Title: Incomplete Neutralization in African American English: The case of final consonant voicing

Author: Charlie Farrington

Dept. of Linguistics

1290 University of Oregon

Eugene, OR 97403-1290

Phone: 802-370-9585

Email: crf@uoregon.edu

Short Title: Incomplete neutralization in AAE
ABSTRACT

In many varieties of African American English (AAE), glottal stop replacement and deletion of word final /t/ and /d/ results in consonant neutralization, while the underlying voicing distinction may be maintained by other cues, such as vowel duration. Here, I examine the relationship between vowel duration, final glottal stop replacement, and deletion of word final /t, d/ to determine if the phonological contrast of consonant voicing is maintained through duration of the preceding vowel. Data come from conversational interviews of AAE speakers in North Carolina, Tennessee, and Washington DC. Results indicate that glottalization and the deletion of word final /t/ and /d/ are widespread across the speakers in the analysis. Additionally, the duration of vowels is significantly longer before underlying /d/ than /t/ for consonant neutralized contexts, thus showing that duration, normally a secondary cue to final voicing, may be becoming a primary cue in AAE.

ACKNOWLEDGEMENTS

I thank Valerie Fridland at the University of Nevada, Reno and Walt Wolfram at North Carolina State University for allowing me access to their data for this research. The data used for this study were collected with support from National Science Foundation grant numbers BCS-1358724, BCS-1129778, and BCS-0001725. I’d also like to thank Tyler Kendall, Mary Kohn, Volya Kapatsinski, Lisa Redford, Jason McLarty, and two anonymous reviewers for helpful comments on earlier versions of this paper.
INTRODUCTION

In African American English (AAE), besides the extensively studied obstruent variation in consonant cluster reduction, there is variation in word final consonant singletons. Word final /t/ and /d/ can be replaced with a glottal stop or deleted, both of which result in coda neutralization. /d/ glottalization, in particular, has been previously analyzed in AAE, and is often referred to as devoicing (Farrington, 2011; Wolfram, 1969). Though this kind of variation does not occur in Mainstream American English (MAE) varieties, coda neutralization processes are common across languages with word final devoicing, a classic example of neutralization (Yu, 2011), devoicing has been called a ‘vernacular universal’ (Chambers, 2000:12). The underlying voicing distinction, however, can be maintained by other cues, such as by the duration of the preceding vowel. Normally a secondary cue to final voicing, vowel duration may be becoming a primary cue in AAE, resulting from coda neutralization.

In the current paper, I examine the relationship between vowel duration, final glottal stop replacement, and deletion of /t, d/ to determine if the phonological contrast of consonant voicing is maintained through the duration of the preceding vowel in these consonant neutralized realizations. In MAE, the effect of consonant voicing on vowel duration is said to be partially phonologized, even when consonant distinctions are maintained (Solé, 2007; Yu, 2011). In their work on AAE in Eastern North Carolina, Holt, Jacewicz, and Fox (2016) showed that AAE speakers have more pronounced duration differences between the vowels preceding word final [t] and word final [d] when compared to MAE speakers, but they did not look at how this variation relates to different surface realizations of /t, d/. Such an increase in durational contrast may be expected to reflect a shift in cues as consonant identity becomes more similar through glottal stop replacement.
BACKGROUND

*Word final neutralization*

Word final neutralization processes are common across a variety of languages, including obstruent devoicing in German, Dutch, Russian, Polish, and many others (see Kharlamov [2012] for an extensive review regarding both the production and perception of devoicing in experimental studies). It has also been found in some American English varieties, particularly as a substrate effect in Wisconsin German English varieties (Purnell, Tepeli, & Salmons, 2005), though the effect here is stop fortition and not glottal stop replacement (Iverson & Salmons, 2006). In the cases of obstruent devoicing above, neutralization is often incomplete, such that consonant voicing is maintained through other cues. Iverson and Salmons (2012) showed that there are multiple phonetic cues available to speakers in making voicing contrasts, even when a final consonant is neutralized, which, in some cases, can lead to contrast enhancement (Kingston & Diehl, 1994). Additionally, Kirby (2010) argued that with devoicing processes, there is a cue-trading relationship between what was a primary cue (consonant voicing) and secondary cue (e.g., vowel duration, stop closure duration, etc.).

In work on the phonological import of neutralization, Yu’s (2011) review chapter on contrast reduction encompasses the notions of neutralization and merger, where the former may or may not have a contrast maintained elsewhere, and in the latter, there is no trace of a contrast in the synchronic system. Yet more systematic investigations are needed to answer questions about the pervasiveness of contrast reduction, especially in non-experimental settings. Silverman’s (2012) extensive cross-linguistic analysis similarly made a distinction between what should be called complete neutralization, where a contrast is completely lost, and incomplete
Incomplete Neutralization in AAE

neutralization, where a contrast is maintained elsewhere. When vowel duration is a primary cue to consonant voicing, the effect could be one of phonologization (Hyman, 1976), where a secondary cue resulting from phonetic differences (e.g., the consonant voicing effect on vowel duration) becomes a primary cue.

In a cross-linguistic study on the effect of consonant voicing on vowel duration in English, Catalan, and Arabic, Solé (2007) suggested that if a feature distinction, such as vowel duration, is phonological rather than phonetic, the feature will vary with speech rate because the targets are phonologically specified. In Catalan, for example, vowel duration is longer before voiced consonants compared to vowels before voiceless consonants, but this difference does not vary as a result of speech rate, highlighting the phonetic basis for the voicing distinction. In English, however, a slower speech rate results in an increase of the difference between the duration of vowels before voiced and voiceless consonants (Solé, 2007), which means there are phonological targets for English speakers. Yu (2011) used this to suggest that English has partially phonologized consonant voicing in vowel duration. A phonological difference like this one in vowel duration is an initial step in Hyman’s (1976) phonologization categories which show that a new distinction (e.g., phonological vowel duration) leads to the loss of another distinction (e.g., coda consonants). AAE is unique among varieties of English because glottal stop replacement occurs for both word final /t/ and /d/, presenting a unique testing ground for the continued phonologization of vowel duration in an American English variety.

