JASA ARTICLE

# Variation in global and intonational pitch settings among black and white speakers of Southern American English[a)]

Aini Li,[1] Ruaridh Purse,[2] and Nicole Holliday[3,b)]
[1]Department of Linguistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
[2]Department of Linguistics, University of Michigan, Ann Arbor, Michigan 48109, USA
[3]Department of Linguistics and Cognitive Science, Pomona College, Claremont, California 91711, USA

**ABSTRACT:**
This article revisits classic questions about how pitch varies between groups by examining global and intonational pitch differences between black and white speakers from Memphis, Tennessee, using data from read speech to control for stylistic and segmental variables. Results from both mixed-effects regression modeling and smoothing spline analysis of variance find no difference between black and white men in mean $F_0$ and pitch range measures. However, black women produced consistently lower mean $F_0$ than white women. These findings suggest that while pitch patterns in black women's speech remain underexplored in the literature, they may play an important role in shaping attitudes and ideological associations concerning black American speakers in general. Moreover, vocal pitch may be a linguistic variable subject to variation, especially in a context of racialized and gendered linguistic standards.
 https://doi.org/10.1121/10.0014906
(Received 20 January 2022; revised 28 September 2022; accepted 4 October 2022; published online 3 November 2022)
[Editor: Matthew B. Winn] Pages: 2617–2628

## I. INTRODUCTION

A long line of research in ethnic identification[1] has centered around the discriminability of African American and European American voices. Understanding what features listeners are able to access while making ethnic identification of voices has significant implications for both speech communication and social justice. For instance, racial and ethnic profiling can be used to deny a candidate job or residence appointments (Purnell *et al.*, 1999). Previous studies have demonstrated the use of multiple cues to distinguish African American voice from European American voice, such as vowel quality (Thomas and Reaser, 2004) and voice quality (Purnell *et al.*, 1999). Among these different types of cues, one of the most heavily discussed is $F_0$: whether African American speakers and European American speakers fundamentally employ different pitch values in some contexts. Here, $F_0$ refers to the repetition rate of voiced speech signals, reflecting vocal fold vibration rate that is heard as pitch. Studies of $F_0$ are motivated by a long-standing folk linguistic notion that black Americans, and particularly black men, have lower-pitched voices than American speakers of other races. This notion is exemplified in metalinguistic discourse among black Americans about "bass" in the voice as a stylistic resource to perform blackness, which also appears to be linked to an assertive and unapologetic stance. Holliday (2016, 57) reports that "bass" is explicitly identified by some black biracial speakers as a feature of their speech and the speech of their black fathers that is eschewed in certain contexts, such as the workplace. This demonstrates the utility of vocal pitch for constructing racialized identity within a group while being conscious about variations in how pitch may be perceived by different groups of interlocutors. Relatedly, a growing body of work examines an ideological link between African American English (AAE) and the performance of certain types of idealized "masculine" personae (e.g., Barrett, 1998; Bucholtz, 1999; Sneller, 2020), which are linked to low $F_0$ due to a well-described connection between vocal pitch and gender identity, from both deterministic and constructivist perspectives (Ohala, 1984; Zimman, 2018).

Despite the shared interest in exploring pitch behavior as a salient impressionistic difference across racial groups, investigations into the empirical basis of pitch differences between black and non-black Americans are not uniform in their results (e.g., Thomas, 2015; Thomas and Reaser, 2004). While several studies find that black men employ a lower overall $F_0$ compared to white men, others claim that the locus of this effect is an overall wider pitch range (Hudson and Holbrook, 1981), and still others fail to find any difference (Walton and Orlikoff, 1994). As we outline in more detail in Sec. I A, in terms of the aims and rationales of these studies, this lack of consensus may be partially due to differences in the measures, measurement sites, and linguistic content comprising the data in these studies. Moreover, there is a paucity of work exploring $F_0$ differences between black and white women, especially as a controlled comparison to men.

[a)]This paper is part of a special issue on Reconsidering Classic Ideas in Speech Communication.
[b)]Electronic mail: nicole.holliday@pomona.edu

We combine both mixed-effects regression modelling and SSANOVA (smoothing spline analysis of variance) techniques, probing the pitch level, pitch range, and intonational contours of different ethnic and gender groups in identical read phrases. We find that differences in pitch level between black and white speakers are limited to the women, with black women using consistently lower pitch than white women. This suggests women's speech plays an important role in ideologically connecting AAE with low-pitched voices and more generally with hegemonically "non-feminine" traits. Indeed, the fact that performances of binary sex are defined exclusive of black women in the hegemony is an example of how black women experience "invisibility" when it comes to prototypical and stereotypical conceptualizations of both race and gender (Babbitt *et al.*, 2018; Coles and Pasek, 2020; Johnson *et al.*, 2012; Morgan, 2005). However, SSANOVA analyses on individual phrases show that black men and white men differ in various contour-related properties, such as peak delay and the number of peaks. Taken together, our methods and findings mark an important turning point in research on the role of $F_0$ in ethnic identification. In reconsidering the unfortunate earlier practice of comparing only male speakers, and in tightly controlling the linguistic and stylistic content of utterances, we can shed a new light on the relationship between race and pitch, especially in terms of their connection to speaker gender. This also represents an important step in redressing the aforementioned invisibility of women in research on sociophonetic differences between racial groups.

### A. Pitch and race

Studies over the past several decades have sought to investigate pitch differences between black and non-black American men, but results have been mixed (Thomas, 2015; Thomas and Reaser, 2004). Although many of the previous studies have not been explicit about how they have defined their racial categories, we employ U.S. Census definitions for the current study. That is, when we discuss black speakers in general terms, we define them as "A person having origins in any of the black racial groups of Africa." Similarly, we define white speakers as "A person having origins in any of the original peoples of Europe, the Middle East, or North Africa" (Holt and Bent, 2019). Importantly, however, for the purposes of the current study, all speakers are classified by race based on their own self-identification. The impetus for many of the previous studies has frequently been to better understand variation between black and white speakers, and how listeners may attune to different types of variables in linguistic profiling and discrimination. On the one hand, a number of older studies have claimed that African Americans, or at least African American males, tend to produce a lower overall fundamental frequency ($F_0$) than their European American counterparts (Hawkins, 1993; Hollien and Malcik, 1962; Hudson, 1977; Hudson and Holbrook, 1981; Wheat and Hudson, 1988). To explore how age, sex, race, and the type of speech activity influences measures of fundamental frequency, Hudson and Holbrook (1981, 1982) conduct a study to compare the fundamental frequency characteristics of young black adults across two contexts: spontaneous speech and reading. They find that not only do African American men have a significantly lower fundamental frequency but they are also inclined to display larger mean frequency ranges, even though some studies indicate that $F_0$ in read speech might be higher than in spontaneous speech (e.g., Richardson, 1973; Snidecor, 1943). However, the finding that black speakers have a lower mean fundamental vocal frequency is not based on analysis of comparable white counterparts, but through simply relating with some previously published data for white speakers. In contrast, some studies have failed to replicate these differences, finding that black and white speakers do not differ in $F_0$ (e.g., Walton and Orlikoff, 1994). Instead of the characterization of black voices, Walton and Orlikoff (1994) conduct a study where listeners are asked to determine the race of the speaker based on vowel samples produced by both black and white speakers to address the issue of speaker race identification. Despite the fact that listeners are 6 speech pathologists who are of various origins in terms of race (both European American and African American are involved), the speech samples are taken from 100 male (50 black and 50 white) prison inmates. In addition to analysis of identification accuracy, they also conduct an acoustic analysis of these vowel productions from both black and white male speakers. Their acoustic analysis shows that black speakers have greater frequency perturbations and a lower harmonics-to-noise ratio than the white speakers, but no significant differences in the mean fundamental frequency were reported between the two groups. With regards to the role of frequency in ethnic identification, Richardson (1973) examines whether listeners can identify the race of black and white speakers from recordings of both read and spontaneous speech manipulated using high- and low-pass filters. He finds that identifying the race of the speaker is associated with the cut-off frequency of filtering. The voices of the black speakers are most identifiable when lower frequencies are more easily heard [i.e., when frequencies lower than the cut-off frequency (510, 1020, 2040 Hz) are attenuated], while the voices of the white speakers are more identifiable than black speakers when higher frequencies are more easily heard [i.e., when frequencies higher than the cut-off frequency (510, 1020, 2040 Hz) are attenuated].

