Journal of Phonetics 94 (2022) 101163

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

American English pitch accents in variation: Pushing the boundaries of mainstream American English-ToBI conventions

Rachel Steindel Burdin [a,*], Nicole R. Holliday [b], Paul E. Reed [c]

[a] *University of New Hampshire, Durham 03824, NH, USA*
[b] *Pomona College, Claremont, 91711, CA, USA*
[c] *University of Alabama, Tuscaloosa 35478, AL, USA*

ARTICLE INFO

Article history:
Received 23 February 2021
Received in revised form 16 May 2022
Accepted 17 May 2022
Available online 18 June 2022

Keywords:
Intonational variation
ToBI
Variationist sociolinguistics

ABSTRACT

Linguists interested in intonation have long struggled to establish a maximally broad set of annotation conventions that function equally well across varieties of American English. The current study tests the advantages and limitations of the widely-used MAE-ToBI conventions, focusing on the H* and L+H* distinction, for three varieties of American English: African American English, Appalachian English, and Jewish English. Results of quantitative analysis of production data from 30 speakers of the three varieties finds major differences in rate of use of the H* and L+H* pitch accent as well as the phonetic realizations of these pitch accents, which may not be captured solely using the MAE-ToBI conventions. These differences appear not only between MAE-ToBI and the other three varieties, but also between the varieties themselves in unique ways that may shed light on the nature of sociolinguistic variation at the level of intonation, as well as the debated status of the distinction of H* vs. L+H* as a phonological or phonetic distinction. These findings provide further motivation for the development and use of annotation systems that explicitly consider sociolinguistic variation as well as phonetic parameters. Such systems will become even more essential as both sociolinguists and phoneticians expand intonational analysis beyond so-called “standard varieties” in order to arrive at a richer and more accurate picture of the intonational system of American English.

1. Introduction

Intonational differences between varieties are often salient to both laypeople and linguists. However, describing and formalizing these differences has posed many challenges. Part of this difficulty comes from the fact that while there is an agreed upon annotation standard for segmental features, the International Phonetic Alphabet, an equivalent set of conventions is lacking for intonation. While the IPA is not truly purely phonetic, as the set of symbols is driven by the goal of being able to notate all of the distinct phonemes in all the world’s spoken languages, it does also allow for researchers to do a range of transcriptions from more broadly phonemic to more closely phonetic. Different types of transcriptions are useful for different things. For example, a very broad phonological transcription allows for researchers to do acoustic analyses on segments which differ significantly in their phonetic realization while accounting for phonological structure. A more narrow transcription allows for initial observations of differences in the field, or for analyses of allophonic variation.

Starting in the 1980s and 1990s and continuing to this day, linguists developed a number of variety-specific Tone and Breaks Indices (ToBI) annotation systems for intonation (see Beckman et al., 2005; Ladd, 2008 on the history of the development of the systems, and Jun, 2005 and 2014 for an overview of different ToBI systems), based on the principles of autosegmental/metrical phonology. ToBI annotation systems are meant to be more phonological than phonetic, with the annotations indicating phonological tones.

The fact that ToBI systems are phonological, and are phonetically underspecified, has both benefits and drawbacks. One drawback is that ToBI systems have always been designed to be variety-specific, in contrast to the purported universality of the International Phonetic Alphabet. For example, English ToBI was intended to cover “general American,

* Corresponding author.
E-mail addresses: rachel.burdin@unh.edu (R.S. Burdin), Nicole.Holliday@pomona.edu (N.R. Holliday), pereed1@ua.edu (P.E. Reed).

https://doi.org/10.1016/j.wocn.2022.101163
0095-4470/

standard Australian, and southern British English" (Beckman & Elam, 1997, pg. 8); to further emphasize this, the variety for American English is specifically called Mainstream American English (MAE) ToBI. As such, researchers using this system to transcribe any variety of spoken American English which differs significantly from this imagined standard must make decisions about whether, how, and to what extent, MAE-ToBI can be used to transcribe the data without modification. Unfortunately, particularly in the United States, there remains a dearth of ToBI systems for non-Mainstream regional and ethnic varieties, so researchers frequently employ MAE-ToBI as the closest available annotation system (Thomas, 2015).

The explicitly phonological nature of ToBI has also led to extensive debates over the line between phonetics and phonology; specifically, whether some pairs of MAE-ToBI labels represent phonetic, rather than phonological, distinctions. In particular, there have been debates over whether the H* and L+H* labels, described in more detail below, represent two distinct phonological categories, or different phonetic realizations of the same category.

However, the phonological aspect of the system can also be a benefit, allowing us to do similar types of analyses to those that are done on segmental features: for example, in looking at phenomenon like /u/ fronting, researchers make an implicit claim that Californian English and Eastern New England both have a phonological category in the word *boot* distinct from *bead, bode,* and *book*; we can then look at variation in the phonetic realization of that category. This underspecification can also be useful in studying the existence of and/or location of category boundaries between phonemes in varieties with very different phonetic realizations of those categories (e.g., the distinction between /ɔ/ and /ɑ/ by a speaker with the Northern Cities Vowel Shift, where /ɔ/ is lowered to [ɑ], and, /ɑ/, fronted to [a], compared to a speaker from New York City, where /ɔ/ can be raised almost to [ʊ^{e}]).

In this paper, we explore to what extent MAE-ToBI is appropriate for use on three varieties of American English. Although some researchers have explored this problem for other varieties of English, including varieties spoken in the United Kingdom, including Glaswegian English (Mayo, Aylett, & Ladd, 1997), New Zealand English (Warren, 2005) and Australian English (Fletcher & Stirling, 2014), this remains an understudied area for varieties of English spoken in the United States. The varieties of interest —African American English (AAE), (American) Jewish English[1] (JE), and Appalachian English (ApE)— have been found to show variation in their intonation systems from more standardized varieties, particularly in their use and phonetic realization of H* and L+H* pitch accents.

In this paper, we focus in on (1) whether these varieties show evidence for phonetically distinct H* and L+H*s, and (2) whether these varieties use these pitch accents in different ways. This study has immediate practical implications, including in how MAE-ToBI training materials, such as the MIT Open Courseware course on transcribing prosodic structures using ToBI (Veilleux, Shattuck-Hufnagel, & Brugos, 2006) should describe the H* / L+H* contrast. It also has more theoretical ones. Previous work has suggested that the presence or lack of a distinction between H* and L+H* may be dialect dependent in the United States (Arvaniti & Garding, 2007). This work adds to that debate by showing that how the distinction is made between these pitch accents appears to be variety dependent.

[1] Jewish English has mainly been studied in North America (specifically, the United States and Canada), but distinctive features of Jewish speech have been noted in other English-speaking countries (Gold, 1985); this paper focuses on American Jewish English.

2. Variation in pitch accents

The current study works within the autosegmental/metrical approach to intonational phonology (see Ladd, 2008 for an overview). In this framework, intonational melodies (tunes) are decomposable into a series of underlying tones. This is in contrast to other theories (e.g., the "British school", see e.g., Crystal, 1969) which take tunes (e.g., a low rise, a rise-fall) as primitives. Building off of autosegmental theories of lexical tone (Goldsmith, 1976), tones are either low (L) or high (H); these tones then align with the segmental string in various ways depending on the language and/or variety in question. As described below, these tones are used for a variety of functions, including marking information structure and status, and the relationships between phrases.

Ladd (1996) outlines four ways in which intonational systems can vary from each other under this framework: semantic, systemic, realizational, and phonotactic variation. Semantic variation is variation in the meaning or function of particular tones and tunes. Systemic variation is variation in the inventory of tones. Realizational variation is variation in the phonetic implementation of tones. Finally, phonotactic variation covers differences in tone/text alignment, and permissible structures within the variety. Fletcher, Grabe, and Warren (2006) add a fifth parameter, based on work on variation in segmental systems, the neutralization of tones. We focus here on variation in pitch accents, which are tones aligned with stressed syllables, touching on variation in edge tones, which are tones aligned with phrasal boundaries, where necessary.

2.1. Systematic variation

While there is debate about the number and types of pitch accents in MAE, most researchers have relied on a general consensus based on the work of Pierrehumbert and Hirschberg (1990). They suggest an inventory of two monotonal pitch accents, H* and L*, and four bitonal pitch accents L+H*, L*+H, H+!H*, and H*+L, with the * indicating the tone associated with the stressed syllable. Later work has modified this inventory somewhat, specifically dealing with how to indicate downstepping: a sequence that Pierrehumbert and Hirschberg (1990), for example, annotate as H*+L H* is now generally annotated as H* !H*, as per the MAE-ToBI guidelines, version 3.0 (Beckman & Elam, 1997). Jun (2014) and the MIT ToBI training course (Veilleux et al., 2006) give an inventory of H*, !H*, L*, H+!H*, L+H* and L*+H. Some have questioned whether H* and L+H*, which can both be realized with a rise in a stressed syllable, represent two distinct phonological categories or are different phonetic realizations of the same underlying category, with, e.g., Ladd and Schepman (2003) arguing that these are in fact one category. Arvaniti and Garding (2007) suggest that part of this debate may be due to regional variation, with some varieties of English potentially lacking the H*/L+H* contrast and others maintaining it.

However, despite this observation about potential variation, research on variation in the number and types of pitch accents, as well as other differences in the intonational phonological systems on the level of phonotactic variation (which would include things like the number and types of phrase levels, and how tones are aligned with a segmental string) in American English is limited. Gooden (2009) suggests an upstepping process for H* and L+H* pitch accents in African American English in Pittsburgh. Other researchers have suggested that African American English may employ different tones or differences in phrasing (Cole, Britt, Thomas, & Coggshall, 2005; Farinella, Yu, Brugos, & Green, 2021). However, neither Reed (2016), Burdin (2016) nor Holliday (2016) suggest a difference in either the total number of tones or differences in prosodic phrasing in the three varieties studied here (ApE, JE, and AAE respectively). But, given the debate over the H* and L+H* contrast, it is still an open question as to whether the tones truly do contrast in these varieties.

