# How Black Does Obama Sound Now?: Testing Listener Judgments of Intonation in Incrementally Manipulated Speech

Nicole R. Holliday and Dan Villarreal*

## 1 Introduction

Recent research in perceptual sociolinguistics has investigated a host of variables—primarily segmental—to assess the extent to which social meanings are constructed in perception similarly to the way they are constructed in ongoing production. Despite production research in sociolinguistics that has demonstrated how speakers use intonational variation to index various ethnic identities and social stances (Burdin 2015, Holliday 2016, Reed 2016), there has been a general lack of perceptual research on the social meanings of intonational variables. In addition, while research has demonstrated US listeners' ability to distinguish African American and white voices (cf. Thomas and Reaser 2004), these studies have also revealed challenges inherent in isolating speaker-specific variables that drive ethnic identification (Holliday and Jaggers 2015); indeed, there has been little research in prosody more generally in non-standard varieties of English (Burdin, Holliday, and Reed 2018). In the present study, we address these gaps in research by investigating the extent to which listeners perceive specific aspects of intonational variation as indexing ethnic identity.

At the same time, research in perceptual sociolinguistics has not confronted the issue of whether social meanings are incremental—that is, does a more phonetically extreme token of a socially marked variable correspond to a stronger social meaning? This gap is due in part to the fact that some socially marked variables are often binned as categorical rather than continuous, such as r-lessness or (ing) alveolarization; however, even when investigating necessarily continuous variables such as vowel quality, research on perceived social meanings also tends to bin variables into discrete categories (Villarreal 2018). In the present study, we address these gaps in research by investigating whether listeners' judgments of aspects of intonation are sensitive to the strength of the variable of interest in the phonetic signal.

We pursued these questions of intonational variation and social meaning via an online matched-guise task in which listeners rated samples of Barack Obama's speech based on how much they perceived samples as "sounding black". Critical stimuli either contained at least one L+H* pitch accent, which has been shown in production studies to be a resource for performance of African American identity in production (McLarty 2011, Holliday 2016), or no L+H* pitch accents; in addition, pitch accents in all critical stimuli varied according to phonetic extremeness (i.e., the magnitude of $F_0$ excursions). Listeners perceived stimuli with at least one L+H* token as sounding more black than those without, but only for phrases with more phonetically extreme realizations of the L+H* contour (i.e., those with a larger difference between $F_0$ maximum and minimum). These findings contribute to our understanding of the construction of social meanings for intonational variation.

## 2 Background

### 2.1 Ethnic Identification in the U.S.

A body of linguistic research on ethnic identification dating back nearly 70 years has found that US listeners are generally rather accurate (70–100%) at distinguishing black speakers from white speakers (cf. Thomas and Reaser 2004). However, despite this fact, it remains unclear which cues listeners are attuned to in ethnic identification and to this end, some more recent studies have attempted to unpack the role of suprasegmentals in ethnic identification. Thomas and Reaser (2004) found that listeners were equally accurate at ethnic identification for monotonized and unmodified stimuli, suggesting that listeners do not rely solely on $F_0$ cues in ethnic identification. They find that some

* The authors would like to express their gratitude to our listeners, Pomona College Linguistics and Cognitive Science, and the New Zealand Institute of Language, Brain, and Behaviour.

cues relevant to pitch accents are recoverable even from monotonized stimuli (i.e., amplitude, duration, and segmental qualities), so it is conceivable that pitch accents aided identification even in monotonized stimuli. Holliday and Jaggers (2015) examined listeners' ability to identify the ethnicity of US politicians based on single-word stimuli in order to assess the effects of voice quality on listener judgments. They found that several suprasegmental variables, including jitter and harmonics-to-noise ratio, influenced ethnic identification though they note that a combination of multiple speakers and contexts may cause challenges in isolating speaker-specific variables that influenced ethnic identification. For this reason, in the present study, we sought to control for the effect of speaker-specific voice quality variation by using a single speaker to create matched-guise stimuli in order to more carefully isolate the prosodic variables that may affect ethnic identification.