Rather than just average pitch level, the expectation of vocal pitch differences between black and non-black Americans may be driven by some particular elements of how pitch is implemented at a phrasal level. Indeed, despite the fact that prosodic variation has not been the subject of heavy empirical scrutiny in sociolinguistic literature until recently (e.g., Clopper and Smiljanic, 2011; Holliday, 2016), it has been widely acknowledged that prosodic features can be sociolinguistically meaningful, differing between speakers of different racial backgrounds (Holliday, 2016; Purnell *et al.*, 1999; Thomas and Reaser, 2004). This includes variables like word stress placement (Baugh, 1983;

https://doi.org/10.1121/10.0014906

Sutcliffe, 2001), speech rate (Kendall, 2013), and prosodic rhythm (Thomas and Carter, 2006), but also a number of intonational variables related to the pitch contours that speakers employ. Holt and Rangarathnam (2018) find differences between black and white speakers in $F_0$ declination and reset as a function of breath grouping, though they do not specifically address mean $F_0$ differences. Additionally, McLarty (2018) investigates similarities and differences between AAE and EAE (European American English) intonation with a particular focus on the differences in type and relative frequency of different types of pitch accents, which are the primary way that English marks prominence. Through comparisons between modern varieties of AAE and EAE to varieties of archival AAE and EAE, he observes that AAE speakers of both generations tend to produce more $F_0$ falls followed by rises to high targets, than their EAE speaking counterparts. Later, Holliday (2019) examines the use of different boundary tones across three question types using data from the Corpus of Regional African American Language (CORAAL). Even within the speech of AAE speakers, she finds significant variation in the realization of questions. These studies lend further support to the idea that prosodic variation can be used to signal racial differences and likely plays a major role in ethnic identification as well.

Possible explanations exist for the mixed results from studies investigating pitch behavior in different racial groups. To begin with, different studies used different tasks, adopted different analysis techniques, and focused on different measurements. Additionally, these contrasting results may highlight potential difficulties in pinpointing the locus of group differences, and the importance of considering multiple relevant social dimensions. Notably, there far fewer studies exist investigating black women's speech (Ducote, 1983). Given the robust link between pitch and gender identity, a consideration of the role of gender is a notable lacuna in the literature that promises interesting results. For instance, Hudson and Holbrook (1982) use a fundamental frequency analyzer (FLORIDA I) to compare absolute, as opposed to proportional, F0 differences between black men and black women and find that in addition to the mean differences, black men have a narrower $F_0$ range and smaller pitch excursions than black women. These results further validate studies like this one that examine both race and gender differences in pitch in a controlled fashion. The present study contributes to the literature on prosodic difference between black and white speakers, exploring both global pitch parameters and the intonational contours of key phrases.

### B. Pitch and gender

Vocal pitch generally differs by talker physiology. On average, the $F_0$ of male speakers is lower than that of female speakers. This physiological basis gives rise to a widespread psychological association between vocal pitch on one hand, and sex and gender on the other, such that a lower pitch is associated with male (or masculine) voices and a higher pitch with female (or feminine) voices (Cartei *et al.*, 2014; Feinberg *et al.*, 2005; Pisanski and Feinberg, 2013; Pisanski and Rendall, 2011). As such, listeners in previous studies have been generally able to accurately categorize speakers in terms of sex (Bachorowski and Owren, 1999) and, given certain cisnormative and gender-binary assumptions, gender. Unfortunately, many earlier studies discussed here have conflated sex and gender as well as made assumptions such that they were directly correlated. The majority of these studies have also excluded made no reference to non-binary identities. While a discussion of queer identities and pitch is beyond the scope of the current data and analysis, it is the case that many of the traditional findings that this paper aims to address were underinformed and underspecified with respect to their treatment of gender. As a result, we encourage the reader to keep in mind that the earlier studies may not have provided sufficient information about the gender identities of their participants. The current study utilizes the findings of those earlier works, but recognizes these limitations and aims to contextualize them properly.

Aside from issues of conflating sex and gender, such deterministic physiological explanations only account for a small amount of actual variance in vocal pitch. Pitch is also used as a sociolinguistic resource, and especially as a resource to perform gender identity. This fact is evident from a number of observations (Zimman, 2018). For example, pitch differences are already present between prepubescent boys and girls (Ferrand and Bloom, 1996; Hasek *et al.*, 1980; Ingrisano *et al.*, 1980). In addition, the magnitude of pitch differences between male and female speakers seems to be culturally specific. To take one example, Japanese men produce a lower average $F_0$ than American men, and Japanese women produce a higher average $F_0$ than American women, which has been suggested to indicate a difference in the sociosemiotic function of pitch to perform gender within different racialized speech communities (Loveday, 1981; Yuasa, 2008). Further, speakers have the capacity to dynamically manipulate pitch in the course of orienting their identity according to the context and audience of a particular speech act. For example, one group of heterosexual speakers engaging in speed-dating were found to converge in pitch according to the perceived attractiveness and likeability of their interlocutor (Michalsky and Schoormann, 2017). Not only does this demonstrate a capacity to control pitch variation, but the results point to an attenuation of the typical differences in pitch between men and women.

Beyond a more global manipulation of mean pitch, speakers consistently exhibit the ability to manipulate the pitch and spectral properties of their voice to create intonational contours. Global manipulation here refers to the general measurement that captures the phrase level mean $F_0$ of all the phrases included in the reading passage. The properties of these contours, too, are linked to the performance of gender identity. Crucially, previous work on gender tends to focus primarily on white speakers, while studies that have looked at race do not often treat gender as a potential

variable. For instance, McConnell-Ginet (1983) claims that masculine speech is often associated with a flattened pitch contour, while feminine speech is associated with a relatively wide pitch range and rapid changes. Much of her reasoning for these associations comes from impersonations of stereotypical masculine and feminine speech because only white speakers are included. Evidence from impersonations also reveals an association between femininity, particularly young white American femininity, and certain stigmatized intonational variables like creaky voice and "uptalk" (Slobe, 2018). At the same time, several studies have found that lower-pitched voices are cross-linguistically evaluated as more credible and attractive, particularly for male speakers (e.g., Gasser *et al.*, 2019; Michalsky and Schoormann, 2017). Thus, not only is pitch used in different ways by speakers of different genders, but the way its use is interpreted is also influenced by listener assumptions about a speaker's perceived gender identity.

### C. Gender and race

There is no single individual experience or behaviour that can be fully defined or accounted for based on only one social category (e.g., "woman" or "black"). Likewise, neither gender nor race can be fully understood without the other. A speaker's performance of gender is informed by their race, and vice versa, in a way that is intersectional (Calder and King, 2020; Crenshaw, 1989; Hooks, 2014; Levon, 2015). However, at a more basic level, the perception and evaluation of linguistic performances of race and gender are affected by pervasive racial stereotypes and understood through the lens of hegemonic gender roles, predominantly informed by white perspectives. To take a relevant example, in the English-speaking world, AAE and other forms of black language are commonly ideologically associated with traits attributed to urban blackness: toughness, danger, and street smarts, as well as coolness (Bucholtz, 2011; Bucholtz and Lopez, 2011; Sneller, 2020). As a consequence of racism, traditional ideas about black men, including stereotypes of physical strength and violence have long given rise to an entrenched association in many people's minds, between black men and a particular type of hegemonic masculinity (Collins, 2004).