2.2. Semantic variation

In varieties of American English, pitch accents are frequently assumed to mark information status. Pierrehumbert and Hirschberg (1990) describe monotonal pitch accents as marking new (H*) compared to given (L*) items. Rising pitch accents (L+H) mark the salience of a scale, with L+H* used for a correction or contrast (i.e., the salience of one item instead of an alternative), and L*+H, for uncertainty about the appropriateness of the scale itself. Some work has called into question a one-to-one mapping between information status and pitch accent choice even for standardized varieties of American English, instead calling for a more probabilistic approach (see, e.g., Chodroff & Cole, 2018; Turnbull et al., 2015). However, the pitch accenting of words and/or the manipulation of prosodic characteristics of stressed syllables still appears to be affected by information status in systemic ways, such that focused items are more likely to be produced with a L+H* pitch accents, and/or with higher peaks, larger rises, and/or steeper rises (ibid).

Previous work has found variation within American English in the distribution of pitch accents. Clopper and Smiljanic (2011), for example, found variation in the number and relative proportion of pitch accents used in reading passages both by gender (male vs. female talkers) and by location (Midwestern vs. Southern United States), and Schmid and Bradley (2019) reported variation between men, women, and non-binary people in the number and type of pitch accents used. This variation suggests possible semantic variation as well: if there is a relationship between focus and using L+H*, especially if that relationship is probabilistic, we may expect variation in how strong the link between focus and L+H* is in different varieties of English; or, to look at it from another angle, there may be differences in to what extent H* exists as a default, neutral pitch accent for new items. As such, a difference in the frequency of use of L+H* and H* may suggest differences in the meanings of these pitch accents.

For the three varieties in question here, previous work has found differences between these varieties and either co-territorial or neighboring varieties of American English in the relative proportion of pitch accent types, along with variation within and across speakers suggesting social meaning attached to that variation. Burdin (2016) found greater use of rising pitch accents (both L*+H and L+H*) in Jewish participants compared to non-Jewish participants in Dayton, Ohio. Greene (2006) found more rising pitch accents from participants from Appalachia as compared to other Southern speakers; Reed (2016) found that within an Appalachian community in East Tennessee, participants who were more “rooted” (a measure of an individual's local place-based attachment; see Reed, 2016, 2020 for more detail) used more rising pitch accents than those participants who were less rooted; there were also realizational differences based on rootedness. Finally, a number of studies have found increased use of rising pitch accents (specifically L+H*) by speakers of African American English compared to non-African American English speakers (McLarty, 2018; Thomas, 2015), with Holliday (2016) finding style shifting in pitch accent use and realization depending on context and topic.

In addition to variation in the relative proportion and frequency of different types of pitch accents, we also may see differences in, for example, which pitch accents are used in which positions (here, intonational phrase final, or nuclear vs. nonnuclear position), as well as the number of pitch accents per intonational phrase. Jun (2014) introduces the concept of macro-rhythm as a parameter of variation in prosodic typology, with more macro-rhythmic varieties having more regular alternations between low and high f0. These alternations can be caused by a variety of factors, including a greater frequency of pitch accenting, and a greater preference for rising or falling pitch accents compared to level pitch accents.

The frequency of use of pitch accents overall may be suggestive of different types of variation. For example, Burdin et al. (2015) found that languages with greater degree of macro-rhythm (more regular alternations in f0) appear to mark focus prosodically in different ways compared to languages with a lesser degree of macro-rhythm: highly macro-rhythmic K’iche’ showed no evidence of using any of the prosodic factors studied (pitch accent type, deaccenting, duration, etc.) to do so, unlike the less macro-rhythmic American English. From slightly different angle, we can also consider Calhoun’s (2010) work on metrical structure, and the role that deviances from more probable patterns play in signaling information structure. Under this model, how frequent pitch accents are, as well as, e.g., what types of pitch accents tend to appear in nuclear vs. non-nuclear positions, may have significant implications for how information structure is signaled. Finally, we can turn to Dainora’s (2002, 2006) work on American English, examining the relative frequency of combinations of pitch accents both with each other and with particular boundary tones. The work generally presents the tendency of particular pitch accents to appear in specific sequences as evidence against a tone-based approach to intonational meaning; however, even if one disagrees with that particular conclusion, it suggests that differences in where in the phrase particular pitch accents tend to occur again, may be indicative of differences in what those pitch accents mean. Thus, differences in the number of pitch accents per phrase, the relative distribution of pitch accent types, and their nuclear status may be reflective of both semantic and phonotactic differences under Ladd’s framework.

2.3. Realizational differences

A large body of research has studied differences in more global measurements of fundamental frequency (f0) and pitch in the United States correlated with a variety of social factors including gender (Schmid & Bradley, 2019; Zimman, 2013), race/ethnicity (Holt & Rangarathnam, 2018; Hudson & Holbrook, 1981; Loman, 1975; Tarone, 1973) and sexual orientation (Zimman, 2013).There has also been work on variation in the production of rising tunes in declarative sentences ("uptalk"; see Warren, 2016 for both a definition of the term and a thorough overview) (Podesva, 2011; Ritchart & Arvaniti, 2014). However, work specifically looking at the phonetic implementation of pitch accents in American English is much more limited. While work has been done on other languages and varieties (see Ladd et al. 2009 for British English and Scottish English; Fletcher et al., 2006; Grabe et al., 2000 for British English; and Atterer & Ladd, 2004 for German), variation in American English remains understudied, with some exceptions (see, e.g., Newmark et al., 2016 on pitch accents in Native American English).

For the varieties studied here, most of the work has focused on intra-group variation or style shifting. Burdin (2017) found that, in Jewish English, L+H* pitch accents were produced with higher peaks, and with larger rise spans (the difference between the peak of the rise and the preceding valley in Hertz; also called excursion) in more socially intimate settings by Yiddish/English bilinguals. Reed (2016, 2020), for Appalachian English, found that more rooted speakers aligned the pitch peak of the rising pitch accent earlier in the syllable than less rooted speakers. Finally, Holliday (2016) found differences by speaker racial alignment with respect to pitch accents and peak alignment relative to the onset of the vowel (hence forth, peak alignment) among Black/biracial men. Much of this work has been focused on the peak, and measurements related to that; however, others have proposed looking at Tonal Center of Gravity, a more global measure of f0 (Barnes, Veilleux, Brugos, & Shattuck-Hufnagel, 2010). Work on Jewish English has incorporated this measure as an area of variation (Burdin, 2017), but to better parallel other work and to limit the scope of the paper, we will focus on more static measurements here.

To summarize this previous work, we find similarities across the three varieties, in that all showed an increased of use of rising pitch accents, specifically L+H*, when compared to more standardized varieties. Secondly, these L+H* are all reported to be phonetically distinct, with differences in the phonetic realization of the pitch accents either compared to more standardized varieties and/or with differences in phonetic realization within the groups studied here.

Our first question for this paper is to what extent these varieties demonstrate this type of variation in ways that are similar or distinct from one another. An analogy can be made here to variation in rhoticity in varieties of American English. A large number of varieties are said to be non-rhotic, including coastal Southern English, Eastern New England English (ENE), New York City (NYC) English (Kurath & McDavid, 1961), and some varieties of African American English (Thomas, 2007). However, there has also been considerable work showing variation both within these regions, either by location or by racial/ethnic group in the phonological rules and constraints governing rhoticity, as well as differences across these regions (see Nagy & Irwin, 2010 for ENE English; Becker, 2014 for NYC English, among others). There is also considerable variation in the phonetic realization of these variables both within and across dialects: compare, for example, New York City's [ɜᶦ] for NURSE and ENE English's fronted [a] in START (Kurath & McDavid, 1961). There are, of course, key differences between segmental and suprasegmental features. In particular, there are more ways for segments to vary—compare the size of English's segmental inventory to its much smaller inventory of pitch accents and edge tones. However, the general principle still holds, in that there may be variation that is broadly similar in some ways (e.g., greater use of and/or more phonetically extreme L+H*s), but different in others (e.g., the exact phonetic implementation of L+H*).

This variation has implications for the appropriateness of the MAE-ToBI system for use on these varieties. We are particularly interested in how H* and L+H* are described in and the official guidelines as outlined in Beckman and Elam (1997), as well as materials like the MIT ToBI training course (Veilleux et al., 2006). Based on our own experiences in both learning ToBI annotation and training others, the MIT training course somewhat of a standard in the field. As such, how H* and L+H* are described in these materials thus will have implications for how annotators approach ToBI labeling. As Gooden et al. (2009) notes, the theoretical biases that researchers have about prosodic structure can impact how structures are perceived and analyzed.

First, with regard to the pitch accents, an informational box in the MIT ToBI training materials broadly describes the differences in meaning between L+H* and H*, with L+H* being associated at least in some cases with indicating contrast, but cautioning that "both contours can be used in a variety of contexts, and a specific context will not necessarily lead all speakers to select the same intonational contour" (Veilleux et al., 2006, Chapter 2.5, page 7). The ToBI labelling guidelines reference the contrast proposed in Pierrehumbert and Hirschberg (1990), with L+H* being "more likely to occur in a contrastive context", and mentions relying on "(theory-dependent) intuitions about meaning" to distinguish between H* and L+H* where a L may not be visible in the f0 contour (Beckman & Elam, 1997, n.p.). If an annotator strictly trained on these guidelines uses them to annotate a variety in which H* and L+H* are used in different ways, there is the possibility of under- or overlabeling particular pitch accents.

As for the phonetics of H* and L+H*, in the MIT ToBI training materials, L+H* is described as having a "sharp rise in pitch", compared to the "gradual" rise for H* (Veilleux et al., 2006, Chapter 2.5, page 5). The official ToBI labeling guidelines describe the difference between the two first in terms of whether or not there is a low preceding the peak: whereas L+H* has "a rise from a fundamental frequency low in the pitch range" that cannot be ascribed to another L tone (a L-, L* or L%), H* as "at most, a small rise from the middle of a speaker's pitch range" (Beckman & Elam, 1997, n.p.). Again, an annotator strictly trained on these guidelines may mislabel H* and L+H*s in varieties where the phonetic realization of these pitch accents differs from standardized varieties of English.