### 2.2 Intonational Variation-Pitch Accents

This study focuses on one particular type of intonational variable as a starting point for under-standing how listeners may react to ethnically-linked suprasegmental features, and we do this using methods based in the auto-segmental/metrical (AM) intonational framework. Though there are a few different methods for describing intonation, much of the work conducted in the last 30 years has employed the AM theory, which itself is based in intonational phonology (cf. Ladd 2008). Essential to the AM theory is the idea that movements in fundamental frequency ($F_0$), the main correlate of what we perceive as pitch, result from an underlying sequence of tones that determine their structure. In the AM theory, these tones are either high (labelled as H) or low (labeled as L), and all movements of the pitch contour are composed of a series of low and high sequences. The AM model is a phonological theory that applies cross linguistically, though its specifications and the inventory of possible tones varies cross-linguistically. The labeling system for intonational phenomena that is based on the AM theory is called the Tones and Breaks Index system (ToBI). Each language, and indeed a number of dialects and varieties, have distinct ToBI systems that reflect the variety's intonational specifications. The ToBI system for Mainstream American English (MAE), originally developed by Beckman and Ayers-Elam (1997) and based on the findings of Pierrehumbert (1980), is the only ToBI system generally in use for examining variation within American English. MAE-ToBI has previously been used for descriptions of Jewish English (Burdin 2016), Appalachian English (Reed 2016), as well as AAL (Jun and Foreman 1996, McLarty 2011, Holliday 2016). Though some scholars have posited that the intonational phonological inventory of AAL may differ from that of MAE, it is still considered a reliable method for analyzing intonation in AAL, at least until researchers further investigate development of an AAL ToBI system (Thomas 2015, Holliday 2016) .

The ToBI system for Mainstream American English contains two types of pitch movements: pitch accents which occur on some stressed syllables, and edge tones that occur at phrase boundaries. The current study focuses only on the movement of pitch accents, though it is important to note that we also tested for the perceptual effects of edge tones. This study focuses on the difference between two types of pitch accents in MAE: a simple high tone, labelled as H*, and a fall-rise, labeled as L+H*. Though more types of pitch accents exist, H* and L+H* are among the most common in several varieties of US English, including AAL (Burdin, Holliday, and Reed 2018). Figures 1a and 1b show characteristic examples of H* and L+H* contours, as extracted from the current data set.

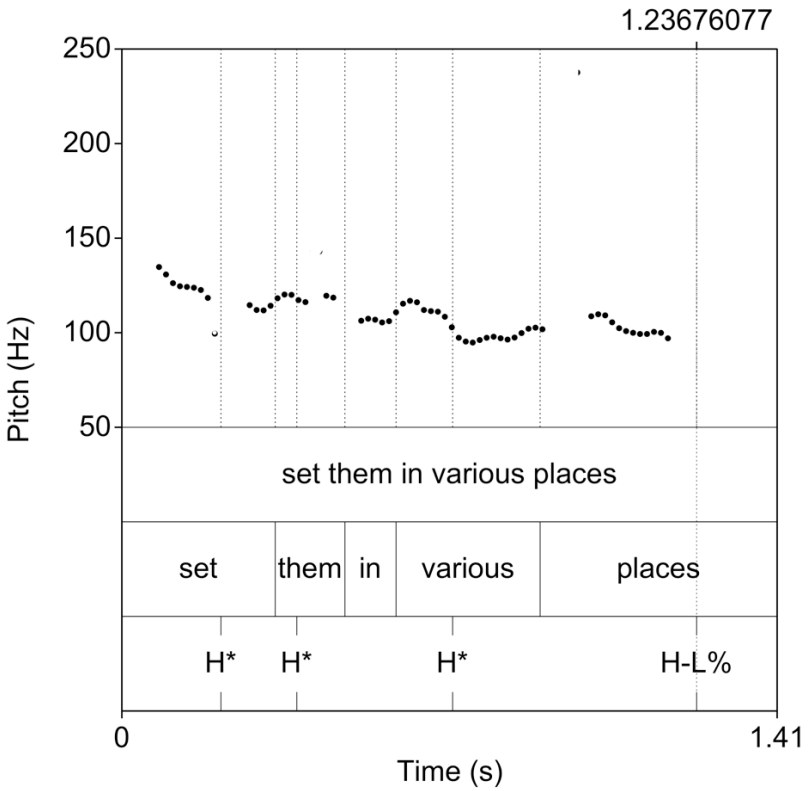


Figure 1a. Three H* pitch accents followed by a H-L% boundary tone.