Glottalization in Mainstream American English

Glottalization of the voiceless coronal stop /t/ is a complex sociolinguistic variable that has been reported on extensively, especially in the United Kingdom (Docherty & Foulkes, 1999; Foulkes...
Most detailed sociolinguistic work on American English glottalization focus on the realization of /t/ in both word medial and word final environments. In general, glottal stop reinforcement ([ʔt] or [tʔ]) is a relatively common feature in American English, for example, *football* realized as [foʔtbal] (Eddington & Taylor, 2009:308). Like British English, the regional varieties of American English studied have consistent internal effects, and different frequencies of use are constrained by phonetic context. Glottal stop replacement and reinforcement are both found across the country, and are socially salient in certain varieties (Eddington & Channer, 2010; Eddington & Taylor, 2009; Roberts, 2006).

Work in laboratory phonology indicates that what is thought of as a glottal stop encompasses different acoustic characteristics, which depend on several variables, including phrasal position, segmental context, gender, dialect, and individual speaker differences (Garellek, 2013; Garellek & Seyfarth, 2016). In addition to these acoustic characteristics, Dilley and Pitt (2007) found that the glottal variant of /t/ accounted for one third of the /t/ tokens in their American English data, and they suggest that lexical content and surrounding contextual information might be most important in interpreting ambiguous information.

*African American English neutralization*

Word final /d/ glottalization is one of the features that consistently differentiates AAE from MAE varieties (Fasold, 1981). Scholars have used the term devoicing to describe the feature, focusing on the voicing differences between the underlying /d/ and the surface form, but in this study, I use the terms *glottal replacement of* /d/ or *glottal /d/* since the most common phonetic realization of word final /d/ is a glottal stop in AAE (Farrington, 2011). Thomas (2007), for example, explained that devoicing of voiced stops is often accompanied by glottalization, and
that these stops could also be deleted with a lengthened vowel to mark consonant voicing. Green (2002) noted that a generalized devoicing rule applies to coda /b, d, g/, rendering these sounds as voiceless stops [p, t, k]. Of the three voiced stops, the focus is often on /d/ because it’s the most frequent member of the voiced stop class (Wolfram, 1969).

Glottal replacement of word final /d/ in AAE has been documented in several geographic settings, including Detroit, MI (Kohl & Anderson, 2000; Nguyen, 2006; Wolfram, 1969), Houston, TX (Koops & Niedzielski, 2009), Durham, NC (Farrington, 2011), and Washington, DC (Fasold, 1972; Grieser, 2014). Additional sources citing glottalization (or devoicing) as a feature of the local variety include New York City (Labov, Cohen, Robins, & Lewis, 1968), Los Angeles, CA (Legum, Pfaff, Tinnie, & Nichols, 1971), and Minneapolis, MN (Pederson, 1967). In addition, developmental clinical literature on AAE focuses on how coda devoicing fits into phonological assessments (Baran & Seymour, 1976; Moran, 1993; Stockman, 2006). Thus, glottal /d/ appears to be one of the features that AAE speakers share across regions. Importantly, /d/ glottalization is also a feature that is used across both working class and middle class speakers of AAE. In Detroit for example, Nguyen (2006) looked at word final /d/ and showed that stratification was conditioned by gender rather than social class. In Detroit, women prefer the glottal variant compared to men, who prefer deletion.

There has been no systematic study looking at the relationship between /d/ glottalization and word final /t/ glottalization in AAE, though there have been a few references in the literature to a relationship between these two features. For example, Luelsdorff (1975), based on elicited speech from one speaker, suggested that a series of phonological processes are at work, including vowel lengthening before voiced stops, followed by consonant weakening.
Final consonant singleton deletion is sometimes discussed as a feature in AAE. For example, Thomas (2007) noted that in addition to glottalization of voiced stops, outright deletion is possible too, and this can be extended to voiceless stops. Most often in studies of /d/ devoicing, deletion is treated as a process in utterance medial position, which can often be accounted for as a connected speech process (Nolan, 1996; Temple, 2014). For pre-pausal deletion, Wolfram (1969) found a social class distinction, such that in the lower working class, speakers exhibit pre-pausal deletion over 16% of the time, compared to zero percent rates in the upper middle-class group (107). More recently, in Atlanta, Harrison (2007) discussed widespread consonant elision, across all obstruent types and word positions, and suggested that elision in AAE is a result of a lexical item’s metrical status in the prosodic domain. Such work demonstrates that more focus is needed on the phonological processes in AAE, especially for consonants.

**AAE Vowel Duration**

Recent papers by Holt and colleagues (Holt, Jacewicz, & Fox, 2015, 2016) provide evidence for vowel duration differences before voiced and voiceless consonants in AAE. Holt et al. (2015) looked at the role of vowel duration in AAE speakers in Eastern North Carolina compared to MAE speakers in Western NC. Speakers read vowels in h_d fames (e.g., heed, hid) in a lab setting. They found that AAE vowels were longer than MAE vowels and vowel duration was not shown to vary by age group (older versus younger speakers). In a follow-up study, the authors looked at how vowel duration before [t] and [d] differed across ethnicities (Holt et al., 2016). Results here showed that AAE speakers have longer vowels before voiced consonants when compared to MAE speakers, but durations before voiceless consonants were not significantly
different between ethnicities. While variation of final stops is mentioned—“glottalization or devoicing of /d/ has been observed so frequently that the variation is considered a ubiquitous feature of vernacular dialects” (Holt et al., 2016:2)—this laboratory experiment focused on how vowel duration relates to the coronal stop realizations, [t] and [d]. Their results suggest that extensive vowel duration lengthening might be a primary cue for consonant voicing distinctions in AAE, at least for coronal stops.