Building on previous work about rates of use and realization of pitch accents in non-MAE varieties, this study compares data from African American English, Jewish English, and Appalachian English. The goal of this comparative analysis is to better describe possible boundaries and parameters of pitch accent variation across varieties of American English. To this end, we present a comparative analysis of production data from speakers of the three varieties of interest, finding that though they each differ from MAE in unique ways, they also differ from each other in nuanced ways in both usage rate and realization. We end with a discussion of implications for the use of the MAE-ToBI system, focusing on H* and L+H*, for these varieties.

3. Methods

Each author interviewed 10 female speakers of each variety. The authors were all speakers of the variety in question for which they did the interviews. For the purposes of this paper, we take a more traditional sociolinguistic approach to linguistic variation (see Eckert, 2012 for a more detailed history of the field) with a speaker's variety (and indeed, the existence of a particular variety) being defined based on set social characteristics: for Appalachian English, residing in Appalachia (specifically, East Tennessee), for African American English, identifying as Black, and for Jewish English, identifying as Jewish[2]. However, as outlined above, all three authors have, at other points, studied these communities following a more "third-wave" approach. That is, rather than looking for differences in speech based on *a priori* broad social categories like race or religion, the authors have looked at variation within these varieties in order to explore how speakers use language to construct their identity(ies), as outlined in Eckert (ibid). None of these communities or varieties is homogenous, and as such, this work represents "strategic essentialism" (see Bucholtz, 2003), in abstracting away from some variation with the goal of studying variation at another level.

The Jewish English speakers were interviewed by Burdin in 2011. All were currently living in the metropolitan New York City area, and were previously known to Burdin to be generally active in Jewish life in the area prior to the interviews. All were English-dominant (used primarily English in their day-to-day lives), although three had grown up speaking Yiddish at home; another three had parents who were Yiddish-speakers, who spoke some Yiddish in the home. All were between the ages of 55 and 75 at the time of the recording. The participants were not explicitly asked about racial identity, but all were of Ashkenazi (Eastern European) heritage, and likely would identify themselves as White and/or Jewish, but not Black. The participants each wore a head-mounted Shure microphone, and were recorded in rooms that were relatively quiet, using a Roland R-05 recorder at a 44.1 kHz sampling rate.

The Appalachian English speakers were interviewed by Reed in 2014. All were living in East Tennessee and had lived exclusively in East Tennessee, and were all previously known to Reed as personal acquaintances and friends. Participants ranged in age from 27 to 77 at the time of recording and all identified as White. The participants each wore either a Shure or Audio-Technica lapel microphone, were recorded in rooms that were relatively quiet, using a Tascam DR-40 at a 44.1 kHz sampling rate.

The African American English speakers were interviewed by Holliday in 2016. All were living in New York City and had previously participated in a perception experiment about their evaluations of African American English. All participants were currently enrolled in at a university, and all self-identified as native speakers of both AAE and MAE. They were all were between the ages of 18–22 at the time of the interview. The participants each wore an Audio-Technica lapel microphone and were recorded in rooms that were relatively quiet, using a Zoom H5N recorder at a 44.1 kHz sampling rate.

These were all initially independent projects and thus had some minor differences in methodology. However, all of the interviews were in the style of a traditional sociolinguistic interview, consisting both of more free ranging ethnographic-style questions, as well as more structured elicitations. All interviews included reading passage data, which is the data analyzed here. The JE speakers were recorded reading "Comma Gets a Cure;" the AAE speakers, "The Rainbow Passage;" the ApE speakers, "Arthur the Rat" (see Appendix A for the texts).

The differences in passage likely caused some differences in the number and types of pitch accents used (see, e.g., Clopper & Smiljanic, 2011), as well as, potentially, the overall pitch range observed. "Arthur the Rat", for example, includes direct quoted speech, which the other two passages do not; "Arthur the Rat" and "Comma Gets a Cure" are both written in a somewhat fable-like style, compared to "The Rainbow Passage", which is a short non-fiction passage. The passages also differed in length, as well as the number of content words (coded following the criteria in Corver and Van Riemsdijk (2001), and bolded in Appendix A). "Arthur the Rat" is 584 words, with 297 content words; "Comma Gets a Cure" is 375 words, with 199 content words, and "The Rainbow Passage", 329 words, with 164 content words. On average, "Arthur the Rat" took 204 s to read; "Comma Gets a Cure", 138 s, and "The Rainbow Passage", 108 s. There are likely also speaking rate differences between the groups; however, exploring this is beyond the scope of this paper.

Despite these differences in the passages, as will be discussed in more detail, we believe that the differences between the three groups is still at least partially attributable to the speakers' varieties, as our findings generally replicate other studies of these varieties looking at other types of speech data, and we will discuss them as such. However, even if the differences we find below are solely due to differences in the data collection method or other factors (e.g., the relative age of the speakers, or geographic location), we believe our general conclusions and implications for ToBI annotation still stand.

Physiological differences in the larynx due to aging may have also affected the f0 results, with the Jewish English speakers being primarily older, the AAE speakers, younger, and the ApE speakers, a range of ages. Studies of the effects of aging on the larynx have generally found a lowering of pitch range for women (Reubold, Harrington, & Kleber, 2010). However, again, we believe that the differences described below are still representative of more general differences between the varieties, as much of the findings parallel previous findings based on non-read speech. And, as will be seen, the results for

[2] As far as we are aware, there are no overlaps between the groups: i.e., none of the Black speakers identified as Jewish, or vice versa.

f0 in particular generally go against what would be expected purely based on physiological factors.

The reading passages were orthographically transcribed and forced aligned, using FAVE for the JE and ApE data (Rosenfelder, Fruehwald, Evanini, & Yuan, 2011); and the Montreal Forced Aligner (McAuliffe, Socolof, Mihuc, Wagner, & Sonderegger, 2017) via the DARLA web interface (Reddy & Stanford, 2015) for the AAE data. The passages were then annotated using MAE-ToBI.

To ensure reliability across the annotations, five files from each variety were double coded by at least two different authors. A comparison of the annotations was done to find any areas of disagreement; after discussion, each author resolved conflicts for their own files. In general, we found that the files were being annotated according to the same general principles. Where there were areas of disagreement, we agreed on a standard choice. For example, final prominences which were ambiguously the result of a combination of a final pitch accent and a boundary tone or a boundary tone alone (e.g., a low rise which might either be a L* L-H% or just L-H %) were resolved in favor of just boundary tones (e.g., L-H %). The remaining files were single coded by the original interviewer.

The full ToBI annotations were extracted by means of a script. Below, we analyze the relative proportion of H* and L +H* pitch accents, as well as their nuclear vs. non-nuclear status, using logistic mixed-effects models. We did not include downstepped versions of these accents in this analysis. As noted above, each of these varieties has been described as having more rising pitch accents than other varieties of English; however, we had no *a priori* theories about how they would compare to the other beyond expecting to see some variation between the varieties based on our own intuitions.

Several phonetic landmarks were then marked for each H* and L+H* pitch accent: the starting point and ending point of the stressed vowel of the syllable, the location and f0 of the peak (H) and the location and f0 of the preceding valley (L). To get the starting and ending points of the vowels, the boundaries as generated by the forced aligners were hand corrected. The location of the peak was semi-automatically located using the “Get highest pitch” function in Praat (Boersma & Weenink, 2016) over the stressed syllable; as such, we include any peak within the stressed syllable, including those appearing after the end of the vowel. Rises with peaks occurring *after* the end of the stressed syllable were taken to be instances of L*+H, and are not included in this analysis. The valley was taken to either be a L within a directly preceding voiced section of speech, or, in the case of H* or L+H* preceded by non-sonorant sounds, occasionally, a L preceding the period without visible voicing. These landmarks were chosen as they have been previously been found to vary in other varieties of English, and have also been found to be distinctive of the varieties in question. Similarly to Ladd and Schepman (2003), we noted both H*s with more obvious rises, and H* that were either quite flat or fell slightly from the start of the stressed syllable; for those flatter H*s with no obvious preceding L, no L was marked. Some examples of the phonetic landmarks are given below, with a flat H* in Fig. 2, a H* with a rise in Fig. 3, and a L+H* in Fig. 4.

These landmarks were then extracted by means of a Praat script.[3] Rises under 2 Hz, or those for which Praat returned an “undefined” measurement for either the L or the H, were hand checked by a trained undergraduate research assistant, with a final pass by the authors, and adjustments made to Praat’s pitch tracking algorithm to get an accurate f0 measurement where possible.

Linear mixed-effects models were then built on looking at four measurements: the peak height, rise span, rise slope, and peak alignment from the start of the stressed vowel. Peak height was the f0 at H; rise span was calculated as the f0 at H minus the f0 at L; rise slope, the change in f0 from H to L divided by the time of H minus the time at L in milliseconds (thus expressed in Hz/milliseconds), and peak alignment, the time at H minus the onset of the stressed vowel (expressed in milliseconds). The results and graphs below show the results for the models built on f0. As noted above, all of the speakers were female, and, as such, there is less of a need to use a measurement such as ERB to control for physiological differences between the speakers. However, statistical models were also built on these measurements in ERB; the pattern of significance was identical for the results reported below, so we present the data in raw Hertz for ease of interpretation.

We chose these measurements as all of them have been found to be a locus of variation for the varieties studied here, as described in section 2.3 above. Likewise, some of these factors have been said to be diagnostic of the L+H* vs. H* distinction. L+H* is described as having a steeper rise than H*: a “sharp rise in pitch” compared to H*, where the rise is “more gradual” (Beckman & Elam, 1997, p. 16). L+H* also has a visible L, so we expect a larger rise span for L+H* compared to H*. Peak location is a potential diagnostic between L+H* and H*: with L+H* having an underlying L, we expect a later peak alignment due to needing to realize both tones. We can see illustrative examples of the difference between these pitch accents, in Fig. 1, adopted from Arvaniti and Garding (2007), with the L+H* showing a later peak alignment, a larger rise span, and a steeper slope compared to the H*. To these three measurements we add peak height, as this has previously been found to be a point of variation within these varieties. Figs. 2, 3, and 4 show examples of the different pitch accents in our data.