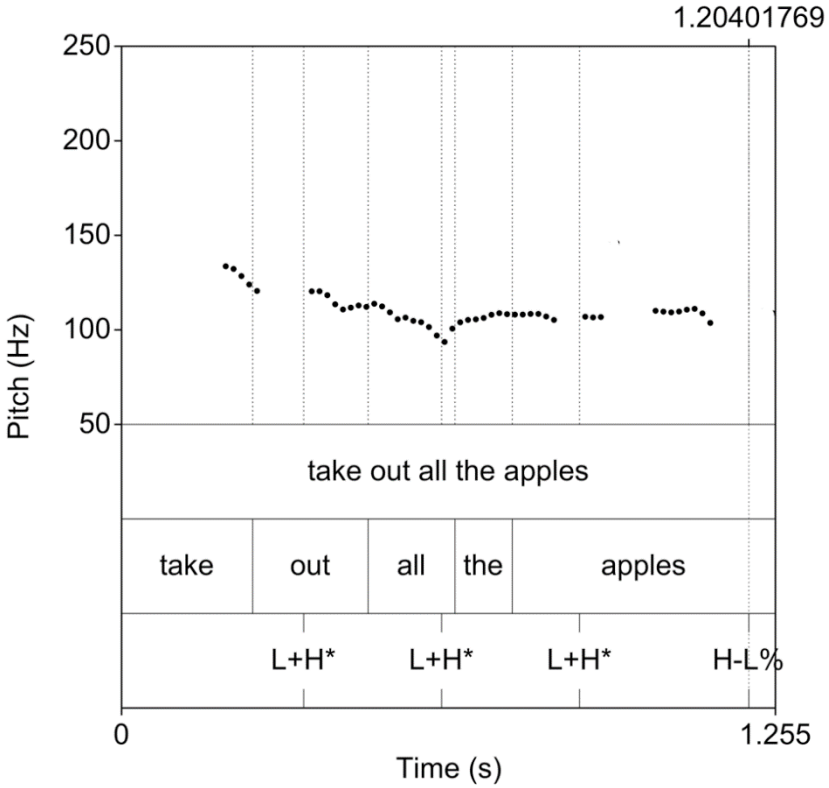


Figure 1b. Three L+H* pitch accents followed by a H-L% boundary tone.

Pitch accents in English can only occur on stressed syllables, though their appearance and use is related to the information structure of the phrase as well as a myriad of other factors (cf. Ladd 2008). Earlier studies have shown that pitch accents are perceptually salient for listeners and that naïve listeners can be trained to identify them quickly (Thomas 2011, McLarty et al 2017). Especially relevant to the current study, Loman (1975), McLarty (2011), and Holliday (2016) have found that differences in rate and context of use for H* vs. L+H* pitch accents in declarative phrases may constitute important differences between intonational patterns in MAE as compared to AAL. In particular, Loman (1975) and McLarty (2011) each found that L+H* pitch accents may be more common in some varieties of AAL.

Recent work by the first author (Holliday 2016) as well as a number of other recent studies (Burdin 2015, Reed 2016) have also found that a greater rate of use of the L+H* contour may also be a resource in production for performance of ethnic identity. In particular, Holliday (2016) recorded 25 men (age 18–32) with one black parent and one white parent in Washington, DC to examine their rates of use of different types of pitch accents. The participants were recorded in casual peer dyad conversations (20 min) with friends, and the analysis of their intonational patterns was taken from these recordings. A sociolinguistic inter-view was also designed and conducted to elicit ideologies about race and self-identifications. Holliday observed a pattern such that the participants who identified more as black, as opposed to multiracial or mixed, were more likely to use greater quantity of L+H* accents than H* accents. This finding supports Loman's (1975) and McLarty's (2011) earlier findings about a greater rate of use of the L+H* contour in AAL than in MAE, and it demonstrates that speakers' production of intonational variation is gradient in terms of frequency.

### 2.3 Incrementality in Intonation and Perception

While sociolinguistic variables tend to be analyzed as either categorical (e.g., negative concord, (ing) alveolarization) or continuous (e.g., /s/ center of gravity, vowel backing), pitch accent variation has both a categorical dimension (e.g., H* vs. L+H*) and a continuous dimension (e.g., peak $F_0$). A recent study comparing L+H* contours in Jewish English, African American English, and Appalachian English shows that both categorical and continuous properties of pitch accents are sites for sociolinguistic differentiation (Burdin, Holliday, and Reed 2018). In particular, the authors found differing rates of use of the L+H* contour among these speech communities, as well as significant cross-community differences in acoustic properties of these L+H* pitch accents: peak $F_0$, rise span, peak offset, and rise slope.

Continuous variables like peak $F_0$ present a different set of possibilities for social meaning than categorical variables like /l/-vocalization. Whereas for categorical variables each individual token presents a discrete set of options (e.g., velar (ing) vs. alveolar (in)), such that fine differentiation via frequency is possible only through the aggregation of several tokens (e.g., 50% (in) vs. 72% (in) vs. 89% (in)), for continuous variables fine differentiation is possible for each individual token (e.g., $F_2$ values of vowels) as well as in the aggregate. As a result, it is possible that incrementality in the phonetic properties of continuous variables is reflected in incremental social meanings of these variables. To this end, Podesva (2011) hypothesizes that salience takes one of two linguistic forms: categorial salience (cued by frequency) and phonetic salience (cued by phonetic extremeness). He argues that a more extreme production signals a stronger social meaning: "If an axis of phonetic variation indexes a particular social meaning, then outliers on that axis can be understood as the strongest indicators of meaning" (254). For example, when examining a gay speaker's intonational contours in declaratives, Podesva shows that two particularly extreme $F_0$ falls appear to be used to lend particular animation to the utterances during which they occur.