Further reinforcing this association, AAE is frequently seen as a useful resource for non-black men to perform this type of masculinity. Because of these entrenched interacting ideologies between gender, race, and language, black men's speech patterns are often resorted to, and considered appropriate, by non-black people to perform masculinity. A substantial body of work has indeed demonstrated how this works in different communities. In her analysis of how a middle-class European American boy who affiliates with African American youth culture tells narratives about interracial conflict, Bucholtz (1999) finds that the narrator tends to use elements of AAE to project an urban youth identity that is influenced by African American youth culture and is associated with urban black masculinity in particular. The use of AAE as an expression of masculinity is reinforced by how features are used in concert. When indexing an African American speaker, the speech of the narrator tends to speak slower, with a lower-pitch, and with reduced pitch range. Similarly, Cutler (1999) observes a white upper-middle class New York City teenager to employ elements of AAE, ranging from segmental features, such as stop pronunciation of inter-dental fricatives, to prosodic features, such as vowel lengthening. These can also be understood as attempts to participate in the complex prestige of African American youth culture embodied as masculine. A further example is found in the context of dominant discourses where Asian Americans have been distinctly positioned with respect to African Americans and European Americans (Chun, 2001, 2013). However, little attention has been paid to the use of female speech in the construction of masculinity and/or femininity. This therefore further raises questions about the role of black female speech in such gender-based performances.

In addition to AAE's utility as a resource for indexing a particular type of masculinity for non-black speakers, black speakers can also draw on these ideological associations in creative ways. For instance, Barrett finds that black drag performers may exaggerate a stylistic contrast when flouting a hegemonically feminine "white woman" persona (Barrett, 1998). In particular, Barrett (1998) describes that African American drag queens are inclined to use AAE as a marked choice, particularly an accentuated drop in pitch to emphasize the act of gender play and links with ideas about physiological masculinity. This association between AAE and pitch is further reinforced by the qualitative impressions of black speakers about their own speech. For instance, Holliday (2016) finds that some speakers report a sense that they "put more bass" in their voices while talking with black interlocutors, even if quantitative analysis does not always confirm these intuitions.

## II. THE CURRENT STUDY

Building on the questions raised in earlier works, in this study, we investigate the $F_0$ differences in a corpus of read speech by black and white speakers in Memphis, Tennessee. Although data derived from spontaneous speech, such as sociolinguistic interviews, may be more naturalistic, it presents difficulties for both segmental and prosodic analysis, and introduces potential questions about stylistic variation and the stylistic repertoires of different groups. For instance, if black and white speakers indeed employ their pitch patterns differently, it is hard to rule out that these differences could be partially driven by the fact that speakers with different linguistic repertoires employ pitch patterns differently during style-shifting when they interact with different interlocutors (Benor, 2011; Tarone, 1973). Here, we take read speech as a point of departure for disentangling some of the complicated issues related to how gender, race, and intonation are correlated, without making claims about other types of speech. This has the advantage of allowing us to focus on findings about $F_0$ properties without possible

https://doi.org/10.1121/10.0014906

confounds of other types of features. Additionally, the use of a corpus of read speech allows us to control for phrase content and length, and permits us to use data from a larger number of speakers. As an exploratory analysis, the goals of our study are twofold: First, we examine the variation of $F_0$ by race and gender within and across phrase boundaries. By focusing on speakers with different racial backgrounds from the same geographical area (Memphis, Tennessee), we explore whether black speakers (men and women) differ from their white counterparts in terms of their use of $F_0$ at different points in comparable phrases through a set of identical read speech declaratives. Second, we investigate whether they implement pitch contours differently in these phrases using SSANOVA analysis. With these goals in mind, we aim to address the following questions:

Do black speakers, both men and women:

(1) use lower/higher $F_0$ than their white counterparts?
(2) display a wider/narrower $F_0$ range than their white counterparts?
(3) implement intonational contours differently than their white counterparts?

The current study aims to (1) develop a better understanding of how $F_0$ data should be analyzed via the lens of comparing differences between talker groups who might have otherwise been combined in previous analyses, and (2) promote a reconsideration of the constructs of race and gender in studies on sociophonetic variation. The remainder of the paper is structured as follows: Sec. III introduces the present study's methods, both in terms of quantitatively comparing the general properties of phrases and qualitatively comparing their intonational contours. The results of each of these approaches are reported in Sec. IV. We then discuss some possible interpretations of our findings in Sec. V and offer some tentative conclusions in Sec. VI.

## III. METHODS

### A. Data and speakers

The data analyzed in this study come from recordings of black and white speakers of Southern American English that were originally collected by Valerie Fridland in Memphis, Tennessee. The Memphis corpus itself was collected between 2001 and 2003 to document sound changes in Southern speech in the Memphis area. It consists of recordings of conversational, reading passage, and word list data from approximately the same number of black and white speakers living in the same geographical area[2] [more detailed information about this dataset can be found in (Fridland, 1999, 2001, 2003)]. The reading passage was constructed to systematically investigate the critical vowel class involved in Southern Vowel Shift (Fridland, 2001). Its full text can be found in the Appendix.

A total of 130 recordings of speakers reading the same passage were downloaded from SLAAP (Sociolinguistic Archive and Analysis Project) (Kendall, 2013). Each recording was 2–3 min long. Some recordings were excluded due to lack of demographic information or other issues with file management in the corpus. In the end, we analyzed 94 recordings of read speech from 47 black speakers (male = 33, female = 14) and 47 white speakers (male = 32, female = 15). All the speakers, with only one exception (age = 41 y), were between the ages of 18–30 y at the time of recording (mean = 21.4 y; standard deviation = 4.19). All speaker demographic information, including age, race, and gender, was self-reported by the study participants.

### B. Measurement and analysis

To begin, we identified phrase boundaries according to the original text of the reading passage, and manually aligned them to each recording using Praat textgrids (Boersma and Weenink, 2022). Phrase boundaries were placed regardless of idiosyncratic pauses or disfluencies to maintain maximum comparability between speakers. In the end, we identified 33 phrases with a mean length of 7.9 words. To better capture the $F_0$ differences between black and white speakers, we adopted acoustic measures primarily from Busà and Urbani (2011) by taking into consideration not only mean $F_0$ but multiple measures of pitch range, since previous literature has shown consistently that languages may differ in this parameter (e.g., Mennen *et al.*, 2007, 2012). Following previous practice, measurements for $F_0$ mean (i.e., pitch level), max, min, and standard deviation were automatically extracted from each phrase for each speaker. To maximally guard against $F_0$ aliasing and ensure accurate $F_0$ tracking, manual inspection was performed on each speaker's recording using a Praat script of dynamics $F_0$ setting. This allows for pitch floor and ceiling values to be customized on an individual basis. After the pitch floor and ceiling values for each speaker were decided, $F_0$ extraction was further conducted with the Praat function "To Pitch," again, on an individual basis. All the data were then combined for final statistical analysis. As a whole, the pitch floor for the whole dataset has values more than 70 Hz and the pitch ceiling is no more than 350 Hz.

Finally, we computed a set of long-term distributional measures using the semitone-transformed (ST) $F_0$ max and min to capture pitch span: 4 standard deviations around the mean (SD4), 80% range, 90% range, and 100% range. To be specific, 100% range was calculated by subtracting the $F_0$ min from $F_0$ max. A 90% range was measured as the difference between the 95th percentile and the 5th percentile in terms of $F_0$ values. Similarly, 80% range was calculated by subtracting the 10th percentile of the $F_0$ value from its 90th percentile. Statistical analyses of these data were conducted in the R version 4.0.5 (R Core Team, 2022); linear mixed effects regression models were run using the `lme4` package version 1.1–27.1 (Bates *et al.*, 2014), and plots were created using *ggplot* package version 3.3.5 (Wickham, 2016). To resolve comparisons of interest in the current study, two coding schemes of GLMM (Generalized Linear Mixed Model) were adopted: sum coding and treatment coding. Sum coding is employed to compare each level to the grand

mean (intercept as the grand mean). That is, sum coding is used when a comparison tests whether the mean of a dependent variable for a given level is significantly different from the overall mean of the dependent variable. Treatment coding is used when each level is compared to a reference level (intercept as the cell mean of the reference group). That is, when a given level of a dependent variable needs to be compared to another level (as opposed to the overall mean), treatment coding is employed.