These measurements are all correlated with each other in the data set: e.g., pitch accents with higher peaks have larger rise spans, steeper slopes, and later peak alignments, though as the results below show, we see differences in the extent to which we see variation on these measurements both between varieties and between the pitch accent types.

4. Results

4.1. Relative use of H and L+H**

As noted above, the data were coded using MAE-ToBI, and thus the researchers approached the data with an *a priori* assumption that there was a distinction between H* and L +H*; acoustic evidence for the existence of this distinction in

[3] All Praat scripts used can be found at https://mypages.unh.edu/rsburdin/praat-scripts.

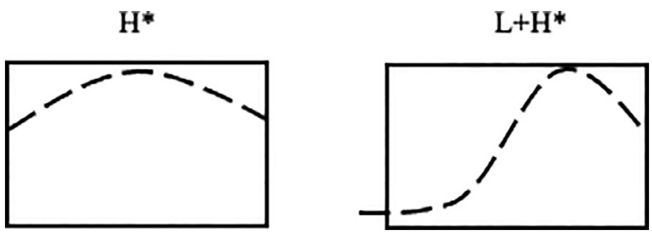

Fig. 1. Schematic of H* and L+H*, showing a steeper slope, later peak, and large rise span on L+H* compared to H*. Adapted from Arvaniti and Garding (2007).

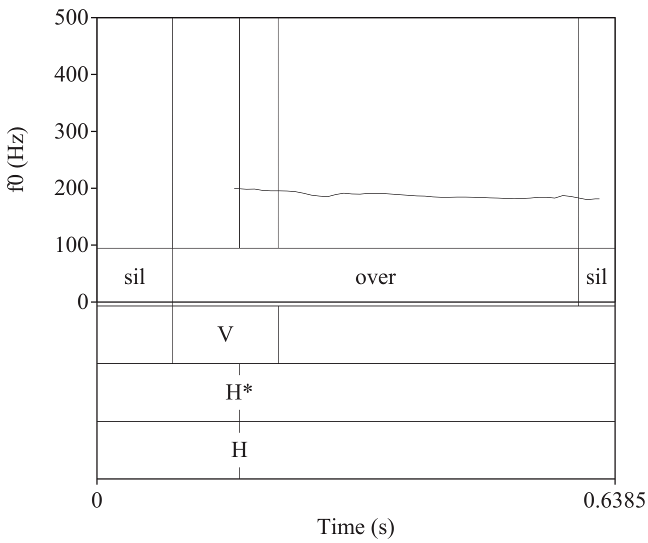

Fig. 2. Flat/slightly falling H* with phonetic landmarks (AAE).

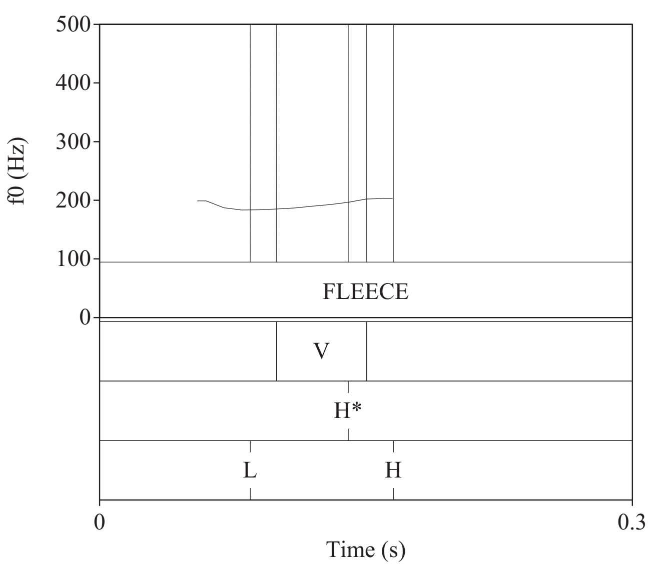

Fig. 3. Slightly rising H* with phonetic landmarks (JE).

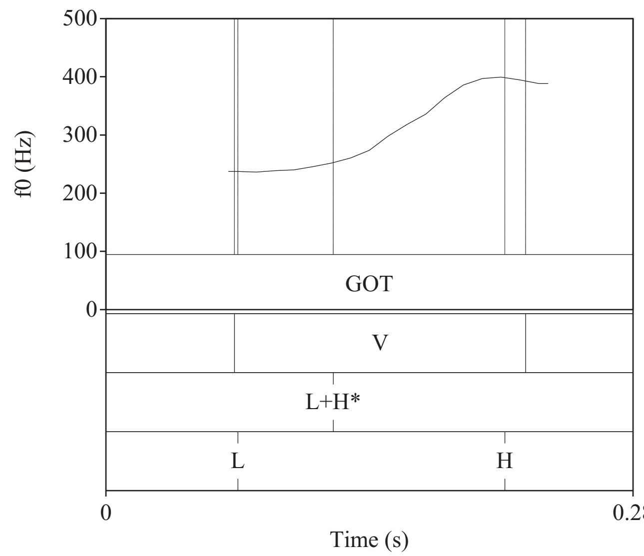

Fig. 4. L+H* with phonetic landmarks (JE).

all three varieties will be laid out in more detail below in Section 4.2.

Table 1 summarizes information about the passages presented above in Section 3, as well as the average length of the passage, average number of IPs, and average total pitch accents (all pitch accents, not just H* and L+H*). The longer passages led to the production of more IPs: ApEng speakers had the most, followed by JE, and then AAE speakers. The longer passages also had more pitch accents.

The Appalachian English and African American English speakers produced slightly more pitch accents per intonational phrase (IP) compared to the Jewish English speakers, with an average of 2.7 PAs per IP for ApE speakers and 2.6 for AAE to 2.0 for the JE speakers. A linear mixed-effects model was built in R (R Core Team, 2013) using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) to see if these differences were significant, with variety as a fixed effect, and speaker as a random effect. Variety was coded as a treatment contrast. Effects with $t > |2|$ were taken to be significant. With Jewish English as the reference level, both African American English ($\beta = 0.5503$, se = 0.1848, $t = 2.977$) and Appalachian English ($\beta = 0.6524$, se = 0.1801, $t = 3.623$) had significantly more pitch accents per phrase than Jewish English. The difference between African American English and Appalachian English was not significant.

These differences may be due to the difference in the number of content words in the reading passages (297 for Appalachian English; 199 for Jewish English; 164 for African American English). However, dividing the number of pitch accents produced by each participant by the total number of content words, the Appalachian English speakers produced the greatest proportion of pitch accents per content word, followed by African American English speakers, and then Jewish English. As such, this may reflect a general difference in the rate of pitch accenting across the varieties, particularly as this aligns with other work finding more pitch accenting in African American English (McLarty, 2018; Thomas, 2015); however, more data with more carefully controlled materials will be needed to study this.

Table 2 gives the total number of pitch accents labeled as H* and L+H* in each variety, along with the relative proportion of H* and L+H*; these data are also shown in Fig. 5.

We compared the overall proportions of H* to L+H* (taking nuclear status into account), as well as what types of pitch accents appeared in nuclear vs. non-nuclear positions. These two analyses get at two different, but related questions. The

Table 1
Summary of length, phrasing, and total pitch accents in each passage.

	Total words	Total content words	Average length	Average IPs	Average pitch accents	Average pitch accents per IP	Average pitch accent/content word
Comma Gets as Cure (JE)	375	199	138 s	67 (range: 59–77)	140.6 (range: 132–159)	2.0 (range 1.7–2.1)	0.70 (range: 0.67–0.95)
Arthur the Rat (ApE)	584	297	204 s	93 (range: 72–104)	245.5 (range: 200–285)	2.7 (range: 1.9–3.4)	0.83 (range: 0.67–0.95
The Rainbow Passage (AAE)	329	164	108 s	50 (range: 59–77)	126.3 (range: 109–147)	2.6 (range: 2.3–3.5)	0.70 (range: 0.64–0.80

Table 2
Number of H* and L+H* in each variety.

	H*	L+H*	Total H* and L+H*
AAE	810 (88%)	108 (12%)	918
ApE	1178 (68%)	547 (32%)	1725
JE	777 (81%)	185 (19%)	962

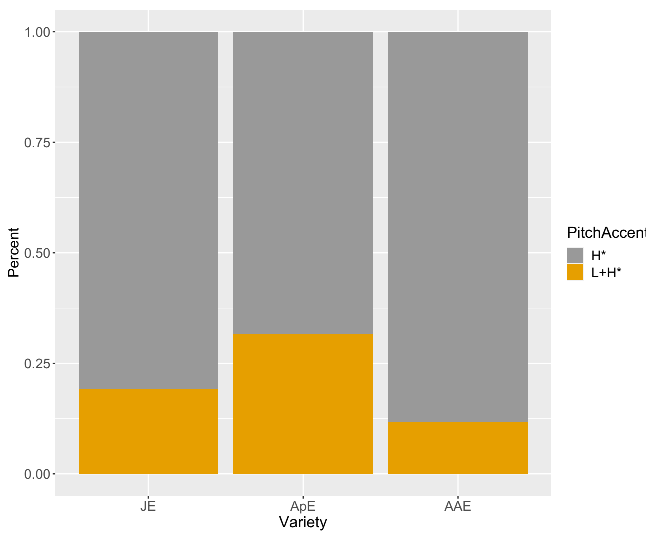

Fig. 5. Overall percentage of H* and L+H* by variety.

Table 3
Nuclear status by pitch accent.