The current paper investigates coda stop glottalization and deletion and the relationship to neutralization in AAE by first examining in detail the distribution of glottalization and deletion of word final /t, d/, paying particular attention to whether speakers from three Southeastern cities vary in terms of their consonant realizations. Next, preceding vowel durations are investigated to determine whether underlying voiced and voiceless consonants are differentiated, and if this contrast is exaggerated, possibly indicating a greater degree of phonologization. Based on work by Holt et al., (2016), and what is known about other languages that exhibit final stop neutralization, a reasonable prediction would be that neutralization is incomplete for AAE, such that a contrast between consonant types is maintained in the vowel duration across different stop realization types.

**METHODS**

**Data Sources**

This study examines data from conversational interviews conducted as part of sociolinguistic studies in three cities in the Southeastern United States: Memphis, TN, Durham, NC, and Washington DC, with a focus on speakers under 30 years of age, listed in Table 1.
TABLE 1. Speaker sample by sex and city

<table>
<thead>
<tr>
<th>City</th>
<th>Year Collected</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memphis, TN</td>
<td>2001</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Durham, NC</td>
<td>2012</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Washington, DC</td>
<td>2015</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Memphis, TN.

Data were collected in Memphis in 2001 as a part of a project conducted by Valerie Fridland (2003). Fieldworkers were African American students at either the University of Memphis or LeMoyne-Owen College, who were recruited and trained for the project. They were asked to record a naturally occurring conversation with family members or close friends from the local community. Speakers were recorded using a Sony MZ-R70 digital minidisc recorder, and an omni-directional Audio-Technica ATR35s lavaliere microphone in an everyday setting, with efforts to ensure that the recording situation was as quiet as possible. Minidiscs were transferred to WAV files using a Sony ATRAC DSP Type-R/ATRAC3 minidisc player in 2015.

Durham, NC.

Recordings from Durham, NC were collected as part of the Frank Porter Graham (FPG) longitudinal study of AAE in 2011-2012 (cf. Van Hofwegen & Wolfram, 2010). Fieldworkers included two African American graduate students and an older European American woman, who was the principle investigator of the study and knew each participant since infancy when they were recruited for the study. Interviews consisted of several tasks, but data in the current analysis come from the sociolinguistic interviews only. A subset of twelve speakers from the post high school time point (~19-21 years of age) were selected based on maintaining residence in Durham for the majority of their lives. Recordings were made as WAV files using a Marantz PMD-660 digital recorder with a lavaliere microphone.

Washington, DC.
The twelve recordings from Washington DC are part of the DCB component of Corpus of Regional African American Language (CORAAL; Kendall & Farrington, 2018). Interviews were primarily conducted by an African American female with ties to the community, but there are some interviews by an African American male who grew up and lives in the area. Twelve younger speakers in the Working Class and Lower Middle Class groups were selected. Recordings were made as WAV files using a Marantz PMD-661 digital recorder with a Shure SM93 lavaliere microphone.

**Coding word final /t, d/**

Tokens of post-vocalic, word final /t, d/ lexical items were extracted, focusing on content words alone. No more than four tokens of each word type were extracted per speaker. For each consonant coded, the preceding vowel’s duration was also collected. In each example presented (Figures 1 through 4), dotted lines on the spectrogram for the vowel immediately preceding the /t/ or /d/ represent where the duration was measured for each example token. These measurements follow Thomas’ (2011) procedures for extracting vowel duration. This includes marking the onset and offset where F2 is visible, and not including the stop gap for coronal stops.

The coding scheme for the token types, detailed below, includes full stops ([t, d]), glottal replaced stops ([ʔ]), glottal reinforced stops ([ʔt], [ʔd]) and deleted (zero coda) consonants. This categorical analysis of variants is an analytical construct that underrepresents the amount of possible variation of final stop realizations, but at the same time uses acoustic and instrumental information in coding each consonant (Docherty & Foulkes, 1999). Additionally, without articulatory data, it must be acknowledged that there could be articulations which are not necessarily auditory or observable from the acoustic signal (Temple, 2014).
[d] and [t]

These represent voiced and voiceless coronal stops, which show no evidence of glottalization. A release may or may not be present. The first two formants will also show transitions from the preceding vowel (F1 lowering and F2 rising). A stop gap is usually present, but its duration varies considerably. Figure 1 shows an unreleased voiced stop on the word *side*, with no glottalization present, clear formant transitions, and voicing present on the stop.

![Graphical representation of a word with an unreleased [d] sound](image)

**Figure 1.** Example of the word ‘side’ with an unreleased [d] by a Washington DC female.

[ʔ]

Glottalization can be realized in several ways in American English, from creaky voicing to a full glottal stop (Thomas, 2011). Figure 2 shows a full glottal stop in a stressed environment. For this token of *iPod*, in particular, there is no glottalization (creaky voicing) leading up to the
glottal stop. In addition to the auditory confirmation, the formants here lack transitions in the way they would for a coronal stop.

**Figure 2.** Example of glottal replaced /d/ in the word ‘iPod’ by a Durham male.

/ʔd/, /ʔt/

Glottalized coronals have similar articulatory processes as full glottal tokens, but there are several differences (Docherty & Foulkes, 1999; Kohl & Anderson, 2000). Glottalized coronal variants have an oral rather than laryngeal stop gap, where articulators are held together in position for a coronal consonant (which can be voiced or voiceless), followed by an acoustic transient. There are a number of acoustic correlates one can look for on a spectrogram, including widening glottal pulses as well as formant transitions (Garellek, 2013). This usually means F1 decreases and F2 increases for coronal stops. Additionally, with a voiced glottalized coronal, there could be acoustic murmur afterward (Thomas, 2011). In the example in Figure 3, there is
glottalization leading up to a stop gap for the oral coronal (unreleased) stop. In the current analysis, glottal reinforced coronals are collapsed with coronal stops due to low token counts \( n = 161 \).