In addition to measuring the general pitch properties of each phrase while controlling highly variable acoustic correlates of stress and accent, we also implemented a SSANOVA analysis. This technique is designed for the comparison of curves along multiple reference points of sound files (Gu, 2013). It has been shown to serve as an effective method for examining contours of syllables, words, and even those over longer time scales, such as phrases (e.g., Morrill, 2015). Following the methodology of Morrill (2015), a 1000-timepoint pitch ($F_0$) contour in the range of between 70 and 350 Hz was extracted from each phrase using the Praat auto-correlation algorithm. Gaps in the contour were interpolated from the points on either side of the gap, and artifacts were removed by smoothing with a bandwidth of 5 Hz. Finally, SSANOVA modeling was implemented with the *gss* package in R (Gu *et al.*, 2014) to plot the contours and allow for qualitative analysis.

## IV. RESULTS

### A. Results from mixed-effects regression

We begin with the quantitative analysis of the general pitch properties of each phrase. To further justify our analysis, if the stereotype that black men (or by extension, black speakers in general) employ a lower $F_0$ than white men (and by extension white speakers) is real, we should expect to see that black speakers employ a lower mean $F_0$ and a narrower pitch range than their white counterparts. Therefore, for each independent measurement, we are interested in whether (1) black men differ from black women, (2) white men differ from white women, (3) women generally differ from men, and (4) black speakers generally differ from white speakers.

#### 1. Pitch level

We start with the results for level (mean $F_0$). Table I presents both the raw mean $F_0$ and median $F_0$ in Hz for both racial groups and across both genders. It is clear that for both black and white speakers, women have higher mean $F_0$ than men. Additionally, white women have an even higher mean $F_0$ than their black counterparts, though this is not the case for white versus black men.

TABLE I. Mean and median $F_0$ in Hz for black and white speakers.

| Speaker groups | Mean $F_0$ (Hz) | Median $F_0$ (Hz) |
|---|---|---|
| Black women | 184.25 | 179.36 |
| Black men | 111.16 | 111.05 |
| White women | 207.49 | 204.07 |
| White men | 112.72 | 110.10 |

A linear mixed-effects model was configured to predict race and gender differences in mean $F_0$, with race (black vs white) and gender (male vs female) as fixed effects (treatment coded in a two-way interaction) and each individual speaker and phrase as random effects to account for different baseline rates of variation across different speakers and phrases. Mean $F_0$ was log-transformed. The results reveal a main effect of race. Compared to white women, black women employ a significantly lower mean $F_0$ ($\beta = -0.12$, $p < 0.01$). There is, as expected, a main effect of gender as white men have a significantly lower mean $F_0$ than white women ($\beta = -0.62$, $p < 0.001$). The interaction between race and gender is also significant, suggesting that for black speakers, the difference in mean $F_0$ between men and women is significantly smaller than it is for white speakers ($\beta = 0.11$, $p = 0.04$).

Notably, this model configuration does not fully resolve all the possible comparisons. Therefore, we performed further planned comparisons by first resetting the reference level to the "white men" and then to "black women." The allows us to make additional comparisons between white men and black men as well as between black women and black men. Results indicate that black men do not significantly differ from white men in terms of mean $F_0$ ($\beta = -0.00$, $p = 0.75$) and again, black women have significantly higher mean $F_0$ than their male counterparts ($\beta = 0.62$, $p < 0.001$). To further test whether there race-based differences exist in $F_0$, regardless of gender, we sum coded gender and treatment coded race. This allows us to control the gender effect while testing whether black speakers generally differ from white speakers. In other words, this coding scheme allows us to narrow in on the race effect, i.e., whether there is a difference between black and white speakers at all with gender being held at the average value. The results further suggest that black speakers have an overall lower $F_0$ than white speakers ($\beta = -0.06$, $p = 0.02$) and by sum-coding race, we find that men generally have significantly lower $F_0$ than women ($\beta = -0.56$, $p < 0.001$).

In short, at least for mean $F_0$, for both black and white speakers, men have an overall lower value than women, as predicted. Additionally, black speakers have an overall lower $F_0$ than white speakers. However, this difference is likely driven by female speakers, as black male speakers do not differ from white male speakers in this respect. In Sec. IV A 2, we test whether similar patterns persist for $F_0$ range measurements.

#### 2. Pitch range

Table II presents all the range measurements except for SD4 (4 standard deviations around the mean) in semitones (in absolute values). It seems that in general, female speakers have wider range than male speakers, both for black and white speakers.

https://doi.org/10.1121/10.0014906

TABLE II. Range measurements for black and white speakers.

| Speaker groups | 100% range (ABS[a] ST) | 90% range (ABS ST) | 80% range (ABS ST) | SD4 (Hz) |
|---|---|---|---|---|
| Black women | 6.94 | 6.25 | 5.55 | 68.58 |
| Black men | 5.61 | 5.05 | 4.49 | 35.29 |
| White women | 7.52 | 6.77 | 6.02 | 81.71 |
| White men | 5.37 | 4.83 | 4.30 | 35.05 |

[a]ABS, absolute values.

For range, we conducted another similar linear mixed-effects model using 90% range (ST) as the dependent variable and race and gender as fixed effects (treatment-coded in a two-way interaction). We focus here on the detailed analysis of 90% pitch range instead of the full 100% $F_0$ range. Expressing the full 100% $F_0$ range can be problematic as a single mistracked frame can throw the entire measurement off kilter. For example, any bit of creakiness will drop the floor of the minimum end of the range, giving a misleading summary statistic. This may have the potential to undermine the analysis of $F_0$ range. Therefore, to make sure our analysis is solid, details of the analysis of 90% pitch range are presented here. The analysis full 100% pitch range was also conducted, but only reported here briefly.

The model output is further summarized in Table III. For 90% range, there is main effect of gender. White men employ a significantly narrower $F_0$ range than white women ($\beta = -1.26$, $p < 0.001$). The effect of race is not significant, suggesting that black women do not differ from white women in $F_0$ range ($\beta = -0.25$, $p = 0.55$). The interaction between race and gender is not significant ($\beta = 0.65$, $p = 0.19$).

We then refit the same model with different reference levels multiple times and used different contrast coding schemes following our previous analysis of mean $F_0$ to generate all the tests that interested us. Further results suggest that, among black speakers, black men also use a significantly narrower 90% $F_0$ range, compared to black women ($\beta = -1.21$, $p < 0.01$). In addition, black men do not significantly differ from white men ($\beta = 0.41$, $p = 0.15$). In general, male speakers tend to use a narrower 90% pitch range than female speakers, regardless of race ($\beta = -0.93$, $p < 0.001$). Additionally, black and white speakers do not differ from each other significantly in 90% $F_0$ range ($\beta = 0.08$, $p = 0.75$).

Similar analyses were implemented to see whether 100% range, 80% range, as well as SD4 (i.e., 4 standard deviations around the mean) differ across different racial and gender groups. Figure 1 shows the SD4 pattern (in ST). Another three linear mixed-effects models with similar configurations were conducted and a set of modelling results similar to those for 90% range were found for all these measurements. For 100% range (ST), black men also use a significantly narrower $F_0$ range, compared to black women ($\beta = -1.34$, $p < 0.01$). In addition, black men do not significantly differ from white men ($\beta = 0.45$, $p = 0.15$). In general, male speakers tend to use a narrower pitch range than female speakers, regardless of race ($\beta = -1.04$, $p < 0.001$). Additionally, black and white speakers do not differ from each other significantly in 100% $F_0$ range ($\beta = 0.09$, $p = 0.75$).