A few perceptual studies also lend further credence to Podesva's hypothesis about phonetic salience. Plichta and Preston (2005) presented listeners with a synthesized continuum from monophthongal to diphthongal /ay/ and asked listeners to identify the speaker's geographic origin along an axis running from the US north to the US south. Listeners not only associated monophthongal /ay/ with the south and diphthongal /ay/ with the north, they also placed successive continuum steps linearly along the north–south axis. In another study, narratives from 36 New Zealand English speakers—24 Māori and 12 Pākehā (non-Māori)—were measured for properties of voice quality; listeners then identified 15-second samples of these narratives as either Māori or Pākehā. The rate

at which speakers were identified as Māori was affected by several voice quality measures; for example, the speakers with the highest mean H1–H2 (a measure of creakiness) were also most likely to be identified as Māori (Szakay 2012). Beyond these two studies, however, little research in perceptual sociolinguistics—and virtually none on intonational variation—has addressed this phonetic salience hypothesis: in what ways can continuous variables carry continuous social meanings?

## 3 Methodology

This study was designed to address two central research questions:

1. How do pitch accents affect listener judgments of ethnic identity? In particular, does the L+H* pitch accent carry a social meaning of blackness in perception, as it does in production?
2. To what extent are the ethnicity-based social meanings of these pitch accents mediated by incremental phonetic differences?

These questions were investigated via a perceptual task in which listeners rated samples of Barack Obama's speech for perceived blackness. This task borrowed from matched-guise methods, in which listeners judge samples that differ by only a single feature, which allows insight into the social meanings of a feature (cf. Campbell-Kibler 2007). The task was administered via an online survey hosted by Qualtrics. In each of 120 randomly ordered trials, listeners heard a single stimulus auto-play twice and responded to the question "How black or white does Obama sound here?" on a unit-less slider bar with *very black* and *very white* on opposite poles. As the recognizability of Obama's voice would have likely rendered ineffective the type of instrumental task framing often used in matched-guise tasks (e.g., rating prospective radio newsreaders, as in Labov et al. 2011), we eschewed such framing, instead telling listeners, "This study is designed to test how people respond to different speech excerpts from the same speaker." Listeners then completed a demographic questionnaire and were invited to comment on the task.

The survey was distributed via social network sampling in May 2017, with a raffle incentive for one randomly selected listener to win an Amazon.com gift card for their participation. The listener sample for analysis contains 79 American English-speaking listeners who self-identified as black and/or white. Of these listeners, 24% self-identified as black and 77% as white (one listener identified as both); 65% identified as female and 35% as male. The majority of listeners also self-identified as liberal and indicated that they approved of Obama's presidency to a high degree.

### 3.1 Stimulus Creation

The 120 clips used as stimuli were based on excerpts of Obama's spontaneous speech from two different 2015 television interviews with Gayle King, a black broadcast journalist who co-anchors the *CBS This Morning* program (Kaplan 2016). The interview was conducted in two parts, both in a private room in the White House, and they aired on February 7, 2016, as part of CBS' promotion of the Super Bowl that was to air later that day. Each stimulus excerpt was based on a single Intonational Phrase unit, ranging from 0.4 to 2.3 seconds in duration (median: 0.9 seconds). We attempted to select short phrases that were fairly semantically bland to avoid overly tilting responses in one direction, though it is impossible to completely control for content in such an experiment.

First, sixty phrases from the interviews were selected as stimulus excerpts: 20 were critical excerpts (from which we created four stimuli apiece), and 40 were filler excerpts (one stimulus apiece). Ten critical excerpts were H* phrases, which contained between 1–3 H* pitch accents and 0 L+H* accents; and ten were L+H* phrases, which contained between 1–3 L+H* pitch accents and 0–2 H* accents. This design was necessary since L+H* pitch accents are relatively rarer, so it was not possible to find enough excerpts that only contained only L+H* accents and no H* accents. Filler excerpts contained 1–3 H* pitch accents and 0 L+H* accents. Excerpts contained either H-L% or L-L% boundary tones, and boundary tone was also tested as a possible predictor.

The critical stimuli were created by manipulating critical excerpts to four manipulation steps, with the original excerpt as Step 1. Steps 2, 3, and 4 were created by modifying each pitch accent in critical excerpts to make $F_0$ minima and maxima successively more extreme. Maxima (in both H* and L+H* accents) were increased by a semitone with each manipulation step; minima (in L+H* accents) were decreased by a half-semitone with each manipulation step. For example, the H* pitch accent in the top panel of Figure 2 has an $F_0$ maximum at 118.3 Hz in step 2 and 125.2

Hz in step 3, a one-semitone difference; the L+H* pitch accent in the bottom panel has an $F_0$ minimum at 101.6 Hz in step 1 and 99.1 Hz, a half-semitone difference. (In some cases, it was not possible to make the manipulations exactly one or one-half semitone.) The $F_0$ manipulations were based on semitones rather than flat magnitudes because semitones are psychoacoustically comparable regardless of initial $F_0$ of the pitch accent (e.g., the difference between 100 and 105 Hz sounds much larger than the difference between 200 and 205 Hz). The first author created stimuli by hand using the Manipulation utility in Praat (Boersma and Weenink 2015). Both authors listened to all manipulated critical stimuli and confirmed that they sounded natural.

