*).

REFERENCES

Anderson, Vicki, and Stuart Davis. 2013 "Where American English Meets German: Devoicing in Pennsylvania Dutchified English." *Journal of the Phonetic Society of Japan* 17, no. 1 (Apr.): 18–25. https://doi.org/10.24467/onseikenkyu.17.1_18.

Appalachian Regional Council. 2021. "Subregions in Appalachia." https://www.arc.gov/wp-content/uploads/2021/11/Subregions-in-Appalachia_2021_Map.pdf.

Avery, Peter, and William J. Idsardi. 2001. "Laryngeal Dimensions, Completion and Enhancement." In *Distinctive Feature Theory*, edited by T. Alan Hall, 41–70. Berlin: Mouton de Gruyter.

Bailey, George. 2016. "Automatic Detection of Sociolinguistic Variation Using Forced Alignment." In "Selected Papers from NWAV 44," edited by Helen Jeoung, *University of Pennsylvania Working Papers in Linguistics* 22, no. 2: 11–20. https://repository.upenn.edu/pwpl/vol22/iss2/3/.

Baum, Shari R., and Sheila E. Blumstein. 1987. "Preliminary Observations on the Use of Duration as a Cue to Syllable-Initial Fricative Consonant Voicing in English." *Journal of the Acoustical Society of America* 82, no. 3 (Sept.): 1073–77. https://doi.org/10.1121/1.395382.

Bayley, Robert, and Cory Holland. 2014. "Variation in Chicano English: The Case of Final (z) Devoicing." *American Speech* 89, no. 4 (Winter): 385–407. https://doi.org/10.1215/00031283-2908200.

Blevins, Juliette. 2006. "A Theoretical Synopsis of Evolutionary Phonology." *Theoretical Linguistics* 32, no. 2 (Dec.): 117–65. https://doi.org/10.1515/TL.2006.009.

Boersma, Paul, and David Weenink. 2017. Praat: Doing Phonetics by Computer (software). Version 6.0.29. http://www.praat.org/.

Carmichael, Katie. 2014. "'I Never Thought I Had an Accent until the Hurricane': Sociolinguistic Variation in Post-Katrina Greater New Orleans." Ph.D. diss., Ohio State University.

Ching, Marvin K. L. 1996. "GreaZy/GreaSy and Other /Z/-/S/ Choices in Southern Pronunciation." *Journal of English Linguistics* 24, no. 4 (Dec.): 295–307. https://doi.org/10.1177/007542429602400405.

Crystal, Thomas H., and Arthur S. House. 1988. "Segmental Durations in Connected-Speech Signals: Current Results." *Journal of the Acoustical Society of America* 83, no. 4 (Apr.): 1553–73. https://doi.org/10.1121/1.395911.

Davidson, Lisa. 2016. "Variability in the Implementation of Voicing in American English Obstruents." *Journal of Phonetics* 54 (Jan.): 35–50. https://doi.org/10.1016/j.wocn.2015.09.003.

Docherty, Gerald J. 1992. *The Timing of Voicing in British English Obstruents*. Berlin: Foris.

Docherty, G. J., D. Watt, C. Llamas, D. Hall, and J. Nycz. 2011. "Variation in Voice Onset Time along the Scottish-English Border." In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII): August 17–21, 2011*, edited by Wai-Sum Lee and Eric Zee, 591–94. Hong Kong: City University of Hong Kong. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Docherty/Docherty.pdf.

Eager, Christopher D. 2015. "Automated Voicing Analysis in Praat: Statistically Equivalent to Manual Segmentation." In *Proceedings of the 18th International Congress of Phonetic Sciences*, edited by the Scottish Consortium for ICPhS 2015. Glasgow: University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0083.pdf.

Gradoville, Michael Stephen. 2011. "Validity in Measurements of Fricative Voicing: Evidence from Argentine Spanish." In *LARP 5: Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, edited by Scott M. Alvord, 59–74. Somerville, Mass.: Cascadilla Proceedings Project. http://www.lingref.com/cpp/larp/5/paper2635.pdf.

Greene, Rebecca Dayle. 2010. "Language, Ideology and Identity in Rural Eastern Kentucky." Ph.D. diss., Stanford University.

Grieser, Jessica A. 2015. "The Language of Professional Blackness: African American English at the Intersection of Race, Place, and Class in Southeast Washington, D.C." Ph.D. diss., Georgetown University.