**Figure 3.** Example of the word ‘squad’ with a glottal reinforced /d/ by a Washington DC female.

**Zero Coda.**

Zero coda shows no evidence for a glottal or coronal stop. Zero coda is especially frequent in utterance medial environments, especially when followed by a consonant. This pre-consonantal deletion process is essentially the same as consonant cluster reduction studies have found (Wolfram, 1969, 1974), but can also be related to more general connected speech processes (Temple, 2014). If the word is pre-pausal, the formants may fade into breathiness, but this isn’t essential. In addition, cases of creaky vowels without a coronal closure or full glottal stop are
also coded as zero coda (Nguyen, 2006). Figure 4 presents a pre-pausal, deleted /d/ in the word *side*.

![Figure 4](image.png)

**Figure 4.** Example of the word ‘side’ with a zero coda /d/ by a Memphis female.

**Independent Variables and Statistical Models**

A total of 3638 tokens were extracted for analysis (2836 monosyllabic, 802 multisyllabic), 1880 of word final /t/ and 1758 of word final /d/. There are 450 unique word types. Two primary analyses were completed, the first addressing consonantal realization and the second the role of vowel duration in maintaining voicing differences.²

To control for interspeaker variation and word specific effects, speaker and word were included as random effects. Possible external variables include fixed effects for sex and city. Sex is included as an external factor because it has been found that men are more likely to delete /d/ and women are more likely to exhibit glottal replacement in Detroit (Nguyen, 2006; Wolfram,
1969). With vowel duration, there is limited evidence that women exhibit longer vowel durations than men in AAE (Holt et al., 2015) and American English more broadly (Jacewicz, Fox, & Salmons, 2007).

With recent work on regional phonologies in AAE (Arnson & Farrington, 2017; Thomas & Bailey, 2015), the question of regional variation continues to be an important one to address within any study of AAE. There are three levels for city: Memphis, Durham, and Washington DC. Each of the cities have large African American populations: as of the 2010 Census, Memphis was 63%, Durham 41%, and Washington DC 50% African American. Overall city demographics only tell part of the story, as racial (de facto) segregation in neighborhoods is apparent in many large urban areas (Massey & Denton, 1989). In terms of sub-dialect regions, Memphis is considered part of the mid-South with strong Southern ties (Fridland, 2003), while Durham is in a South Atlantic state (North Carolina). DC, generally considered between the Mid-Atlantic dialect area and the South, has had a large African American population since the 19th century. Resulting from the Great Migration, there was an influx of African Americans from the South Atlantic states (Virginia, North Carolina, South Carolina) to Washington DC in the mid-twentieth century, which led to Southern cultural alignment as well as more Southern-like vowel phonologies (Arnson & Farrington, 2017).

The remaining variables include fixed effects for consonant type (t, d), utterance position (medial, final), syllable stress (stressed, unstressed), syllables in a word (monosyllabic, multisyllabic), following segment (pause, consonant, vowel), and lexical frequency.\(^4\,^5\)

**ANALYSIS**

*Word final stops in AAE*
We begin our analysis by looking at the overall distribution of glottalization and deletion in each of the cities represented, plotted in Figure 5. Bar plots are organized by consonant type and city, with *glottal stop* (glottal stop realization), *zero coda* (deleted), and *coronal stop* (fully articulated coronal stop and glottal reinforced coronal stops) realizations.

**Figure 5. Distribution of final stop realization by City.**

Glottal stop replacement is the most frequent realization of /t/ and /d/ across each of the three cities, especially for /t/, comprising 73 percent of the /t/ tokens and 39 percent of the /d/ tokens. Deletion is more common for /d/ than /t/ overall (Table 3). Considering the three cities, the overall patterns appear more similar between Durham and DC compared with Memphis. This makes sense from an historical perspective, since the population changes from the mid-twentieth century in DC resulted in African American populations moving from the Carolinas and Virginia...
to DC. In sum, glottal stops and zero codas are both common in the three varieties of AAE, with glottal stops accounting for 56 percent of the data and zero coda 21 percent.

To evaluate these differences statistically, logistic mixed-effects regression models were constructed for consonant realization for the 36 speakers in the study using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R. Interactions between consonant type and all other factors were tested in model build up to tease out any differences between consonants. Two comparison models were built. The first involves regressions comparing only glottal variants and the full coronal stop form, which shows factors favoring glottal stops. The second comparison collapses the two stop realizations (glottal stop, coronal stop) into a realized category, and is compared to zero coda tokens.

**Glottal vs Coronal**

The best fit model included effects for consonant type (reference level = /t/), following segment (reference level = pause), syllable stress (reference level = stressed syllable) and city (reference level = Memphis). An interaction between consonant type and realization improved the fit on the model. Statistical results are listed in Table 2. Model testing included interactions between city and internal effects, but these interactions never significantly improved the model. Additionally, utterance position did not improve model fit. While position of the coda consonant within an utterance is likely important, that information is essentially captured in the following environment, where most, but not all, tokens of pre-pausal /t/ and /d/ occur at the end of an utterance. Surprisingly, sex and frequency were not significant factors in any of the models tested.