Filler stimuli were created by modifying the final syllable of filler excerpts to include percepts of creaky voice: low $F_0$ and damped pulses (Keating et al. 2015). In particular, a Praat script modified alternating cycles of the final syllable by lengthening their duration and lowering their amplitude, the effect of which is evident in Figure 3. As with critical stimuli, both authors listened to all manipulated filler stimuli and confirmed that they sounded natural.

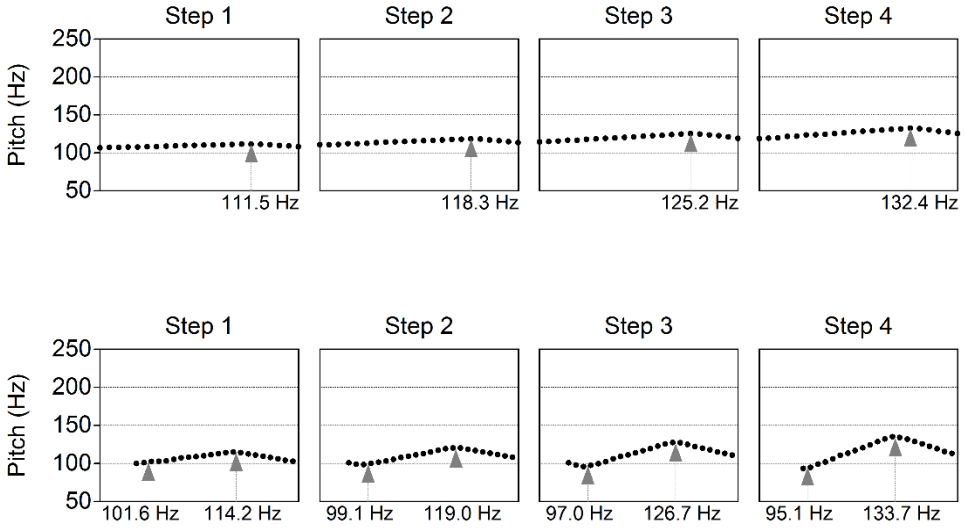


Figure 2: Manipulated versions of pitch accents in stimuli: H* pitch accent in would (top) and L+H* pitch accent in all (bottom)

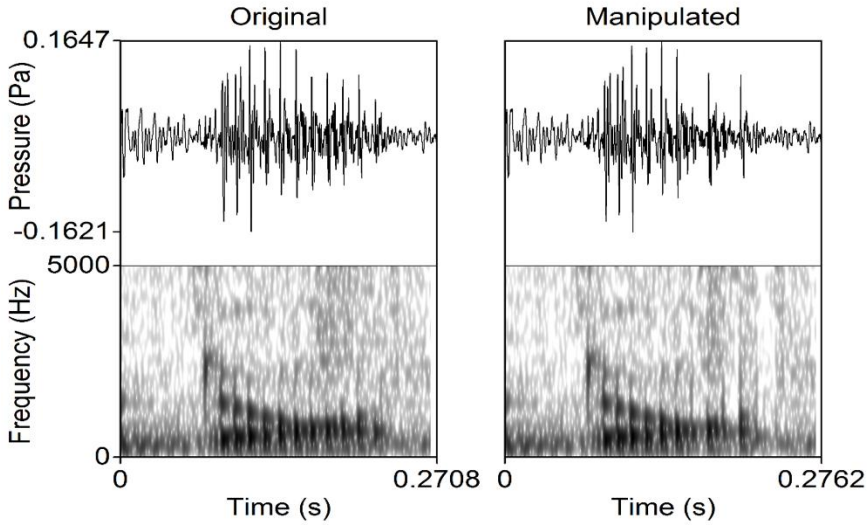


Figure 3. Original and manipulated versions of *doing* from a filler stimulus.

### 3.2 Analysis and Hypotheses

The only question posed to the listeners for each stimulus was "How black or white does Obama sound here?" and they responded using a continuous slider bar. After surveys were completed, the slider-bar positions were converted to real numbers between 0 (*very white*) and 100 (*very black*), then standardized by listener to control for the effect of variable usage of the continuous slider bar. All results are reported here in unit-less standard deviations (i.e., *z*-scores); a difference of 1 standard deviation can be interpreted as a difference of roughly one sixth of the length of the slider bar for the average listener in this survey.