Hay, Jennifer, and Katie Drager. 2010. "Stuffed Toys and Speech Perception." *Linguistics* 48, no. 4: 865–92. https://doi.org/10.1515/ling.2010.027.

Hazen, Kirk. 2002. "Identity and Language Variation in a Rural Community." *Language* 78, no. 2 (June): 240–57. https://doi.org/10.1353/lan.2002.0089.

Hazen, Kirk, and Sarah Hamilton. 2008. "A Dialect Turned Inside Out: Migration and the Appalachian Diaspora." *Journal of English Linguistics* 36, no. 2 (June): 105–28. https://doi.org/10.1177/0075424208317127.

Hazen, Kirk, Jordan Lovejoy, Margery Webb, Emily Vandevender, and Kiersten Woods. 2015. "The Sociophonetics of Z Devoicing." Paper presented at Southeastern Conference on Linguistics 82, Raleigh, N.C., Apr. 9–11, 2015.

Herd, Wendy. 2020. "Sociophonetic Voice Onset Time Variation in Mississippi English." *Journal of the Acoustical Society of America* 147, no. 1 (Jan.): 596–605. https://doi.org/10.1121/10.0000545.

Holmes, Janet. 1996. "Losing Voice: Is Final /z/ Devoicing a Feature of Maori English?" *World Englishes* 15, no. 2 (July): 193–205. https://doi.org/10.1111/j.1467-971X.1996.tb00105.x.

Hughes, George W., and Morris Halle. 1956. "Spectral Properties of Fricative Consonants." *Journal of the Acoustical Society of America* 28, no. 2 (Mar.): 303–10. https://doi.org/10.1121/1.1908271.

Hunnicutt, Leigh, and Paul A. Morris. 2016. "Prevoicing and Aspiration in Southern American English." In "Proceedings of the 39th Annual Penn Linguistics Conference," edited by Sunghye Cho. *University of Pennsylvania Working Papers in Linguistics* 22, no. 1: 215–24. https://repository.upenn.edu/pwpl/vol22/iss1/24/.

Iverson, Gregory K., and Joseph C. Salmons. 1995. "Aspiration and Laryngeal Representation in Germanic." *Phonology* 12, no. 3 (Dec.): 369–96. https://doi.org/10.1017/S0952675700002566.

Iverson, Gregory K., and Joseph C. Salmons. 2003. "Legacy Specification in the Laryngeal Phonology of Dutch." *Journal of Germanic Linguistics* 15, no. 1 (Mar.): 1–26. https://doi.org/10.1017/S1470542703000242.

Iverson, Gregory K., and Joseph C. Salmons. 2011. “Final Devoicing and Final Laryngeal Neutralization.” In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 3: 1622–43. Malden, Mass.: Wiley-Blackwell. https://doi.org/10.1002/9781444335262.wbctp0069.

Jacewicz, Ewa, Robert Allen Fox, and Samantha Lyle. 2009. “Variation in Stop Consonant Voicing in Two Regional Varieties of American English.” *Journal of the International Phonetic Association* 39, no. 3 (Dec.): 313–34. https://doi.org/10.1017/S0025100309990156.

José, Brian. 2010. “The Apparent-Time Construct and Stable Variation: Final /z/ Devoicing in Northwestern Indiana.” *Journal of Sociolinguistics* 14, no. 1 (Feb.): 34–59. https://doi.org/10.1111/j.1467-9841.2009.00434.x.

Kurath, Hans, and Raven I. McDavid, Jr. 1961. *The Pronunciation of English in the Atlantic States: Based upon the Collections of the Linguistic Atlas of the Eastern United States*. Ann Arbor: University of Michigan Press.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William. 2001. *Principles of Linguistic Change*. Vol. 2, *Social Factors*. Oxford: Blackwell.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin : Mouton de Gruyter.

Landis, J. Richard, and Gary G. Koch. 1977. “The Measurement of Observer Agreement for Categorical Data.” *Biometrics* 33, no. 1 (Mar.): 159–74. https://doi.org/10.2307/2529310.

Lisker, Leigh. 1986. “‘Voicing’ in English: A Catalogue of Acoustic Features Signaling /b/ versus /p/ in Trochees.” *Language and Speech* 29, no. 1: 3–11. https://doi.org/10.1177/002383098602900102.

Lisker, Leigh, and Arthur S. Abramson. 1964. “A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements.” Word 20, no. 3: 384–422. https://doi.org/10.1080/00437956.1964.11659830.