Variety	Nuclear status	H*	L+H*
JE	Non-Nuclear	425 (55%)	67 (36%)
	Nuclear	352 (45%)	118 (63%)
ApE	Non-nuclear	714 (61%)	375 (69%)
	Nuclear	464 (39%)	172 (31%)
AAE	Non-nuclear	462 (57%)	75 (69%)
	Nuclear	348 (43%)	33 (31%)
TOTAL	Non-nuclear	1601 (59%)	517 (62%)
	Nuclear	1164 (41%)	323 (38%)

Table 4
Pitch accent by nuclear status.

Variety	Pitch accent	Non-Nuclear	Nuclear
JE	H*	425 (86%)	352 (75%)
	L+H*	67 (14%)	118 (25%)
ApE	H*	714 (66%)	464 (73%)
	L+H*	375 (34%)	172 (27%)
AAE	H*	462 (86%)	348 (91%)
	L+H*	75 (14%)	33 (9%)
TOTAL	H*	1601 (76%)	1164 (78%)
	L+H*	517 (24%)	323 (22%)

first allows us to explore the overall proportion of H* and L+H* pitch accents in the varieties; the second allows us to see differences in where these pitch accents tend to be in the intonational phrase.

As seen in Table 2, overall, there were more H*s than L+H*s. Tables 3 and 4 break down the number of H* and L+H* by variety and nuclear status. Table 3 shows the relative proportion of each pitch accent type by nuclear status (e.g., out of the H*s, how many are nuclear?); these data are also presented in Fig. 6. Table 4 shows the relative proportion of nuclear and non-nuclear pitch accents by each pitch accent type (e.g., what proportion of the nuclear pitch accents were H* or L+H*?); these data are also presented in Fig. 7.

Table 2 shows that, overall, H*s were more common than L+H* in each variety. However, Appalachian English has proportionately more L+H*s than African American English and Jewish English.

Table 3 shows that, overall, H*s were primarily non-nuclear. However, this differed slightly by variety, with ApE having more non-nuclear H*s than the other two varieties. Overall, L+H*s were also primarily non-nuclear; this also holds for both ApE and AAE, but not JE, where L+H*s were primarily nuclear.

As can be seen in Table 4, both non-nuclear and nuclear pitch accents tended to be H* rather than L+H*. However, again, there were differences by variety: nuclear pitch accents in AAE were much more likely to be H* compared to the other two varieties. Non-nuclear pitch accents in ApE, while still overall more commonly H*, had a relatively higher proportion of L+H* than the other two varieties.

Two logistic mixed-effects models were built testing to see if the differences in Tables 3 and 4 were significant: first, a model predicting nuclear status based on the pitch accent type and the variety, and an interaction between the two, and then, a model predicting pitch accent type based on nuclear status and the variety, and an interaction between the two. Random intercepts by speaker were included in both models. Contrasts were treatment coded, with non-nuclear as the reference level for nuclear status, H* as the reference level for pitch accent type, and Jewish English as the reference level for variety. Variety was then releveled to enable comparisons between African American English and Appalachian English.

For the model predicting nuclear status (Nuc $\sim$ PAType*Variety + (1|Speaker)) the intercept was near significant, with more non-nuclear pitch accents than nuclear ones (β = −0.2075, SE = 0.1122, z = −1.850, p = 0.0643). Pitch accent type was a significant predictor of nuclear status, with L+H* being more likely to be non-nuclear (β = 0.7859, SE = 0.1735, z = 4.529, p < 0.001). There was also a significant interaction between variety and pitch accent type with L

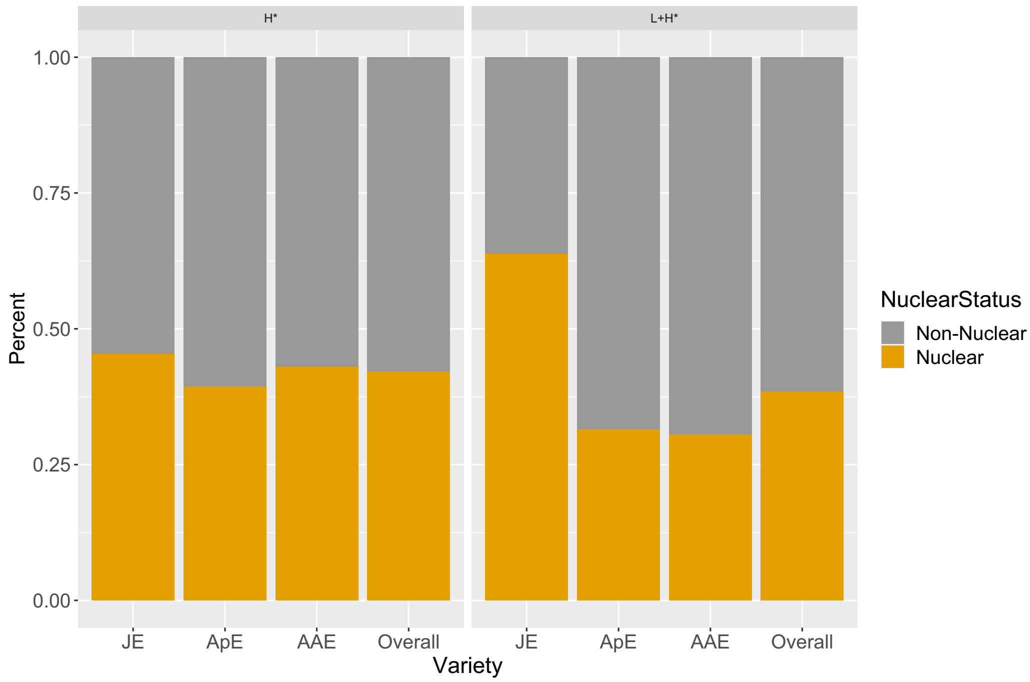

Fig. 6. Nuclear status by pitch accent and variety.

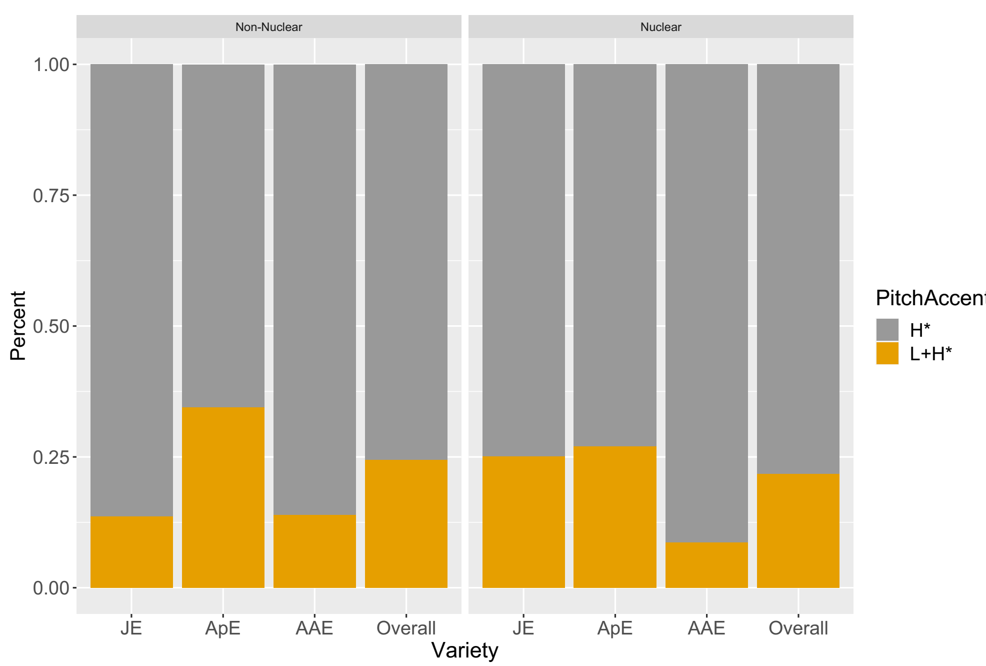

Fig. 7. Pitch accent by nuclear status and variety.

+H* being more likely to be nuclear in Jewish English compared to Appalachian English (β = -1.3255, SE = 0.2881, z = -4.601, p < 0.001) and African American English (β = -1.1431, SE = 0.2084, z = -5.486, p < 0.001). There were no significant differences between African American English and Appalachian English.

For the model predicting pitch accent type (PAType $\sim$ Variety*Nuc + (1|Speaker)) the intercept was significant, with more H* than L+H* pitch accents (β = -1.9196, se = 0.2186, z = -8.780, p < 0.001). There was a significant effect of variety, with Appalachian English having more L+H*s than Jewish English (β = 1.2783, se = 0.2852, z = 4.482, p < 0.001). Nuclear pitch accents were more likely to be H* (β = 0.7908, se = 0.2878, z = 4.549, p < 0.001). There was a significant effect of nuclear status by variety, with nuclear pitch accents in Appalachian English (β = -1.3345, se = 0.2878, z = -4.637, p < 0.001) and African American English (β = -1.1625, se = 0.2091, z = -5.561, p < 0.001) being less likely to be L+H* compared to Jewish English. Appalachian English also had significantly more L+H*s than African American English (β = 1.1768, se = 0.3055, z = 4.178, $p < 0.005$).

4.2. Phonetic realization of H* and L+H*

A smaller set of data was used for the acoustic analysis. A number of the H* tokens had pitch tracking errors which made extracting f0 impossible; this was particularly true for the African American English speakers who were generally creakier than the Appalachian English and Jewish English speakers. For H*, 11 tokens were excluded from the Jewish English speakers, 76 from the African American English speakers, and 50 from the Appalachian English speakers, for a total of 2629 tokens remaining. For L+H*, 1 token was excluded from Jewish English, 2 tokens from African American English, and 3 tokens from Appalachian English, for a total of 835 tokens remaining.

A significant portion of the H* tokens were produced with relatively high, flat f0 with no obvious peak or preceding low. A little under half (1116) of the remaining H*s fit this profile: 398 of the African American English tokens (out of 810; 49%), 540 of the Appalachian English tokens (out of 1178; 45%), and 178 of the Jewish English tokens (out of 777; 23%). These pitch accents were excluded from the analyses related to the rise (rise span and slope), but were included in the analyses for peak height and peak alignment.