Table 2. Summary of best mixed effects regression model: Glottal versus Coronal (overall percentage glottal 71%) (n = 2889)
<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates (Std. Err.)</th>
<th>Number of tokens and (percentage glottal) for factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consonant (vs /t/)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/d/</td>
<td>1.29*** (0.21)</td>
<td>1675 (81)</td>
</tr>
<tr>
<td>City (vs Memphis)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durham</td>
<td>-0.72*** (0.20)</td>
<td>1214 (56)</td>
</tr>
<tr>
<td>DC</td>
<td>-0.57** (0.20)</td>
<td>989 (75)</td>
</tr>
<tr>
<td>Following Segment (vs Pause)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consonant</td>
<td>0.95*** (0.20)</td>
<td>985 (66)</td>
</tr>
<tr>
<td>Vowel</td>
<td>2.96*** (0.19)</td>
<td>915 (72)</td>
</tr>
<tr>
<td>Syllable Stress (vs Stressed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstressed</td>
<td>-0.71*** (0.20)</td>
<td>288 (74)</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consonant=d : Realization=Consonant</td>
<td>0.96*** (0.26)</td>
<td></td>
</tr>
<tr>
<td>Consonant=d : Realization=Vowel</td>
<td>-0.06 (0.26)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *p < .05; **p < .01; ***p < .001. Other predictors tested in model buildup include Utterance Position, Speaker Sex, Word Frequency, Number Syllables in Word.

There are significant main effects for consonant, city, following segment and syllable stress. For consonant type, /d/ favors coronal stops to glottal stops when compared to /t/. Durham and DC disfavor coronal stops to glottal stops when compared to Memphis, which means that Memphis preserves coronal stops most among the three cities. Following consonants and vowels favor coronal stops compared to following pauses. Figure 6 shows the proportion of consonant realization (glottal versus coronal) by following environment. And finally, unstressed syllables favor glottalization.
Figure 6. Distribution of final stop realization (glottal versus coronal) by following environment.

*Deleted vs Realized Stops*

In this analysis, deleted (zero coda) variants (reference category) are compared to realized stops, collapsing both the glottal and coronal stop realizations together. Table 3 displays the summary of the predictors in the best fit model constructed for /t/ and /d/. Main effects are found for consonant, following segment and syllable stress. For following segment, consonants and vowels both strongly favor deletion compared to following pauses. This result is expected as it relates to connected speech processes (Nolan, 1996; Temple, 2014) and has been consistently found in work on consonant cluster reduction in AAE (Thomas & Bailey, 2015; Wolfram, 1969). Finally, unstressed syllables favor realized stops over deleted stops. Two interactions improved the model, consonant type and city, with a significant interaction for Durham, and consonant type
and syllable stress. For /t/, unstressed syllables favor deletion. This is somewhat unexpected, and shows a different pattern than /d/, and also a different pattern than coronal versus glottal stops above. Once again, sex and frequency did not improve model fit.6

**TABLE 3. Summary of best mixed effects regression model for Deleted vs Realized (overall percentage deleted 20.6%) (n = 3638)**

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates (standard error)</th>
<th>Number of tokens and (percentage deleted) for factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-3.64*** (0.22)</td>
<td></td>
</tr>
<tr>
<td>Consonant (vs /t/)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/d/</td>
<td>1.13*** (0.18)</td>
<td>1758 (31)</td>
</tr>
<tr>
<td>Following Segment (vs Pause)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consonant</td>
<td>2.91*** (0.15)</td>
<td>1454 (43)</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.84*** (0.19)</td>
<td>714 (10)</td>
</tr>
<tr>
<td>Syllable Stress (vs Stressed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstressed</td>
<td>1.26*** (0.29)</td>
<td>380 (24)</td>
</tr>
<tr>
<td>City (vs Memphis)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durham</td>
<td>-1.17*** (0.27)</td>
<td>1206 (18)</td>
</tr>
<tr>
<td>DC</td>
<td>-0.50 (0.26)</td>
<td>1148 (20)</td>
</tr>
<tr>
<td>Interactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consonant=d : City=Durham</td>
<td>0.95*** (0.26)</td>
<td></td>
</tr>
<tr>
<td>Consonant=d : Realization=DC</td>
<td>0.30 (0.25)</td>
<td></td>
</tr>
<tr>
<td>Consonant=d : Stress=Unstressed</td>
<td>-1.71*** (0.36)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *p < .05; **p < .01; ***p < .001. Other predictors tested in model buildup include Utterance Position, Speaker Sex, Word Frequency, Number Syllables in Word.

**Consonant Realization Summary**

Within the three cities in this study, glottal stop replacement is a frequent variant for both /d/ and /t/. It is worth noting that within this sample of 36 speakers of three AAE varieties, /d/ glottalization is present in every speaker. Whereas devoicing is often seen as a categorical process in other languages, in AAE, the final stop realization is clearly more probabilistic, with speakers realizing coda /t/ and /d/ as glottal stops, coronal stops, or deleted.
Consonant deletion is not unique to AAE, and likely results from connected speech processes especially when another consonant follows (Nolan, 1996; Temple, 2014). In the consonant cluster reduction literature, though, it is clear that AAE varieties often show high rates of deletion, especially in pre-vocalic or pre-pausal environments (Thomas & Bailey, 2015). In addition to internal effects, social factors are significant for /t/, showing that /t/ deletion is more common in Memphis than DC and Durham. Regarding /d/ deletion, only internal factors are included in the best model, showing that connected speech processes are a major factor in the deletion of /d/, and not social categories like city or sex. Glottal stop replacement of /d/ is an unmarked feature in terms of sex differentiation.

The fact that similar factors appear to work on glottal stops and coronal stops for both /t/ and /d/ suggests that there are similar underlying processes, though the rate at which glottal stop replacement occurs is much higher for /t/. As a feature of AAE, /d/ glottalization is geographically widespread (Fasold, 1981), though it does appear in other American English varieties (see Wolfram and Christian, 1976). With /t/, glottalization is so frequent across all varieties of MAE that it is less commented on as a feature, except in certain regional areas (Roberts, 2006).