The results for 80% range (ST) are similar. Black women do not behave differently compared to white women ($\beta = -0.22$, $p = 0.55$) and black men pattern with white men ($\beta = 0.36$, $p = 0.15$). White men have narrower 80% range than white women ($\beta = -1.12$, $p < 0.001$) and black men also have narrower 80% range than black women ($\beta = -1.07$, $p < 0.01$). Overall, men have narrower range than women ($\beta = -0.83$, $p < 0.001$), whereas black speakers do not differ from white speakers ($\beta = -0.07$, $p = 0.75$).

As for SD4 (Hertz), black women are similar to white women ($\beta = -3.35$, $p = 0.09$) and black men are similar to white men ($\beta = 0.32$, $p = 0.80$). White men use a narrower pitch range than white women ($\beta = -14.76$, $p < 0.001$) and black men also use a narrower pitch range than black women ($\beta = -11.09$, $p < 0.001$). In general, men exhibit smaller SD4 than women ($\beta = -12.92$, $p < 0.001$), and black speakers are similar to white speakers ($\beta = -1.51$, $p = 0.20$).

In sum, similar to 90% $F_0$ range, for 100%, 80% $F_0$ range, and SD4, we observe a robust gender difference. Women tend to employ a wider $F_0$ range, both for black and white speakers. On the other hand, there is no difference based on race for any of these measures. Black and white speakers have similar $F_0$ range distributions. This differs from the results for mean pitch level across phrases, in which we observe both a main effect of gender and an interaction effect between gender and race. That is, women tend to have a higher mean $F_0$ than men, and white women

TABLE III. LMER results for 90% range (ST) ~ race * gender + (1 | speaker) + (1 | phrase). Asterisks in the table indicates levels of significance given the *p*-values.

| Fixed effects | Estimate | Std. Error | df | t value | Pr (>\| t\|) |
|---|---|---|---|---|---|
| (Intercept) | 6.49 | 0.33 | 134.39 | 19.89 | <0.001*** |
| Race (vs white speakers) | −0.25 | 0.41 | 89.91 | −0.60 | 0.55 |
| Gender (vs female) | −1.26 | 0.35 | 89.93 | −3.64 | <0.001*** |
| Race: Gender | 0.65 | 0.50 | 90.05 | 1.32 | 0.19 |

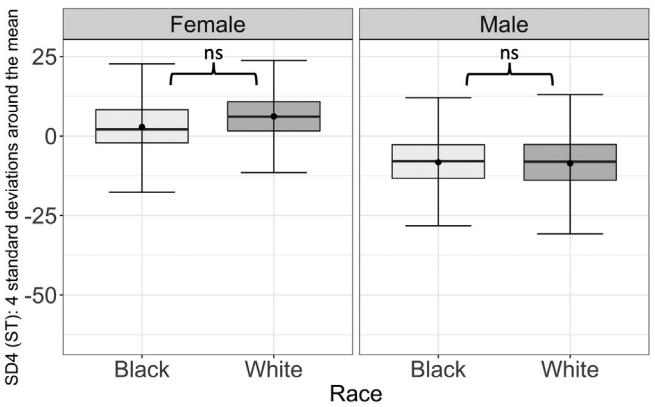


FIG. 1. Box plot comparing the SD4 (4 standard deviations around the mean) for black and white speakers.

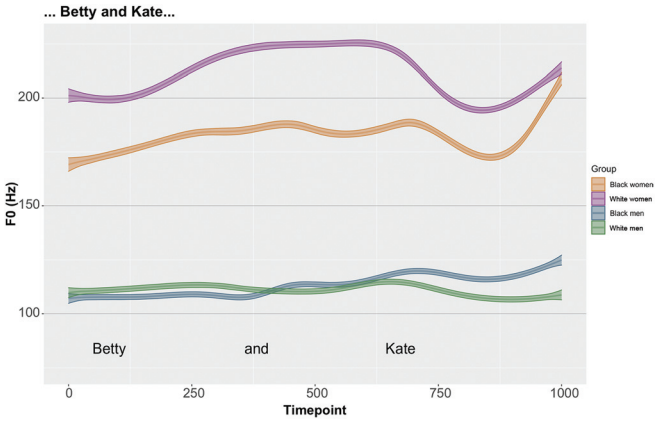


FIG. 2. Modelled pitch contour of the phrase *Betty and Kate*.

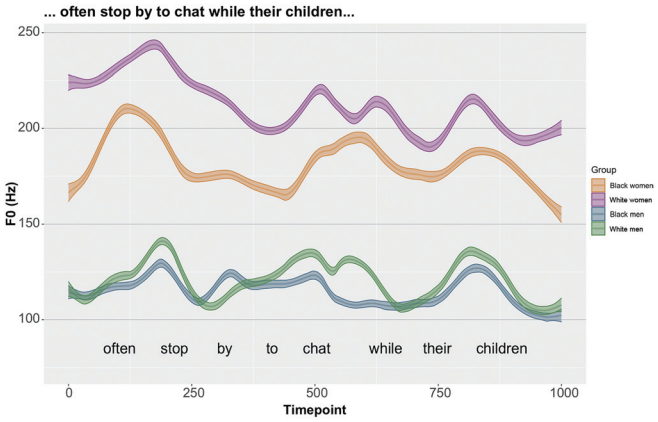


FIG. 4. Modelled pitch contour of the phrase *Often stop by to chat while their children*.

employ an even higher mean $F_0$ than black women. White men and black men tend to pattern with each other for both mean $F_0$ and $F_0$ range measures. Thus, our findings fail to support the basic stereotype of black men having a lower $F_0$ than white men, and underline the need to further explore these variables in women's speech. In Sec. IV B, we turn to a qualitative SSANOVA analysis of the $F_0$ contour shapes of individual phrases to investigate group differences in $F_0$ implementation across different phrases.

### B. SSANOVA results

We provide here a mixed methods analysis of the findings. The figures below show modeled pitch contours of the phrases of interest for each race/gender combination. The figures demonstrate notable differences in pitch contour shape between the groups. Whenever contours do not overlap, it suggests that the groups have produced reliably distinct pitches across the entire phrase. Our analysis includes an examination of 17 phrases; we excluded phrases in the original data set that were shorter than 3 or longer than 10 syllables, to control for differences caused by phrase length. Overall, results indicate some patterned differences between the four groups, as well as differences along lines of race and gender.

First, the most significant differences appear to be related to the location of peaks in the contour. Nine of the phrases have instances where black speakers realize a similar peak at a later time than the white speakers do, as suggested by Fig. 2. These findings support those of earlier studies that have found that black speakers may realize $F_0$ peaks at later locations than white speakers (Holliday, 2016).

Second of all, the height of the $F_0$ peaks also appears to differ by group across some phrases. Specifically, the white women typically employ higher peaks than the other groups, as displayed in Fig. 3. In addition, as suggested by Fig. 4, black women also display a greater tendency to use falling or less pronounced rising melodies at phrase boundaries than the other groups. Interestingly, we also observe instances where black speakers use double peaks in shorter phrases where white speakers use only one peak, as shown in Fig. 5.

There were few differences aside from peak timing and this double peak phenomena between the white and black men, indicating that perhaps differences in contour shape between groups are more evident in the speech of female speakers. This could be partially due to the fact that the

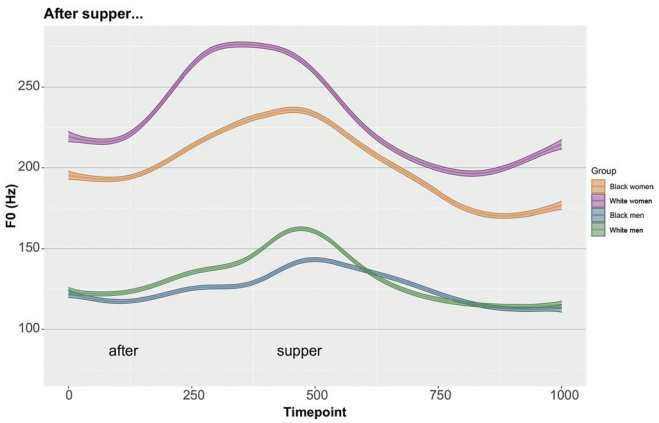


FIG. 3. Modelled pitch contour of the phrase *After supper*.