Montgomery, Michael. 1997. “The Scotch-Irish Element in Appalachian English: How Broad? How Deep?” In *Ulster and North America: Transatlantic Perspectives on the Scotch-Irish*, edited by H. Tyler Blethen and Curtis W. Wood, 189–212. Tusclaoosa: University of Alabama Press.

Pratt, Teresa. 2018. “Embodying Toughness: LOT-Raising, /l/-Velarization, and Retracted Articulatory Setting.” In “Selected Papers from New Ways of Analyzing Variation (NWAV 46),” edited by Jordan Kodner and Lacey Wade. *University of Pennsylvania Working Papers in Linguistics* 24, no. 2: 107–66. https://repository.upenn.edu/pwpl/vol24/iss2/13/.

Purnell, Thomas, Eric Raimy, and Joseph Salmons. 2009. “Defining Dialect, Perceiving Dialect, and New Dialect Formation: Sarah Palin’s Speech.” *Journal of English Linguistics* 37, no. 4: 331–55. https://doi.org/10.1177/0075424209348685.

Purnell, Thomas, Joseph Salmons, and Dilara Tepeli. 2005. “German Substrate Effects in Wisconsin English: Evidence for Final Fortition.” *American Speech* 80, no. 2 (Summer): 135–64. https://doi.org/10.1215/00031283-80-2-135.

Purnell, Thomas, Joseph Salmons, Dilara Tepeli, and Jennifer Mercer. 2005. "Structured Heterogeneity and Change in Laryngeal Phonetics: Upper Midwestern Final Obstruents." *Journal of English Linguistics* 33, no. 4: 307–38. https://doi.org/10.1177/0075424205285637.

Raphael, Lawrence J. 1972. "Preceding Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word-Final Consonants in American English." *Journal of the Acoustical Society of America* 51, no. 4B (Apr.): 1296–1303. https://doi.org/10.1121/1.1912974.

Reed, Paul E. 2016. "Sounding Appalachian: /ai/ Monophthongization, Rising Pitch Accents, and Rootedness." Ph.D. diss., University of South Carolina.

Rohena-Madrazo, Marcos. 2015. "Diagnosing the Completion of a Sound Change: Phonetic and Phonological Evidence for /ʃ/ in Buenos Aires Spanish." *Language Variation and Change* 27, no. 3 (Oct.): 287–317. https://doi.org/10.1017/S0954394515000113.

Rosenfelder, Ingrid, Joe Fruehwald, Keelan Evanini, and Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite (software). https://web.archive.org/web/20120124025539/http://fave.ling.upenn.edu/.

Ryalls, John, Allison Zipprer, and Penelope Baldauff. 1997. "A Preliminary Investigation of the Effects of Gender and Race on Voice Onset Time." *Journal of Speech, Language, and Hearing Research* 40, no. 3 (June): 642–45. https://doi.org/10.1044/jslhr.4003.642.

Saito, Hiroko. 1994. "Devoicing of Word-Final /z/ in English: A Comparison of Inflectional and Lexical Instances of /z/." *Area and Culture Studies* 49: 139–60. http://repository.tufs.ac.jp/bitstream/10108/23598/1/acs049008.pdf.

Salmons, Joseph. 2020. "Germanic Laryngeal Phonetics and Phonology." In *The Cambridge Handbook of Germanic Linguistics*, edited by Michael T. Putnam and B. Richard Page, 119–42. Cambridge: Cambridge University Press.

Scobbie, James M. 2006. "Flexibility in the Face of Incompatible English VOT Systems." In *Laboratory Phonology 8*, edited by Louis Goldstein, D. H. Whalen, and Catherine T. Best, 367–92. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197211.2.367.

Smith, Caroline L. 1997. "The Devoicing of /z/ in American English: Effects of Local and Prosodic Context." *Journal of Phonetics* 25, no. 4 (Oct.): 471–500. https://doi.org/10.1006/jpho.1997.0053.

Soli, Sigfrid D. 1982. "Structure and Duration of Vowels Together Specify Fricative Voicing." *Journal of the Acoustical Society of America* 72, no. 2 (Aug.): 366–78. https://doi.org/10.1121/1.388080.

Sonderegger, Morgan, Jane Stuart-Smith, Thea Knowles, Rachel Macdonald, and Tamara Rathcke. 2020. "Structured Heterogeneity in Scottish Stops over the Twentieth Century." *Language* 96, no. 1 (Mar.): 94–125. https://doi.org/10.1353/lan.2020.0003.