Similar effects were also found on realized versus zero coda forms across both /t/ and /d/. Deletion of the final consonant, unsurprisingly, is most common when there is a following consonant and much less frequent when there is a following vowel. Since these are consonant singletons, outright deletion would be unexpected for many varieties of MAE and even AAE (Wolfram, 1969). However, especially for /d/, deletion is possible when there is a following pause. Extensive work on consonant cluster reduction in AAE speaking communities (e.g., Wolfram, 1969, 1974) has also shown that deletion is more common in pre-pausal and pre-
vocalic position in AAE than in MAE varieties. Our finding on deletion suggests that there are systematic internal differences that are shared across the examined varieties, including by syllable stress and following segment, but for /t/ in particular, a significant effect for city was also found.

Because neutralized realizations of final /t/ and /d/ account for a majority of the data, the next step in the analysis is to look at whether consonant voicing distinctions are maintained across neutralized tokens, such that a listener could determine consonant voicing based on the duration of the preceding vowel.

Vowel Duration Differences

While there are various cues available for consonant voicing distinctions (Kharlamov, 2012), this study focuses on vowel duration, as the difference in vowel duration between voiced and voiceless consonants has been found to be more distinct for African American speakers when compared to European American speakers in citation form (e.g., ‘bead’ [bi:d] compared to ‘beat’ [bit]) (Holt et al., 2016). Table 4, which collapses across vowel qualities and utterance position, shows raw vowel durations by phonetic realization for each consonant. In comparing the realizations across underlying consonants, it is the case that all /d/ realizations have longer durations than /t/. In the current analysis, log transformed duration (in milliseconds) values are used to transform the duration data from a right-skewed distribution into a more normal-like distribution.

<table>
<thead>
<tr>
<th>Consonant</th>
<th>Realization</th>
<th>n</th>
<th>Mean Duration (ms)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>/t/</td>
<td>Glottal Stop</td>
<td>1290</td>
<td>176.96</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>Zero Coda</td>
<td>178</td>
<td>145.11</td>
<td>0.069</td>
</tr>
</tbody>
</table>
The primary interaction of interest is between consonant type and phonetic realization. Figure 7 shows the log duration by phonetic realization, grouped by underlying consonant and utterance position. What is apparent in the figure is that, in general, vowel duration before /d/ is longer than before /t/ for each realization, as is expected in English. Within each consonant type in medial position, it is surprising that the glottal stop realization is longer than zero codas and coronal stops. In utterance final position, vowels before /d/ are not different by realization, but for /t/, vowels before zero coda appear longer, though only eleven tokens are represented.
There are several factors that could affect the vowel duration, including the underlying following consonant (t, d), the phonetic realization (glottal, deleted, full stop), as well as vowel type. Vowel quality differences such as vowel height, vowel tenseness, and vowel glide quality all affect the intrinsic duration of a vowel, and are tested in model building as fixed effects. For example, tense vowels will be longer than lax vowels, low vowels will be longer than mid vowels (which are longer than high vowels), and diphthongs will be longer than monophthongs (Holt et al., 2015). Utterance position is also expected to have an effect because of phrase final lengthening—syllables at the end of intonation phrases will be longer than those within an
intonation phrase—so it is included in the model (Thomas, 2011). Consonants on unstressed syllables (e.g., *wicked, bucket*) are excluded from the vowel duration analysis, but remain in the realization analysis.

The best fit model for vowel duration is summarized in Table 5. There is a significant main effect of consonant type, such that vowels before */d/* are longer than vowels before */t/*. This effect is an expected one, based on what is known about English (e.g., Yu, 2011) and the duration patterns found for the effect voicing across languages (Solé, 2007). Beginning with main effects that are not included in any interactions, the model identifies that there are expected effects for inherent duration differences for Vowel Height (such that low vowels are longer than mid vowels and high vowels) and Vowel Tenseness (tense vowels are longer than lax vowels). Vowel diphthongality was not a significant predictor, and thus, not included in the model. Additionally, females have longer vowel durations than males, which has been found to be a consistent effect in studies of vowel duration in American English (Hillenbrand, Getty, Clark, & Wheeler, 1995; Holt et al., 2015; Jacewicz et al., 2007), though Deser (1990), a study of Detroit AAE, did not find this pattern.

**Table 5. Summary of best mixed-effects model with interactions for Log Duration (overall mean duration in milliseconds (ms) 192.4) (n = 3258)**

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates (standard error)</th>
<th>Number of tokens (mean duration in ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>5.26*** (0.04)</td>
<td></td>
</tr>
<tr>
<td>Consonant (vs t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>0.48*** (0.03)</td>
<td>1492 (222.4)</td>
</tr>
<tr>
<td>Realization (vs Glottal Stop)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero Coda</td>
<td>0.12* (0.06)</td>
<td>657 (175.1)</td>
</tr>
<tr>
<td>Coronal Stop</td>
<td>0.04 (0.04)</td>
<td>766 (179.4)</td>
</tr>
<tr>
<td>Utterance Position (vs Final)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial</td>
<td>-0.29*** (0.02)</td>
<td>1984 (164.9)</td>
</tr>
<tr>
<td>Vowel Height (vs Low)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>-0.10*** (0.03)</td>
<td>877 (187.8)</td>
</tr>
<tr>
<td>High</td>
<td>-0.19*** (0.03)</td>
<td>958 (180.0)</td>
</tr>
</tbody>
</table>
Regarding significant interactions, there is a relationship between underlying consonant type, realization, and utterance position (pictured in Figure 7). In medial position, vowels before zero coda and coronal stops are shorter than vowels before glottal stops for both /t/ and /d/. For interactions with consonant type, there is a relationship where vowels before zero coda are shorter than before glottals regardless of utterance position for /d/ compared to /t/, while coronal stop realizations do not share a significant interaction with consonant. While utterance effects are expected as a result of utterance final lengthening, which increases the duration of the final syllable, the interactions are not necessarily expected. In any case, vowel duration in utterance final position does not vary by phonetic realization the way it does in medial position (Figure 7).