Using the lmerTest package (Kuznetsova et al. 2016) for R (R Core Team 2015), we compared linear mixed-effects models of standardized ratings to find the predictor structure that best modeled the data in critical trials. The predictors that we tested were phrase type (H* vs. L+H*), manipulation step, boundary tone (L-L% vs. H-L%), nuclear pitch accent, stimulus duration, and numerous listener effects (race, gender, political ideology, approval of Obama's presidency, education, use of desktop vs. mobile to complete survey, hometown, geographic mobility, experience with linguistics, musical experience). Unfortunately, the distribution of H* and L+H* tokens in stimuli precluded predictors for the number of H* and number of L+H* pitch accents in critical trials.

We also included random intercepts for excerpts as nested within phrase type, as each excerpt exclusively belonged to one of the two phrase types. Since ratings were standardized by listener, by-listener random intercepts would be redundant; attempting to include random slopes for listeners led to convergence errors.

## 4 Results

The table below presents a summary of the best model for listener rates of blackness, which included predictors of phrase type, manipulation step, and their interactions.

| | Estimate | SE | df | t | Pr(>\|t\|) | |
|---|---|---|---|---|---|---|
| (Intercept) | -0.0409 | 0.1065 | 23.1 | -0.384 | 0.7045 | |
| PhrTypeL+H* | 0.0172 | 0.1507 | 23.1 | 0.114 | 0.9103 | |
| Step2 | 0.0041 | 0.0458 | 6056 | 0.089 | 0.929 | |
| Step3 | -0.0214 | 0.0459 | 6056 | -0.466 | 0.6415 | |
| Step4 | 0.0132 | 0.0459 | 6056 | 0.287 | 0.7737 | |
| PhrTypeL+H*:Step2 | 0.0297 | 0.065 | 6056 | 0.457 | 0.6475 | |
| PhrTypeL+H*:Step3 | 0.1314 | 0.065 | 6056 | 2.02 | 0.0434 | * |
| PhrTypeL+H*:Step4 | 0.1047 | 0.065 | 6056 | 1.609 | 0.1076 | |

Table 1: Summary of best model of listener ratings of blackness (* = $p < .05$). Degrees of freedom obtained via Satterthwaite approximations (Kuznetsova et al. 2016).

### 4.1 Results By Phrase Type

Model results show that there was no main effect of phrase type (H* vs. L+H* phrases) on listener ratings of blackness of the clips, indicating that pitch accent alone did not trigger different blackness ratings. Figure 4 shows this result, with the panel on the left indicating the results for the manipulation of the H* phrases, and the panel on the right indicating the results for the manipulation of the L+H* phrases. As is visible in these figures, listener ratings of blackness were remarkably similar for the H* phrases at each step, though the L+H* phrases showed greater differences between manipulation step, which will be addressed further in the next section. There was also no main effect of manipulation step on listener ratings of blackness.

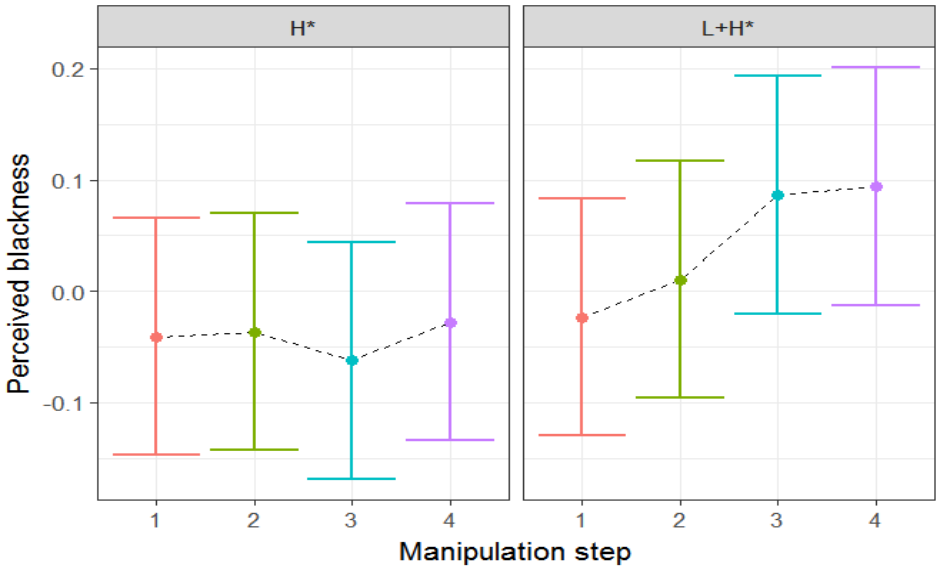


Figure 4. Results of regression model by phrase type and manipulation step, with H* phrases in the left panel, and L+H* phrases in the right panel.