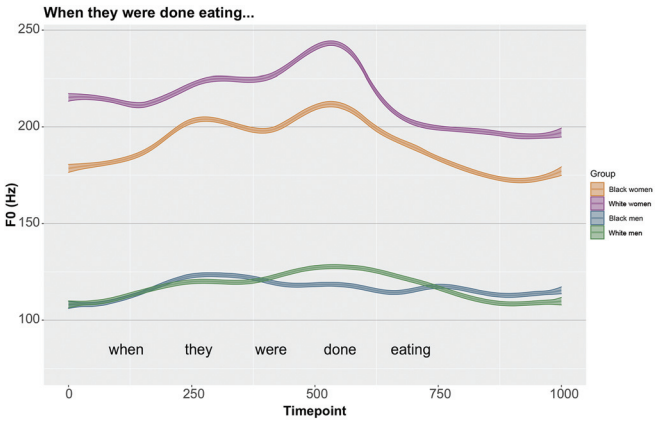


FIG. 5. Modelled pitch contour of the phrase *When they were done eating*.

https://doi.org/10.1121/10.0014906

women use a wider pitch range that extends into higher $F_0$s (Traunmüller and Eriksson, 1995), allowing them more space to move within a comfortable range. However, it is also notable that the semitone-transformed $F_0$ contours of the white and black women never cross or overlap for any of the phrases studied. For the white and black men, the contours overlap for almost every phrase. This result underlines the importance of including women in investigations of pitch differences between racial groups. Moreover, while more research is certainly needed to better model these differences, the current study provides at least some evidence for pitch contour differences between the groups of interest that cannot be fully captured through mean $F_0$ and $F_0$ range measurements. To date, there is no existing research that has posited such double peak phenomena as subject to ethnolinguistic variation, but future studies should examine whether such a pattern may reflect differences between ethnolects.

## V. DISCUSSION

The key results of our analysis demonstrate that the stereotype that black men employ a lower $F_0$ than white men is not necessarily supported. Rather, we discovered that there exists a robust difference in the pitch behaviour among women of different races. Mixed-effects regression model results show black and white men do not differ in either mean $F_0$ nor in pitch range measurements, and exhibit—at a group level—overlapping pitch contours in every phrase we investigated. On the other hand, we find that black women consistently employ a lower vocal pitch than white women, both in terms of mean $F_0$ and pitch contour, despite the fact that black and white women do not differ in measures of pitch range either. We, first of all, offer a line of interpretation of these results that they indicate a complex picture of the relationship between gender and race and the use of pitch as a stylistic resource in the traditional Labovian (Labov *et al.*, 2011) sense. We then relate our findings and interpretations back to the ethnic identification literature, demonstrating that better consideration of how multiple identity dimensions interact may help us better understand ethnolinguistic variation in pitch.

### A. Pitch as a stylistic resource to perform gender and race

The key result in this study, which differs from others of its type, is the discovery of a robust difference in the pitch behaviour of women of different races that is not found for men. While several (but by no means all) previous studies have found $F_0$ differences between black and white men (Hawkins, 1993; Hollien and Malcik, 1962; Hudson, 1977; Hudson and Holbrook, 1981; Wheat and Hudson, 1988), our study fails to replicate this effect. Instead, we find that black women exhibit consistently lower pitch than white women, even though the groups utilize similar pitch ranges. Our interpretation of the men's results center around an important methodological difference between our study and several others; the data for our study come from read speech while many previous studies look at spontaneous conversations. Although this decision has allowed us to extend our analysis to compare aggregated intonational contours for different social groups, it is likely to have induced more self-monitoring and meta-linguistic awareness on the part of speakers. Classically, we would expect this to induce linguistic convergence towards a standard and more formal style, associated with a higher socioeconomic class (Labov, 2006). Indeed, previous studies have claimed that the pitch behaviour of black men is sensitive to style (Hudson and Holbrook, 1981, 1982), such that this group in particular uses a much wider range of pitches in spontaneous speech compared to read speech. Tarone (1973) is more specific still, claiming that the black men in her study only utilize a wider pitch range than her white subjects during "competitive" speech acts. These observations provide a possible explanation for our result that black and white men do not differ in pitch level or range in a read speech task. The context surrounding the reading task may induce more "standard" settings for pitch level and range, where what is standard is established by a dominant white model. In other words, we can recast classic task-based effects of style and formality in terms of referee design (Bell *et al.*, 2002; Rickford *et al.*, 1994), with the black men employing a narrower pitch range and/or higher $F_0$ to avoid pitch behaviors, or stereotypes surrounding them, that could be considered "non-standard."

The question then remains of why the read speech style reveals no differences between the black and white men in the data, but highlights differences between the black and white women. One possible interpretation of this result is that, because the racist stereotype linking AAE to low pitch is rooted in ideas of both blackness and masculinity, black men may be more aware of pressure to converge along $F_0$ dimensions. Alternatively, we should consider that white womanhood—the prevailing model of hegemonic femininity in modern America—is not necessarily understood or performed the same as black womanhood (Cole and Zucker, 2007). Our results indicate that vocal pitch, which is frequently used as a stylistic resource for the performance of gender (e.g., Zimman, 2018), may not be used in the same way for women of different races. In short, stereotypical notions of feminine voices as higher in pitch, and their resulting utility for performance of femininity, may not apply to many black female voices. Thus, perhaps the robust pitch differences we observe between black and white women may be a result of black women's non-participation in a linguistic performance of hegemonic white femininity, at least in a read speech task that is likely to elicit more "formal" or "standard" styles. Framed in this way, it is unavoidable that dimensions of formality and "standardness," whose roots are fundamentally rooted in racist and classist ideologies, are still navigated with reference to an ethnolinguistic landscape rife with inequity. Moreover, aspects of a speaker's identity, for example their race and gender, and their combination, relevant for understanding their place in this landscape and the appropriate linguistic targets for speakers to perform "standardness."

### B. The role of pitch in ethnic identification

To return to the central question, couched in the context of ethnic identification literature, of how pitch differs between black and white speakers in American English, we have demonstrated that this question is significantly more complex than it seems. To a first approximation, since black men and white men do not differ in terms of pitch-related measures in our study, it suggests that $F_0$ might not be as robust a cue to discriminate the voices of black men from white men as previously thought (cf. Hawkins, 1993; Hollien and Malcik, 1962; Hudson, 1977; Hudson and Holbrook, 1981; Wheat and Hudson, 1988). However, this does not mean that $F_0$ has no role to play in ethnic identification. Of course, there are other salient acoustic cues that may be used in combination with $F_0$ to accomplish this task, including voice quality and vowel quality. But more fundamentally, our results suggest that the task of ethnic identification cannot be done in a vacuum, and depends on other sociostylistic dimensions of a speaker and their speech. In our (read speech) data, $F_0$ would serve as a very reliable cue for distinguishing the black women from white women, but not necessarily for discriminating between other black and white speakers. This suggests that, if listeners use $F_0$ for ethnic identification, they must simultaneously make informed judgments about many other sociolinguistic aspects of the signal.

On a larger, ideological note, our results and analysis may also relate to the folk intuition that black Americans have lower pitched voices than non-black Americans, i.e., that lower pitch is a cue to "sounding black." We propose that the differences between black and white women observed may, in fact, play a key role. That is, perhaps the expectation for black speakers to exhibit lower pitch can be attributed to associations, not exactly between AAE and certain masculine traits, but between AAE and *non-femininity* as it is ideologically modeled by white women. This interpretation is coherent with the use of AAE, along with low pitch, juxtaposed in explicit contrast with white feminine styles (Barrett, 1998). Since black men are frequently taken as the prototypical black speakers in research (Morgan, 2004), fewer studies exist that include black women (cf. Hudson and Holbrook, 1982). However, these will be key to the further exploration of this theory, and of a model of ethnic identification that better accounts for the multidimensionality of speaker identity.