To summarize, when factors that affect vowel duration are controlled for, there is still a statistically significant pattern in Figure 7, with vowels before glottal stops being longer in utterance medial position.

**The Relationship between vowel duration and consonant realization**

Earlier, internal variables such as following segment and syllable stress were found to be predictive of final consonant realization. Other than the larger proportion of glottal stops for /t/ overall, the internal factors were in the same direction. Since glottal stops, coronal stops, and

<table>
<thead>
<tr>
<th>Vowel Tense</th>
<th>(vs Tense)</th>
<th>1720 (191.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lax</td>
<td>-0.07** (0.02)</td>
<td>1538 (193.5)</td>
</tr>
<tr>
<td>Sex</td>
<td>(vs Male)</td>
<td>1697 (178.9)</td>
</tr>
<tr>
<td>Female</td>
<td>0.10* (0.05)</td>
<td>1561 (207.1)</td>
</tr>
</tbody>
</table>

**Interactions**

| Consonant=d : Realization=Zero Coda | -0.13*** (0.04) |
| Consonant=d : Realization=Coronal Stop | -0.02 (0.03) |
| Consonant=d : UttPosition=medial | -0.06 (0.03) |
| Realization=Zero Coda : UttPosition=medial | -0.22*** (0.05) |
| Realization=Coronal Stop : UttPosition=medial | -0.19*** (0.04) |

Notes: *p < .05; **p < .01; ***p < .001. Other predictors tested in model buildup include City, Word Frequency, Vowel Diphthongality.

*a Model used log duration (in ms)*
zero coda are common across all speakers across the three cities, vowel duration was
investigated as a cue for consonant voicing, which is neutralized in two of the three realizations
of the coronal stops. While there were expected effects for vowel inherent differences like vowel
height and tenseness, there was not an effect for vowel diphthongization, which has been found
elsewhere (Holt et al., 2015). This could be related to the reflexes of the African American
Vowel System and the Southern Vowel Shift, where glide weakened or monophthongal variants
are more common for vowels like /ai/ and /e/ (Thomas, 2007). Speaker sex is a main effect here,
with women exhibiting longer vowels than men, which is in line with previous work on AAE
(Holt et al., 2016) and other varieties of American English (Jacewicz et al., 2007). A consistent
difference between the consonants is that vowels before /d/ are longer than vowels before /t/ (cf.
Solé, 2007). In utterance medial position, vowels before glottal stops are longer than coronal
stops, but this trend is not apparent in final position.

Ultimately, though, the vowel duration results point to the fact that in AAE, vowel
duration could be becoming a primary cue for consonant voicing, especially when the consonant
is realized as a glottal stop. Given the claims about vowel duration phonologizing in MAE (Yu,
2011), our finding shows that like other varieties of American English, AAE maintains the
expected distinctions between consonant voicing. The fact that AAE neutralization results from
deletion of the consonant and glottal stop replacement suggests that AAE varieties are slightly
further along in phonologization of vowel duration, because the voicing of the consonant does
not need to be apparent since the vowel stably reflects underlying voicing.

**DISCUSSION**
The current analyses add to our understanding of consonantal variation in Southern AAE varieties. The first set of analyses showed that the distribution between three realizations, glottal stops, coronal stops, and zero codas (i.e., deleted stops) exhibit a similar pattern across the three cities under investigation. This similarity is especially true for Durham and Washington DC, which could result from historical population movements. Memphis exhibits more coronal stop realizations throughout the data. Since glottalization of word final /d/ in AAE is found throughout the country, extending this kind of analysis to other cities would be an important next step. These three cities also share similar demographics, that is, large African American populations, so extending the analysis to cities with smaller African American populations would also contribute to understanding the viability and role of an ethnic marker where there are fewer speakers within the same ethnicity. Evidence from Farrington (2011) suggests that cities with such demographic differences exhibit much less glottalization of word final /d/. /t/ glottalization, on the other hand, is in widespread usage among both MAE and AAE speakers. Though rates of /t/ glottalization between ethnic varieties of English have not yet been investigated, it makes sense that glottalization would be the most frequent realization of /t/ in the data.

While work has begun to look at regular variation of /t/ glottalization in terms of storage and representation, the inclusion of AAE varieties would add an important layer to that body of research. For /t/, Sumner and Samuel (2005) found distinctions between immediate activation on the semantic level and long-term activation, such that regular variation does not disrupt activation on the short term but does on the long term. For varieties, like AAE, that have glottalization of word final /d/ in addition to /t/, does vowel duration become the main cue in terms of semantic activation, or are there feature mismatches based on voicing? Additionally, recent work by Chong and Garellek (2018:2) on the online perception of glottalized coda stops in
American English suggests that glottalized variants of voiced sounds “are not attested in American English”. They find that glottalization is not associated with voiced stops because of slower recognition of target words. Inclusion of AAE speaking listeners, who do exhibit glottalization of voiced stops, could help to either reinforce conclusions, or introduce variability that could help researchers develop more complex, but more broad-reaching, theoretical phonological models.

One striking result from this study is the fact that each speaker across the three regional sites exhibits a range of variation for final /d/, including the three kinds of phonetic realizations analyzed here: coronal stops, glottal stops, and deleted stops. On top of that, the lack of sex differences in the statistical analysis for type of phonetic realization indicates that there aren’t differences between men and women.