Stevens, Kenneth N., Sheila E. Blumstein, Laura Glicksman, Martha Burton, and Kathleen Kurowski. 1992. "Acoustic and Perceptual Characteristics of Voicing in Fricatives and Fricative Clusters." *Journal of the Acoustical Society of America* 91, no. 5 (May): 2979–3000. https://doi.org/10.1121/1.402933.

Thompson, Roger M. 1977. "Mexican-American English: Social Correlates of Regional Pronunciation." *American Speech* 50, no. 1–2 (Spring–Summer 1975): 18–24. https://doi.org/10.2307/3087859.

University of Virginia Demographics Research Group. 2022. "Virginia's Demographic Regions." https://demographics.coopercenter.org/sites/demographics/files/media/files/2022-05/RegionsMap_DemographicsCooperCenter.pdf.

Verhoeven, Jo, Allen Hirson, and Kavya Basavaraj. 2011. "The Devoicing of Fricatives in Southern British English." In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII): August 17–21, 2011*, edited by Wai-Sum Lee and Eric Zee, 2070–73. Hong Kong: City University of Hong Kong. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Verhoeven/Verhoeven.pdf.

Walker, Abby. 2019. "The Role of Dialect Experience in Topic-Based Shifts in Speech Production." *Language Variation and Change* 31, no. 2 (July): 135–63. https://doi.org/10.1017/S0954394519000152.

Walker, Abby. 2020. "Voiced Stops in the Command Performance of Southern US English." *Journal of the Acoustical Society of America* 147, no. 1 (Jan.): 606–15. https://doi.org/10.1121/10.0000552.

Walker, Abby, Andrew Burlile, and Katherine Askew. 2016. "Investigating the Interaction between Speaker Dialect and Listener Differences across Two Tasks." Poster presented at the 15th Biennial Conference on Laboratory Phonology (LabPhon), Ithaca, N.Y., July 13–16, 2016.

Walker, Abby, and Yourdanis Sedarous. 2015. "Automating Categorical Coding of Phonological Variables: Implementation and Evaluation." Paper presented at the 89th annual meeting of the Linguistic Society of America, Portland, Ore., Jan. 8–11, 2015.

Weatherholtz, Kodi, Kathryn Campbell-Kibler, and T. Florian Jaeger. 2014. "Socially-Mediated Syntactic Alignment." *Language Variation and Change* 26, no. 3 (Oct.): 387–420. https://doi.org/10.1017/S0954394514000155.

Wolfram, Walt, and Donna Christian. 1976. *Appalachian Speech.* Arlington, Va.: Center for Applied Linguistics.

Yuan, Jiahong, and Mark Liberman. 2008. "Speaker Identification on the SCOTUS Corpus." *Journal of the Acoustical Society of America* 123, no. 5 (May): 3878. https://doi.org/10.1121/1.2935783.

Yuan, Jiahong, and Mark Liberman. 2011. "Automatic Detection of 'g-Dropping' in American English Using Forced Alignment." In *2011 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU 2011): Proceedings; December 11–15, 2011, Waikoloa, Hawaii, U.S.A.*, edited by Thomas Hain and Kai Yu, 490–93. http://doi.org/10.1109/ASRU.2011.6163980.

RACHEL HARGRAVE holds a B.A. in English from Virginia Polytechnic Institute and will complete their literature M.A. from Virginia Tech in spring 2018. They are interested in mythopoeic literature and adaptation studies as well as multimedia pedagogy, queerness in anime, and queer-centered pedagogy. Email: rachh15@vt.edu.

AMY SOUTHALL graduated summa cum laude from Virginia Polytechnic Institute in 2017, receiving her B.A. in professional and technical writing with a certificate in literature and language, preeducation, and minors in agriculture and applied economics, peace studies and violence prevention, and language sciences. In 2020, Southall received her master's degree in public administration with a certificate in local government management. She currently works in local government in public relations. Email: als26@vt.edu.

ABBY WALKER is an associate professor in the Department of English at Virginia Polytechnic Institute. She got her M.A. in linguistics in 2007 from the University of Canterbury in New Zealand, and her Ph.D. in linguistics in 2014 from Ohio State University. Her interests include speech perception, speaker perception, intraspeaker variation, and how these processes interact. Email: ajwalker@vt.edu.