### 4.2 Results By Manipulation Step

Though the main effect of phrase type failed to reach significance, the model results do indicate a significant interaction between phrase type and manipulation step, with more extreme L+H* phrases rated as sounding blacker than less extreme L+H* phrases, and no perceived blackness difference for H* phrases regardless of step. Figure 5 presents these results with each panel representing the difference between the ratings of the H* phrases and the L+H* phrases at each manipulation step. We can see from these results that listeners appear to interpret the more extreme L+H* realizations (greater difference between $F_0$ minimum and $F_0$ maximum within a L+H* contour) as blacker, but this is not the case for the more extreme H* realizations (which only had higher $F_0$ maxima).

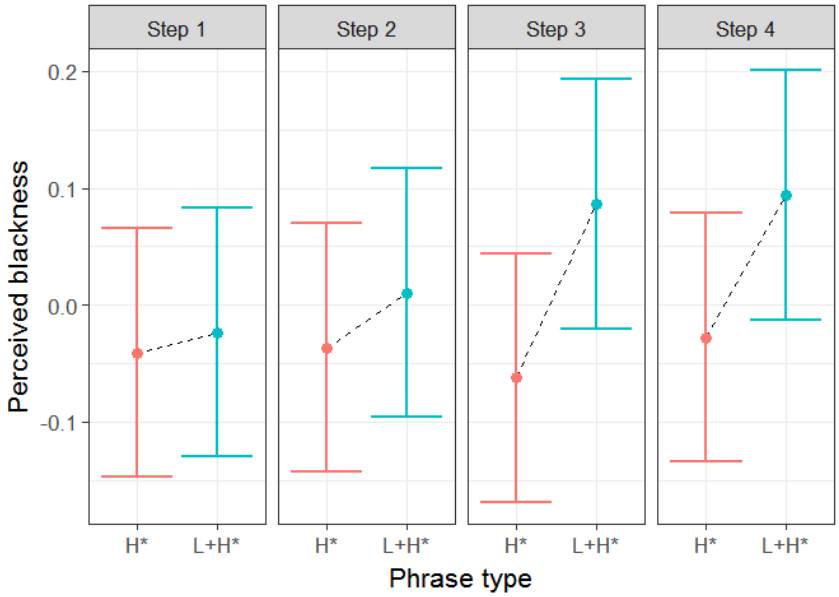


Figure 5. Results of regression model by manipulation step, with each panel representing a manipulation step, and both phrase types compared in each panel.

Also notable is that the model reveled no significant listener effects for gender, race, region, education, or political affiliation. This indicates that listeners were remarkably similar in their ratings regardless of a number of potentially influential demographic factors.

## 5 Discussion

This study has demonstrated that listeners are sensitive to the phonetic realizations of the H* and L+H* pitch contours in declaratives, and that a larger difference between the $F_0$ maximum and minimum within L+H* contours appears to cause listeners to rate a speaker, in this case, Obama, as sounding blacker. However, the difference between H* and L+H* pitch accent phrases is not sufficient to trigger this judgment; it is the actual realization of the contours themselves that listeners seem to attune to.

These results have several promising implications for our understanding of intonational variation and speaker judgments of race. First, the study has shown that a methodology of using a speaker (or speakers) known to listeners can yield results in a study that tests intonational differences. Second, we have seen that single-speaker stimuli may be useful for testing gradient manipulations, and this may especially be the case for speakers such as Barack Obama, who are already known to employ multiple styles. This latter implication indicates an exception to the conventional wisdom in matched-guise research that these tasks only work if listeners believe they are judging different speakers (Giles and Billings 2004).

We have also demonstrated that incrementality in intonation affects social meaning of variables, supporting earlier production claims by Podesva (2011) among others. These results have implications for studies of linguistic prejudice, since they indicate that listener impressions may also be incremental as opposed to categorical. Practically, this means that listeners may not simply negatively evaluate a speaker for "sounding black," but rather that the phonetic details of their intonational use may bear on the extent to which they are linguistically profiled. This result is exciting because it represents a first step at quantifying the phonetic details of intonational variation that may bear on how listeners may make judgments of ethnicity in everyday life as well as engage in linguistic profiling.

## References

Alim, H. Samy and Geneva Smitherman. 2012. Articulate While Black: Barack Obama, Language, and Race in the U.S. Oxford: Oxford University Press.

Beckman, M. E. and Ayers Elam, G. 1997. Guidelines for ToBI labelling, version 3.0.

Benor, Sarah Bunin. 2010. Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. Journal of Sociolinguistics 14(2). 159-183.

Boersma, Paul and David Weenink. 2015. Praat [phonetic analysis software], vers. 5.4.01. `http://www.fon.hum.uva.nl/praat/`

Burdin, R. S. 2015. Phonological and phonetic variation in list intonation in Jewish English. Presented at New Ways of Analyzing Variation (NWAV) 44. October 24th, Toronto, ON, Canada.