Regarding consonant deletion (zero coda forms of /t, d/), pre-pausal deletion is not frequent, nor widespread (Harrison, 2007), but there is deletion in pre-consonantal environments. This deletion can be attributed, in part, to connected speech processes, in the same way that consonant cluster reduction occurs in all varieties of American English, especially when a consonant follows (Thomas & Bailey, 2015). A closer look teasing apart connected speech processes from higher deletion rates resulting from dialect differences is warranted. Both deletion and glottalization of word final /d/ have been found to index sex differences in Detroit AAE (Wolfram, 1969). In Wolfram’s (1969) work, men are more likely than women to use zero coda variants, while women are more likely to use glottal variants of /d/. In the current analysis, sex was not a significant predictor in any model for any of the consonantal analyses. This suggests that at least for younger speakers in the South, sex differences in /t/ and /d/ realizations
are not as relevant, which highlights the possible use of these variants by both sexes as an index of ethnic identity.

The finding that vowel duration varies by consonant adds to the existing literature on contrast reduction and incomplete neutralization by strictly using conversational data. Kharlamov (2012) has shown that some incomplete neutralization effects result from lab methodologies (e.g., whether words are read, whether minimal pairs are included in the data). The fact that differences are significant in the conversational setting suggests that it is a cue that listeners could actually use in everyday communication to differentiate consonant type if a word was ever ambiguous.

Vowel duration contrasts are generally maintained for the oral coronal consonants, with consistent distinctions between underlying /t/ compared to underlying /d/, regardless of whether the consonants surface as a fully articulated coronal stop or a glottal stop. When the consonant is deleted, the contrast is reduced, but still significant. In the Urban South (e.g., Atlanta), some AAE speakers also exhibit glottal stop replacement or elision of the velar stop consonants, /k, g/ (Harrison, 2007). In addition to duration, it is likely that vowel quality is also a cue for speakers in differentiating consonants. For example, in the framework of phonological rules, vowel processes resulting from voicing occur prior to glottal stop replacement. Take, for example, the /ai/ diphthong, as in bite and bide. In AAE varieties, the non-pre-voiceless /ai/ (in bide) undergoes monophthongization (Thomas, 2007), while the pre-voiceless /ai/ does not. When both /d/ and /t/ become glottal stops, the vowel quality difference is maintained, making bide [baːʔ] and bite [buɪʔ]. With velar consonants, it is possible that vowel quality differences are maintained prior to glottal stop replacement. Future work will investigate not only the relationship between velar and coronal stops, and the cues that condition each category, but also
the relationship between vowel quality, vowel duration and underlying consonant type. Determining how all of these pieces fit together will help us better understand the complex role of vocalic and consonantal variation in AAE.

CONCLUSION
The results of this study show that coda coronal (oral) stops /t, d/ have variable phonetic realizations, including, most prominently, the glottal stop. Because of the apparent widespread nature of /d/ glottalization in AAE, speakers from three cities, Memphis, Durham, and Washington DC, were included in the analysis, all of whom replace word final /t/ and /d/ with glottal stops to an extent, and more frequently for coda /t/. Internal factors like following segment and syllable stress account for much of the variation, while social and dialect differences (i.e., main effects for city) are more limited within the data. Consonant elision can partially be attributed to connected speech processes (Temple, 2014), but there is some coda, pre-pausal consonant elision that warrants further investigation. After determining the effect of factors on phonetic realizations of final /t/ and /d/, I looked at vowel duration as a cue for consonant voicing. When comparing the duration of vowels before voiced and voiceless consonants, there were significant differences in the expected direction, with longer durations before underlying /d/ consonants compared to underlying /t/. In sum, AAE exhibits coda stop neutralization through glottal stop realization and deletion, the result of which is incomplete neutralization, as vowel duration is a prominent cue for consonant voicing.
NOTES

1. Although the information about glottalization comes from literature considering the glottal stop an allophone of /t/, there have not yet been any studies on differences between [ʔ] as an allophone of /t/ and [ʔ] as an allophone of /d/. In the current analysis, a glottal stop replacing /d/ is treated as the same glottal stop that also replaces /t/.

2. In model selection, for both logistic and linear regressions, more complete models were compared to less complete models with likelihood ratio tests, using the anova function in R (Baayen, Davidson, & Bates, 2008). Starting with a simple model with only random effects, factors were added and included if they significantly improved the model fit ($P = .05$), with additional comparisons of AIC and BIC values.

3. After comparing different ways to code following segment (e.g., a finer grained separation of following consonants into obstruents, liquids, glides) in the statistical analysis, a simpler coding scheme of following pause, vowel, or consonant, was used, following earlier AAE consonant cluster reduction studies (e.g., Wolfram, 1974).

4. Speech rate as a variable of interest is not included in the current analysis although vowel duration can be expected to be a function of speech rate (Kendall, 2013). In our statistical analyses, including speaker as a random effect helps to control for consistent differences in speech rate across speakers. A closer look at how speech rate interacts with consonant realization and vowel duration is left for future work.

5. In order to calculate lexical frequency, the SUBTLEXus corpus (Brysbart & New, 2009) was used. Following Forrest (2017), I implemented a base-10 logarithmically transformed measure for use in model build-up. For a handful of lexical items (e.g., Northside, Dipset, Fleetwood), a value of 1 was assigned as a value for inclusion in the analysis.
6. In this model buildup, the log lexical frequency came closest to improving model fit ($p = 0.08$). An effect would be compatible with findings on consonant cluster reduction and lexical frequency (e.g., Bybee, 2002). Future work incorporating function words into the analysis may shed light on whether frequency has an effect on neutralization.

7. Log duration values were modeled using a linear mixed-effects regression analysis with lmerTest (Kuznetsova, Brockhoff, & Christensen, 2016). All reasonable two-way interactions were tested, but no higher order (e.g., three-way) interactions. Underlying assumptions of the final linear model were tested, including confirmation that the distribution of the residuals had a mean of zero and residuals did not exhibit overdispersion.
References


Incomplete Neutralization in AAE