Linear mixed-effects models were built using the lmer package in R. The initial model for all acoustic factors included a three-way interaction among nuclear status, variety, and pitch accent type as fixed effects, and random slopes by nuclear status, and random intercepts by speaker; models with more complex random effects structures failed to converge. Models were stepped down using log-likelihood comparison tests, removing effects that did not improve the model. Significance for interaction terms and fixed effects that were not involved in interaction terms were assessed using log-likelihood comparison tests between models with and without the effects. As this procedure could not be done on fixed effects involved in interaction terms, these fixed effects were taken to be significant if $t > |2|$.

Effects were coded as treatment contrasts, as above. The reference level was H* for the pitch accents, non-nuclear pitch accents for nuclear status, and Jewish English as the reference level for variety to compare this variety to African American English and Appalachian English. The factor levels were then releveled to compare African American English and Appalachian English.

4.2.1. Peak height

As noted above, the peak height model included all tokens for which a reliable f0 measurement could be taken of the peak; this was 3463 tokens. The final model included fixed

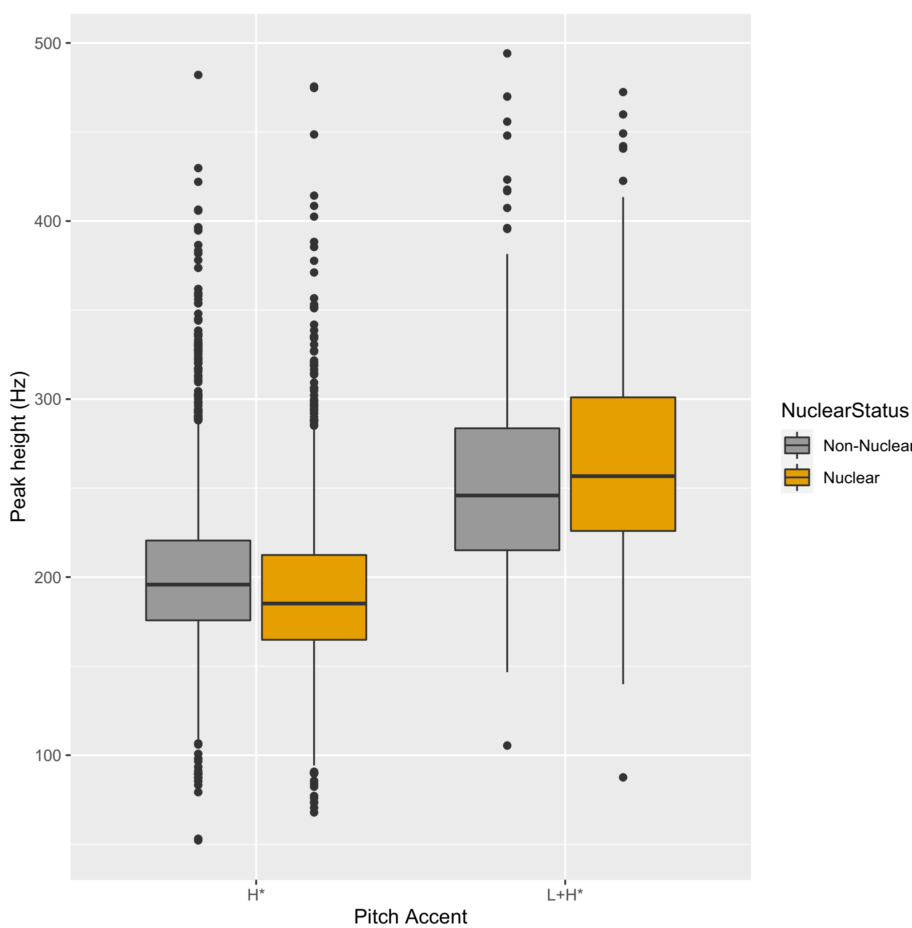

Fig. 8. Peak height by pitch accent and nuclear status.

effects for variety, nuclear status and pitch accent type, as well as interactions between pitch accent type and variety, and nuclear status and pitch accent type (H $\sim$ PAType*Variety + Nuc*PAType + (1|Speaker)). Pitch accents labeled as L+H* had significantly higher peaks than pitch accents labeled as H* (β = 51.05, SE = 4.127, t = 12.368). Likewise, nuclear pitch accents had slightly lowered peaks compared to non-nuclear pitch accents (β = −10.075, SE = 1.691, t = −5.957); however, there was a significant interaction between nuclear status and pitch accent type, with H*s being, overall, slightly lower in nuclear position, and L+H*s, slightly higher, compared to non-nuclear pitch accents of the same type, as can be seen in Fig. 8 (β = 15.504, SE = 3.568, t = 4.339, $p < 0.001$). There was a significant interaction of pitch accent by variety, with African American English (β = −24.853, SE = 5.769, t = −4.308) and Appalachian English (β = −10.58, SE = 4.339, t = −2.441) both having lower L+H* peaks compared to Jewish English, and Appalachian English having higher peaks than African American English (β = 14.262, SE = 4.976, t = 2.866), as can be seen in Fig. 9. The lower peak heights may be reflective of lower f0 in African American English, as documented in Holt & Rangarathnam (2018) (but see also other studies suggested an expanded pitch range overall, including the use of falsetto, e.g., Tarone, 1973; as such, this may be a stylistic choice to use the lower end of a wider range).

4.2.2. Rise span

As noted above, rise span was only looked at for tokens with a clear rise; this was 2347 tokens. Rise span here is defined as the f0 of the peak minus the f0 of the preceding valley. The final model included fixed effects for pitch accent type and variety, and an interaction between the two (RiseSpan $\sim$ PAType*Variety + (1 + PAType|Speaker)). As expected based on prior research, L+H* had larger rise spans than H* (β = 32.947, SE = 4.077, t = 8.590). Overall, the rise spans in Jewish English were wider than those in African American English (β = −15.818, SE = 4.309, t = −3.671) and Appalachian English (β = −9.252, SE = 4.309, t = −2.220). Finally, there was a significant interaction between pitch accent type and variety, with Jewish English having a smaller difference in rise spans between H* and L+H*s compared to Appalachian English (β = 13.538, SE = 5.416, t = −2.500), but a larger difference compared to African American English (β = −18.017, SE = 6.204, t = −2.904). Appalachian English had larger difference in rise spans between H* and L+H* than African American English (β = 31.555,

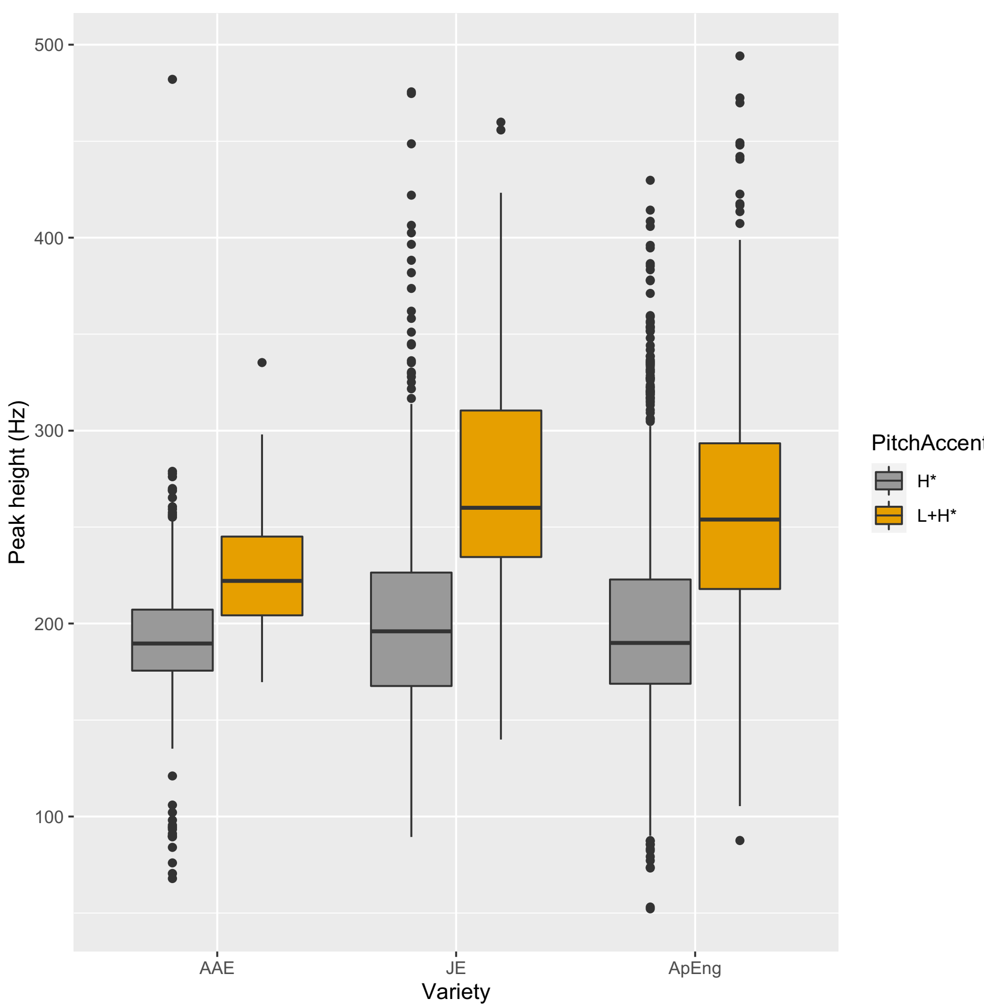

Fig. 9. Peak height by variety and pitch accent type.