## VI. CONCLUDING REMARKS

This study combines both quantitative and qualitative analysis to empirically test vocal pitch differences between black and white men and women from Memphis, Tennessee. Based on our quantitative results, the prediction that black speakers employ lower pitch does not obtain where it was most expected: between black and white men. However, we find that black women consistently exhibit a lower pitch level than white women. We argue that these results, and the fact that they do not match the widespread perception of black voices, is best understood through a lens that simultaneously considers speaker race, gender, and the speech style under examination. Specifically, while a reading task may trigger convergence towards a standard, moderate pitch level for black men, the data from black women suggest the active non-participation in a white feminine model of pitch. Based on these findings, we advocate for a sociolinguistically informed approach to ethnic identification—one that explicitly examines the role played by women's speech—that considers the complex interplay between race, gender, and style.

## ACKNOWLEDGMENTS

The authors wish to thank Valerie Fridland for granting us access to the data and Tyler Kendall for his assistance with SLAAP corpus. Special thanks also goes to Charlie Farrington, Kirby Conrod, Matthew Winn, three anonymous reviewers, and editors for their comments and suggests. Part of this work was presented at PLC 45; we thank the audience for their feedback.

## APPENDIX: THE FULL TEXT OF THE READING PASSAGE

Some mornings in the summertime, when the sky is fair and the lawn covered in dew, the good Duke Post and his wife Peg walk down to the brook by their house. There, beside the trees, is their favorite place to sit, talk, and sip coffee. Her father, Don, and his dog, Bookie, often stop by to chat while their children, Betty and Kate, toss off their shoes and leap headfirst into the deep brook. It makes Peg feel like a kid again to watch them dive, shout, and slosh around in the water and swing off the old black tire tied to the oak tree. One hot hazy, dull afternoon, she gave a call to their friends Pam and Ben Powder, inviting them over for supper. On the way, their truck got stuck in the mud, and they showed up an hour late, for which they caught a good deal of teasing. But soon the crowd was having fun and the good hosts put out tunafish sandwiches, hot dogs, a big pot of bean soup, and beer bread. When they were done eating, it was a sin that no one had saved room for Peg's tasty spice cake that was yet to come. After supper, Duke, Ben, and his pal Bill went out on Duke's inflatable boat. Unfortunately, the sky got gray and started to pour rain. Bill lost his footing on the slick bank and fell in the water. After ten minutes he finally got into the boat. Once back on shore, the sudden weather shift sent everyone home, and the party was over.

[1]We use "ethnic identification" (instead of "racial identification") to refer to the field of study, following the body of previous work. Otherwise, "race" is used to refer to the social category.

[2]Participants were recorded using standard cassette tapes by a portable high fidelity Marantz PMD201 series tape recorder with an external microphone.

Babbitt, L. G., Gaither, S. E., Toosi, N. R., and Sommers, S. R. (**2018**). "The role of gender in racial meta-stereotypes and stereotypes," Soc. Cogn. **36**(5), 589–601.

Bachorowski, J.-A., and Owren, M. J. (**1999**). "Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech," J. Acoust. Soc. Am. **106**(2), 1054–1063.
Barrett, R. (**1998**). "Markedness and styleswitching in performances by African American drag queens," in *Codes and Consequences: Choosing Linguistic Varieties*, edited by C. Myers Scotton (Oxford University Press, Oxford, UK), pp. 139–161.
Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2014**). "Fitting linear mixed-effects models using lme4," arXiv:1406.5823.
Baugh, J. (**1983**). *Black Street Speech: Its History, Structure and Survival* (University of Texas Press, Austin, TX).
Bell, A., Eckert, P., and Rickford, J. R. (**2002**). "Back in style: Reworking audience design," in *Style Sociolinguistic Variation*, edited by P. Eckert and J. R. Rickford (Cambridge University Press, New York), Vol. 1, pp. 139–169.
Benor, S. B. (**2011**). "Mensch, bentsh, and balagan: Variation in the American Jewish linguistic repertoire," Lang. Commun. **31**(2), 141–154.
Boersma, P., and Weenink, D. (**2022**). "Praat: Doing phonetics by computer (version 6.2.23) [computer program]," available at http://www.praat.org/.
Bucholtz, M. (**1999**). "You da man: Narrating the racial other in the production of white masculinity," J. Socioling. **3**(4), 443–460.
Bucholtz, M. (**2011**). "Race and the re-embodied voice in Hollywood film," Lang. Commun. **31**(3), 255–265.
Bucholtz, M., and Lopez, Q. (**2011**). "Performing blackness, forming whiteness: Linguistic minstrelsy in Hollywood film 1," J. Socioling. **15**(5), 680–706.
Busà, M. G., and Urbani, M. (**2011**). "A cross linguistic analysis of pitch range in english l1 and l2," in *ICPhS*, pp. 380–383.
Calder, J., and King, S. (**2020**). "Intersections between race, place, and gender in the production of /s/," Univ. Pennsylvania Work. Papers Linguistics **26**(2), 5, available at https://repository.upenn.edu/pwpl/vol26/iss2/5.
Cartei, V., Bond, R., and Reby, D. (**2014**). "What makes a voice masculine: Physiological and acoustical correlates of women's ratings of men's vocal masculinity," Horm. Behav. **66**(4), 569–576.
Chun, E. W. (**2001**). "The construction of White, Black, and Korean American identities through African American vernacular English," J. Linguist. Anthropol. **11**(1), 52–64.
Chun, E. W. (**2013**). "Ironic blackness as masculine cool: Asian American language and authenticity on YouTube," Appl. Linguist. **34**(5), 592–612.
Clopper, C. G., and Smiljanic, R. (**2011**). "Effects of gender and regional dialect on prosodic patterns in American English," J. Phon. **39**(2), 237–245.
Cole, E. R., and Zucker, A. N. (**2007**). "Black and white women's perspectives on femininity," Cultur. Divers. Ethnic Minor. Psychol. **13**(1), 1–9.
Coles, S. M., and Pasek, J. (**2020**). "Intersectional invisibility revisited: How group prototypes lead to the erasure and exclusion of black women," Transl. Issues Psychol. Sci. **6**(4), 314–324.
Collins, P. H. (**2004**). *Black Sexual Politics: African Americans, Gender, and the New Racism* (Routledge, New York).
Crenshaw, K. (**1989**). "Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics," U. Chi. Legal f. 139.
Cutler, C. A. (**1999**). "Yorkville crossing: White teens, hip hop and African American English," J. Socioling. **3**(4), 428–442.
Ducote, C. A. (**1983**). "A study of the reading and speaking fundamental vocal frequency of aging black adults," Ph.D. thesis, Louisiana State University, Louisiana.
Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (**2005**). "Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices," Anim, Behav. **69**(3), 561–568.
Ferrand, C. T., and Bloom, R. L. (**1996**). "Gender differences in children's intonational patterns," J. Voice **10**(3), 284–291.
Fridland, V. (**1999**). "The southern shift in Memphis, Tennessee," Lang. Var. Change **11**(3), 267–285.
Fridland, V. (**2001**). "The social dimension of the Southern Vowel Shift: Gender, age and class," J. Socioling. **5**(2), 233–253.
Fridland, V. (**2003**). "'Tie, tied and tight': The expansion of /ai/ monophthongization in African-American and European-American speech in Memphis, Tennessee," J. Socioling. **7**(3), 279–298.
Gasser, E., Ahn, B., Napoli, D. J., and Zhou, Z. (**2019**). "Production, perception, and communicative goals of American newscaster speech," Lang. Soc. **48**(2), 233–259.
Gu, C. (**2013**). *Smoothing Spline ANOVA Models* (Springer Science & Business Media, New York), Vol. 297.
Gu, C. (**2014**). "Smoothing Spline ANOVA models: R package gss," J. Stat. Softw. **58**(5), 1–25.
Hasek, C. S., Singh, S., and Murry, T. (**1980**). "Acoustic attributes of preadolescent voices," J. Acoust. Soc. Am. **68**(5), 1262–1265.
Hawkins, F. D. (**1993**). "Speaker ethnic identification: The roles of speech sample, fundamental frequency, speaker and listener variations," Ph.D. thesis, University of Maryland, College Park, MD.
Holliday, N. R. (**2016**). "Intonational variation, linguistic style and the black/biracial experience," Ph.D. thesis, New York University, New York.
Holliday, N. R. (**2019**). "Variation in question intonation in the corpus of regional African American language," Am. Speech **94**(1), 110–130.
Hollien, H., and Malcik, E. (**1962**). "Adolescent voice change in southern Negro males," Speech Monogr. **29**(1), 53–58.
Holt, Y. F., and Bent, T. (**2019**). "Racial categorization and word identification: The influence of sex, race and regional dialect," J. Acoust. Soc. Am. **145**(3), 1819.
Holt, Y. F., and Rangarathnam, B. (**2018**). "$F_0$ declination and reset in read speech of African American and White American women," Speech Commun. **97**, 43–50.
Hooks, B. (**2014**). *Ain't I a Woman: Black Women and Feminism* (Routledge, New York).
Hudson, A. I. (**1977**). "A study of the speaking and reading fundamental vocal frequency of young black adults," Ph.D. thesis, ProQuest Information & Learning, Florida.
Hudson, A. I., and Holbrook, A. (**1981**). "A study of the frequency reading fundamental vocal of young black adults," J. Speech. Lang. Hear. Res. **24**(2), 197–201.
Hudson, A. I., and Holbrook, A. (**1982**). "Fundamental frequency characteristics of young black adults: Spontaneous speaking and oral reading," J. Speech. Lang. Hear. Res. **25**(1), 25–28.
Ingrisano, D., Weismer, G., and Schuckers, G. H. (**1980**). "Sex identification of preschool children's voices," Folia Phoniatr. Logop. **32**(1), 61–69.
Johnson, K. L., Freeman, J. B., and Pauker, K. (**2012**). "Race is gendered: How covarying phenotypes and stereotypes bias sex categorization," J. Pers. Soc. Psychol. **102**(1), 116–131.
Kendall, T. (**2013**). "New tools and speech databases," in *Speech Rate, Pause and Sociolinguistic Variation* (Springer, New York), pp. 37–48.
Labov, W. (**2006**). *The Social Stratification of English in New York City* (Cambridge University Press, New York).
Labov, W., Ash, S., Ravindranath, M., Weldon, T., Baranowski, M., and Nagy, N. (**2011**). "Properties of the sociolinguistic monitor 1," J. Sociolinguistics **15**(4), 431–463.
Levon, E. (**2015**). "Integrating intersectionality in language, gender, and sexuality research," Lang. Linguist. Compass **9**(7), 295–308.
Loveday, L. (**1981**). "Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae," Lang. Speech **24**(1), 71–89.
McConnell-Ginet, S. (**1983**). "Intonation in a man's world," JSTOR **3**(3), 541–559.
McLarty, J. (**2018**). "African American language and European American English intonation variation over time in the American south," Am. Speech **93**(1), 32–78.
Mennen, I., Schaeffler, F., and Docherty, G. (**2007**). "Pitching it differently: A comparison of the pitch ranges of German and English speakers," in *16th International Congress of Phonetic Sciences*, pp. 1769–1772.
Mennen, I., Schaeffler, F., and Docherty, G. (**2012**). "Cross-language differences in fundamental frequency range: A comparison of English and German," J. Acoust. Soc. Am. **131**(3), 2249–2260.
Michalsky, J., and Schoormann, H. (**2017**). "Pitch convergence as an effect of perceived attractiveness and likability," in *INTERSPEECH*, pp. 2253–2256.
Morgan, M. (**2004**). "'I'm every woman': Black women's (dis)placement in women's language study," in *Language and Woman's Place: Text and Commentaries*, edited by M. Bucholtz and R. T. Lakoff (Oxford University Press, New York), pp. 252–259.
Morrill, T. H. (**2015**). "The implementation of phrasal prosody by native and non-native speakers of english: Ss anova for multi-syllabic intonation contours," in *ICPhS*.