Burdin, R., Nicole Holliday and Paul Reed. Rising Above the Standard: Variation in L+H* contour use across 5 varieties of American English. Proceedings of Speech Prosody 2018,

Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American Speech* 82.32-64.

Giles, Howard, and Andrew C. Billings. 2004. Assessing language attitudes: Speaker evaluation studies. *The handbook of applied linguistics*, ed. by Alan Davies and Catherine Elder, 187-209. Malden, MA: Blackwell.

Holliday, Nicole. 2016. Intonational Variation, Linguistic Style, and the Black/Biracial Experience.New York University. PhD Dissertation

Holliday, Nicole R. and Zachary S. Jaggers. 2015. Influence of suprasegmental features on perceived ethnicity of American politicians. Paper presented to 18th International Congress of Phonetic Sciences, Glasgow, Scotland, 2015.

Jun, Sun-Ah, and Christina Foreman. 1996. "Boundary Tones and Focus Realization in African American English Intonations." *The Journal of the Acoustical Society of America* 100 (4): 2826–2826.

Kaplan, Rebecca, director. Obamas Share Super Bowl Traditions With Gayle King. CBS News. CBS Interactive, 7 Feb. 2016, www.cbsnews.com/news/obamas-share-super-bowl-traditions-with-cbs-gayle-king/.

Keating, Patricia, Marc Garellek and Jody Kreiman. 2015. Acoustic properties of different kinds of creaky voice. Paper presented to 18th International Congress of Phonetic Sciences, Glasgow, Scotland, 2015.

Kuznetsova, Alexandra, Bruun Brockhoff and Haubo Bojesen Christensen. 2016. lmerTest [R package], vers. 2.0-33. `https://CRAN.R-project.org/package=lmerTest`

Ladd, D. Robert. *Intonational phonology*. Cambridge University Press, 2008.

Loman, Bengt. 1975. Prosodic patterns in a Negro American dialect. Style and text: studies presented to Nils Erik Enkvist, ed. by Håkan Ringbom, Alfhild Ingberg, Ralf Norrman, Kurt Nyholm, Rolf Westman, and Kay Wikberg, 219–42. Stockholm, Sweden: Språkförlaget Skriptor AB.

McLarty, Jason A. 2011. An Apparent Time Analysis of Intonation Using Ex-Slaves: Comparing Raleigh, North Carolina AAE and EAE Pitch Accent Types and Frequencies, North Carolina State University. M.A. Capstone

McLarty, Jason, Charlotte Vaughn and Tyler Kendall. 2017. Acoustic correlates of perceived prosodic prominence in African American English and European American English. New Ways of Analyzing Variation 46. Madison, Wisconsin.

Pierrehumbert, Janet Breckenridge. 1980. The phonology and phonetics of English intonation. Massachusetts Institute of Technology. PhD Dissertation.

Plichta, Bartek and Dennis R. Preston. 2005. The /ay/s have it: The perception of /ay/ as a North-South stereotype in US English. *Acta Linguistica Hafniensia* 37. 243-285.

Podesva, Robert J. 2011. Salience and the social meaning of declarative contours. Journal of English Linguistics 39(3). 233-264.

Purnell, Thomas, William J. Idsardi and John Baugh. 1999. Perceptual and phonetic experiments on American English dialect identification. Journal of Language and Social Psychology 18(1). 10-30.

R Core Team. 2015. R: A language and environment for statistical computing, vers. 3.2.2. `https://www.R-project.org/`

Reed, Paul. 2016. Sounding Appalachian: /aI/ Monophthongization, Rising Pitch Accents, and Rootedness: University of South Carolina. PhD dissertation.

Szakay, Anita. 2012. Voice quality as a marker of ethnicity in New Zealand: From acoustics to perception. Journal of Sociolinguistics 16(3). 382-397.

Thomas, Erik R. 2011. Sociophonetics: An Introduction. Basingstoke, U.K./New York: Palgrave. 356 + xiv pp.

Thomas, Erik R. and Jeffrey Reaser. 2004. Delimiting perceptual cues used for the ethnic labeling of African American and European American voices. Journal of Sociolinguistics 8(1). 54-87.

Villarreal, Dan. 2018. The construction of social meaning: A matched-guise investigation of the California Vowel Shift. *Journal of English Linguistics*. 46.52-78.

Department of Linguistics and Cognitive Science
Pomona College
185 E. Sixth St, Room 211
Claremont, CA 91711
*nicole.holliday@pomona.edu*

New Zealand Institute of Language, Brain, and Behaviour
University of Canterbury
Private Bag 4800, Christchurch
*daniel.villarreal@canterbury.ac.nz*