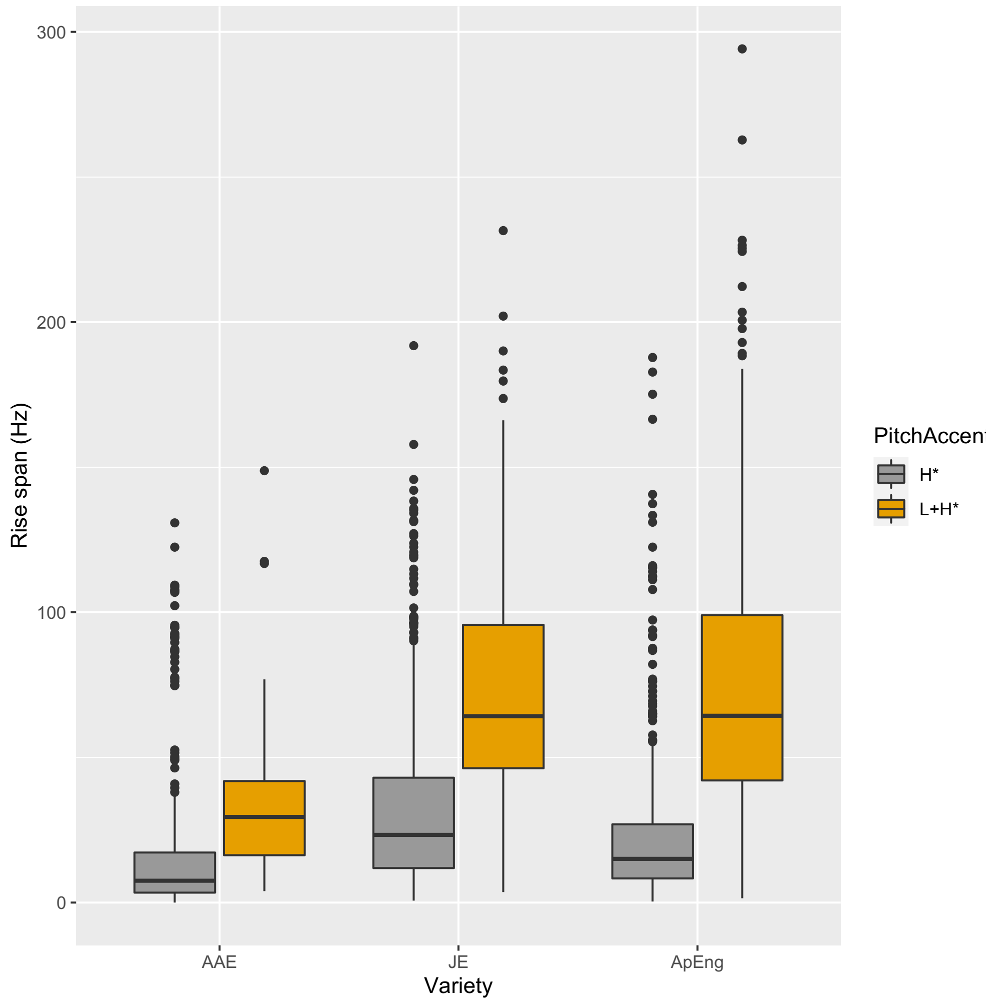

Fig. 10. Rise span by pitch accent and variety.

SE = 5.880, t = 5.367). The differences in rise span by variety and pitch accent can be seen in Fig. 10.

4.2.3. Slope

As noted above, slope was only looked at for tokens with a clear rise; this was 2347 tokens. Slope was calculated as the difference, in Hz, between the peak and the preceding valley, divided by the difference in the time of the peak and valley, in milliseconds. The final model for slope included fixed effects for pitch accent type and data set (Slope ~ PAType + Variety + (1 + PAType|Speaker)). L+H* had steeper slopes than H*s (β = 1.6294, SE = 0.1515, t = 10.758). Jewish English had steeper slopes than African American English (β = −1.3837, SE = 0.3829, t = −3.652), as did Appalachian English (β = 1.7937, SE = 0.3783, t = 4742). The differences between the pitch accents by variety can be seen in Fig. 11.

4.2.4. Peak alignment

As noted above, peak alignment was only looked at for tokens with a clear rise; this was 2348 tokens. Peak alignment was calculated as the difference between the peak of the rise and the onset of the stressed vowel. A larger value indicates a later peak; the peaks were also occasionally negative due to the peak occurring, e.g., in a sonorant sound prior to the stressed vowel. As noted above, these peaks could occur anywhere within the stressed syllable, including after the offset of the vowel[4].

The final model for peak alignment included fixed effects for pitch accent type, variety, and nuclear status, and all two- and three-way interactions between the fixed effects (PeakAlignment ~ (PAType + Variety + Nuc)^3 + (1 + PATyp e|Speaker)). L+H* had a later peak alignment than H*, as expected, and can clearly be seen in Fig. 10 (β = 7.22038, SE = 0.79257, t = 9.110). Nuclear pitch accents had earlier peak alignments than non-nuclear pitch accents (β = −3.55, SE = 0.399, t = −8.897) as expected due to tonal crowding from the adjacent boundary tones. There was a significant two way interaction between variety and nuclear status, with Jewish English having a larger difference between nuclear and non-nuclear pitch accents than both African American English (β = 3.625, SE = 0.573, t = 6.302) and Appalachian English (β = 2.08, SE = 0.52401, t = 3.982). Finally, there

[4] An anonymous reviewer asked if some of these late aligned L+H* were L*+H; we were mindful of the potential for ambiguity between these two pitch accents especially on one syllable words, and this was a topic of discussion in determining inter-rater reliability.

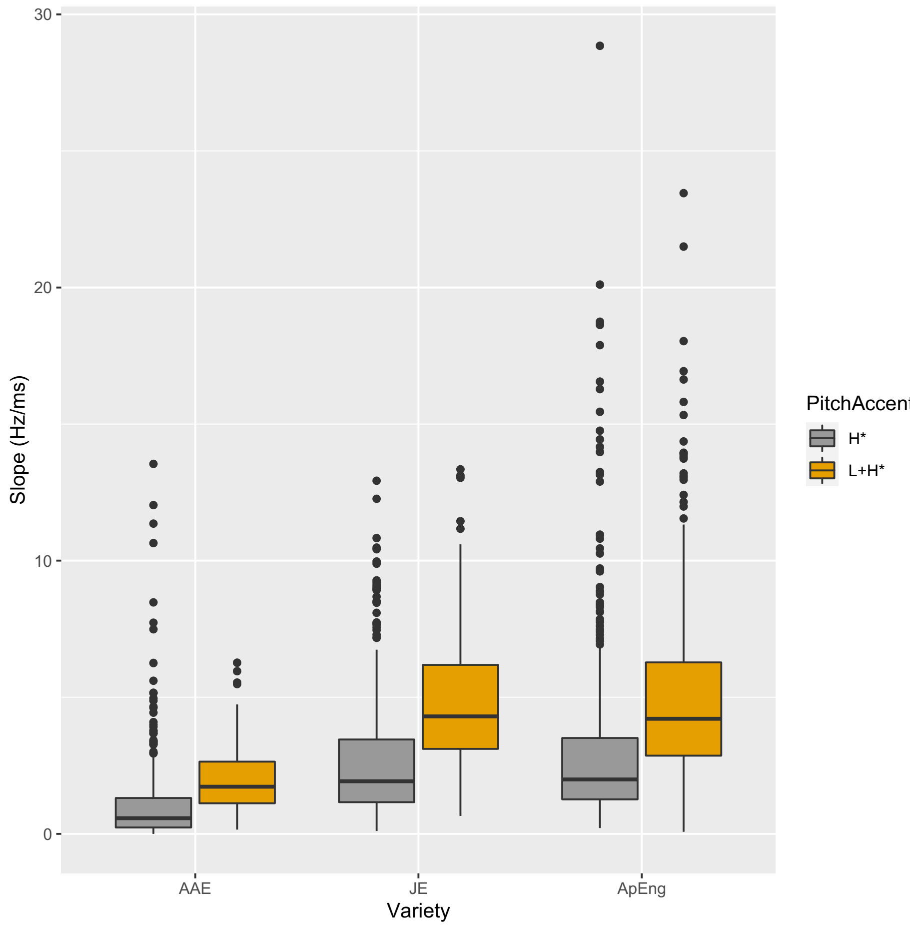

Fig. 11. Difference in slope by pitch accent and variety.

was a three-way interaction between variety, nuclear status, and pitch accent, with the difference between nuclear and non-nuclear H* and L+H* being greater in Jewish English due to earlier peak alignments for H*, compared to both African American English (β = −3.84799, SE = 1.538, t = −2.501) and Appalachian English (β = −2.40094, SE = 1.112, t = −2.158), as can be seen in Fig. 12.

5. Discussion

The common finding across all of these models were differences between H* and L+H* within each variety; however, the phonetic details and use of these pitch accents appeared to differ between the varieties. First, there are statistically significant differences between the pitch accents coded as H* and those coded as L+H*, with L+H* having higher peaks, steeper slopes, larger rise spans, and later peak alignments than H*.

However, we see differences in the exact phonetic implementation of these contrasts (realizational variation, under Ladd's typology) as can be seen by the significant interactions of variety and pitch accent type in several of the models. We can see differences in the size of the phonetic difference between these two varieties differs, as summarized in Table 5. Jewish English had the largest difference between H* and L +H* on three of the measures (peak height, slope, and peak alignment); African American English, the smallest difference across all four measures. We also see overlaps between the categories across the varieties. This overlap is most notable in the peak height, and slope of the rise, as shown in Figs. 13 and 14, where we see almost complete overlap between the slope of the rise for H* in Jewish English and Appalachian English with the slope of the rise for L+H* in African American English. We also see overlap, but not quite to the same degree, between the peak height of L+H*s in African American English and H*s in the other two varieties.