Ohala, J. J. (**1984**). "An ethological perspective on common cross-language utilization of $F_0$ of voice," Phonetica **41**(1), 1–16.
Pisanski, K., and Feinberg, D. R. (**2013**). "Cross-cultural variation in mate preferences for averageness, symmetry, body size, and masculinity," Cross Cult. Res. **47**(2), 162–197.
Pisanski, K., and Rendall, D. (**2011**). "The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness," J. Acoust. Soc. Am. **129**(4), 2201–2212.
Purnell, T., Idsardi, W., and Baugh, J. (**1999**). "Perceptual and phonetic experiments on American English dialect identification," J. Lang. Soc. Psychol. **18**(1), 10–30.
R Core Team (**2022**). R: *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria), available at https://www.R-project.org/.
Richardson, J. A. C. (**1973**). "The identification by voice of speakers belonging to two ethnic groups," Ph.D. thesis, The Ohio State University, Ohio.
Rickford, J. R., and McNair-Knox, F. (**1994**). "Addressee- and topic-influenced style shift: A quantitative sociolinguistic study," in *Sociolinguistic Perspectives on Register*, edited by D. Biber and E. Finegan (Oxford University Press, New York), pp. 235–276.
Slobe, T. (**2018**). "Style, stance, and social meaning in mock white girl," Lang. Soc. **47**(4), 541–567.
Sneller, B. (**2020**). "Phonological rule spreading across hostile lines: (TH)-fronting in Philadelphia," Lang. Var. Change **32**(1), 25–47.
Snidecor, J. C. (**1943**). "A comparative study of the pitch and duration characteristics of impromptu speaking and oral reading," Speech Monogr. **10**(1), 50–56.
Sutcliffe, D. (**2001**). "The voice of the ancestors: New evidence on 19th-century precursors to 20th-century African American English," in *Sociocultural and Historical Contexts of African American English*, edited by S. L. Lanehart (John Benjamins, Amsterdam), pp. 129–168.
Tarone, E. E. (**1973**). "Aspects of intonation in Black English," Am. Speech **48**(1/2), 29–36.
Thomas, E. R. (**2015**). *Prosodic Features of African American English* (Oxford University Press, Oxford).
Thomas, E. R., and Carter, P. M. (**2006**). "Prosodic rhythm and African American English," Engl. World-Wide. **27**(3), 331–355.
Thomas, E. R., and Reaser, J. (**2004**). "Delimiting perceptual cues used for the ethnic labeling of African American and European American voices," J. Socioling. **8**(1), 54–87.
Traunmüller, H., and Eriksson, A. (**1995**). "The frequency range of the voice fundamental in the speech of male and female adults" (unpublished), available at http://www.ling.su.se/staff/hartmut/f0_&mf.pdf.
Walton, J. H., and Orlikoff, R. F. (**1994**). "Speaker race identification from acoustic cues in the vocal signal," J. Speech. Lang. Hear. Res. **37**(4), 738–745.
Wheat, M. C., and Hudson, A. I. (**1988**). "Spontaneous speaking fundamental frequency of 6-year-old black children," J. Speech. Lang. Hear. Res. **31**(4), 723–725.
Wickham, H. (**2016**). *ggplot2: elegant Graphics for Data Analysis* (Springer, New York).
Yuasa, I. P. (**2008**). *Culture and Gender of Voice Pitch: A Sociophonetic Comparison of the Japanese and Americans* (Equinox Publishing, Indonesia).
Zimman, L. (**2018**). "Transgender voices: Insights on identity, embodiment, and the gender of the voice," Lang. Linguist. Compass. **12**(8), e12284.