This finding has implications for the debate over the category boundary between H* and L+H*, as well as whether or not this distinction is present at all in varieties of American English. At the very least, these findings show that the annotators made a reliable distinction in what pitch accents were labeled as H* vs. L+H* in each of the three varieties. The possibility still exists, however, that what are being labeled are two different phonetic realizations of the same category, rather than two separate phonological categories. If we are in fact dealing with two distinct phonological categories, these findings show that, just as for segmental features, different varieties may place category boundaries in different locations, and what counts as intercategorical variation in one variety may become cross-categorical variation in another. If we are dealing with phonetic realizations within the same category, we still see that while all three varieties do have a distinction, the distinction is made in different ways in all three varieties. In any case, we

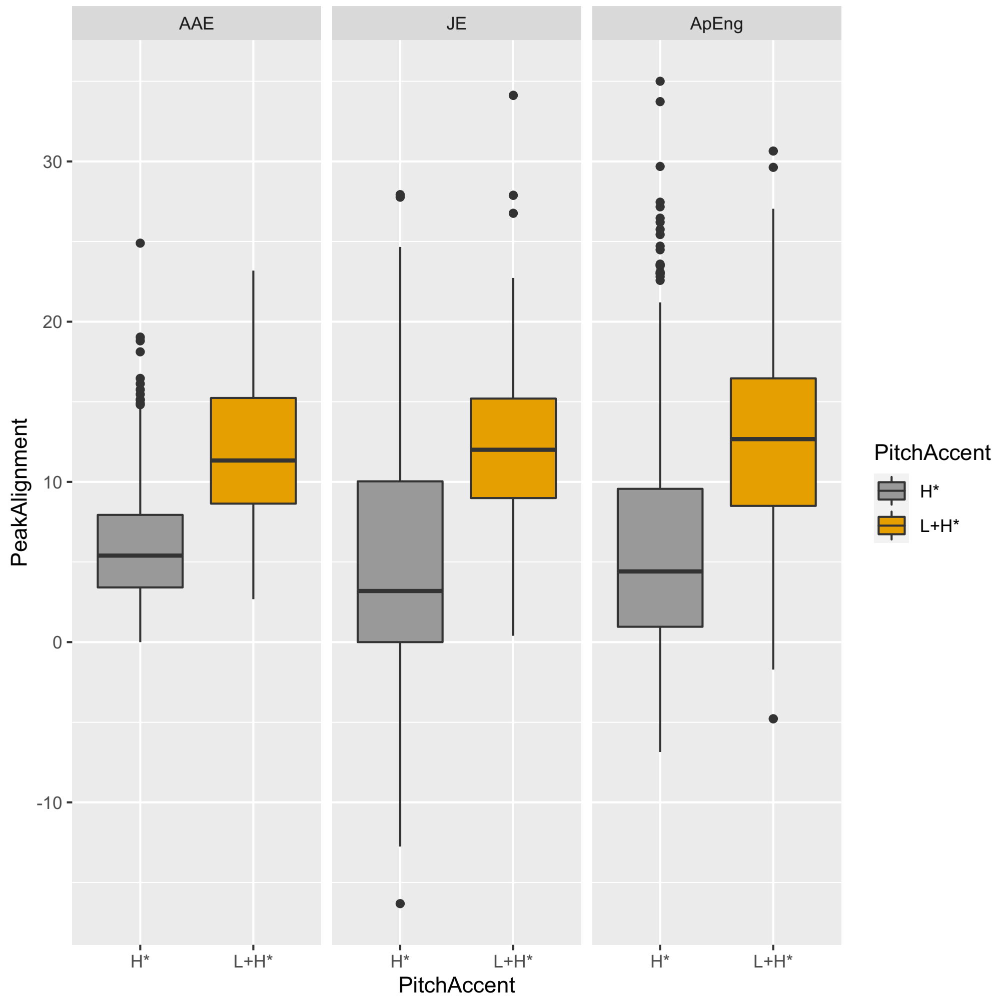

Fig. 12. Peak alignment, by variety, pitch accent type, and nuclear status.

Table 5
Summary of phonetic measurements of H* and L+H*.

	H*	L+H*	Difference between H* and L+H*
Peak height	AAE: 192 Hz JE: 200 Hz ApE: 202 Hz	AAE: 227 Hz JE: 273 Hz ApE: 259 Hz	AAE: 35 Hz JE: 73 Hz ApE: 57 Hz
Slope	AAE: 1.135 Hz/ms JE: 2.56 Hz/ms ApE: 2.97 Hz/ms	AAE: 2.01 Hz/ms JE: 4.88 Hz/ms ApE: 4.93 Hz/ms	AAE: 0.86 Hz/ms JE: 2.32 Hz/ms ApE: 1.96 Hz/ms
Rise span	AAE: 17 Hz JE: 32 Hz ApE: 23 Hz	AAE: 32 Hz JE: 73 Hz ApE: 76 Hz	AAE: 16 Hz JE: 41 Hz ApE: 53 Hz
Peak alignment	AAE: 6.01 ms JE: 4.95 ms ApE: 6.01 ms	AAE: 11.6 ms JE: 12.47 ms ApE: 12.71 ms	AAE: 5.59 ms JE: 7.52 ms ApE: 6.69 ms

found that H* and L+H* were useful as *labels* for annotation, in that we saw variation in how they both were produced and were used. As such, while this work, like Arvaniti and Garding (2007), shows the importance of taking a range of varieties into account when looking for phonological distinctions, without more carefully controlled materials, it is still somewhat of an open question what distinction is being annotated here.

We also see differences in the overall rate of pitch accent use, relative proportions of L+H* and H*, and nuclear vs. non-nuclear status of the pitch accents across the varieties. AAE and ApE use more pitch accents per IP phrase than JE. ApE has relatively more L+H*s than JE and AAE; however, JE has more nuclear L+H*s than ApE and AAE. As noted above, some of these differences may be driven by the passage choice; however, these results align with previous work using other data from these interviews, and in some cases, other speakers of the same varieties. For example, the greater proportion of nuclear L+H*s in JE aligns with work showing greater use of rising-falling nuclear contours (L+H* L-L% and L+H* !H-L%) in JE both from interview data from these same speakers as well as another group of JE speakers from Ohio (Burdin, 2016). For the AAE data, the lack of difference in rate of use of L+H* vs. H* with nuclear status, as well as the later peak alignment, parallel the findings of Holliday (2016). For the ApE data, finding relatively more L+H*s and differing phonetic implementation parallels work in Reed (2016, 2020).

These differences are suggestive of differences in the meanings and/or functions of these pitch accents, and perhaps pitch accenting in general, across these three varieties (semantic variation, under Ladd's typology, as well as possibly phonotactic variation). As discussed in the introduction, previ-

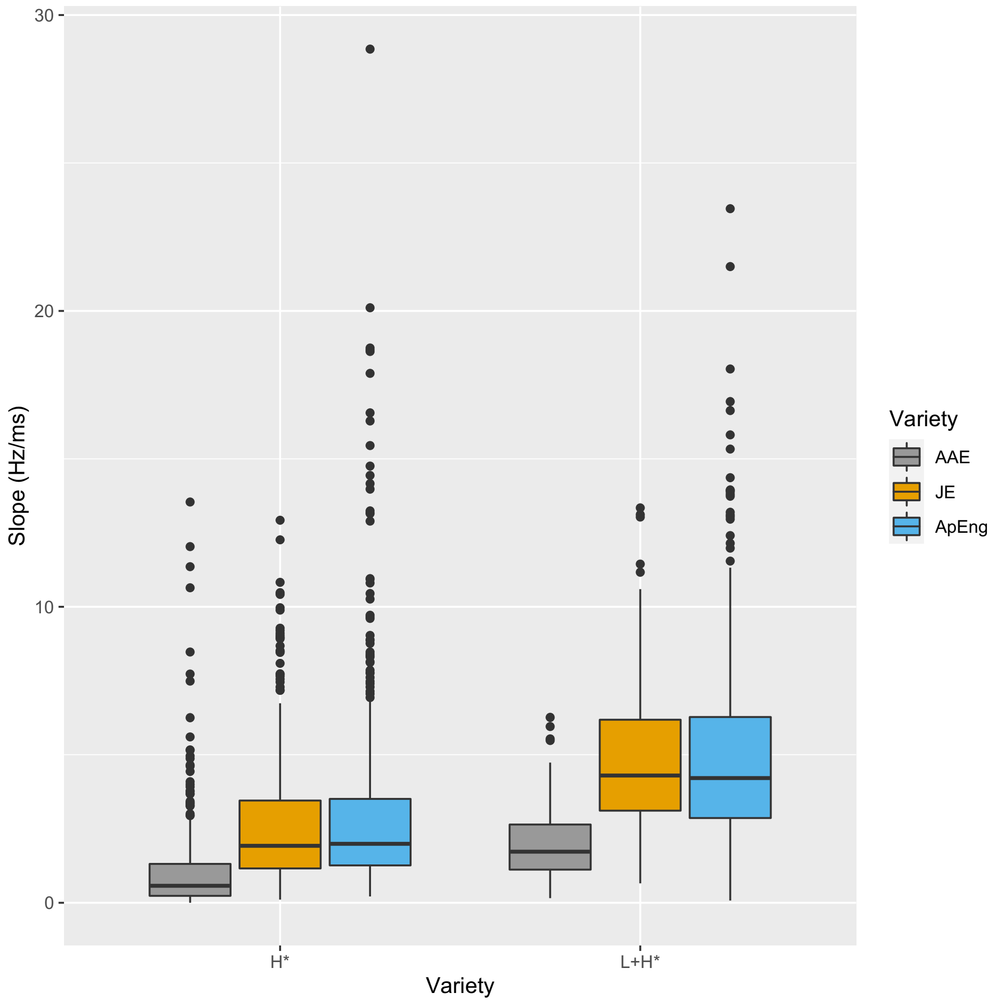

Fig. 13. Slope by pitch accent and variety.

ous work has called into a question a lack of a one-to-one link between pitch accenting and information status or structure in American English, in favor of more probabilistic approaches. We could thus imagine considerable variation in how strong or weak, e.g., the link between L+H* and focused items, or givenness and deaccenting, is in each variety.

We maybe then see an interaction with phonetic distinctiveness and semantic use. Take the AAE data here, in which the H* and L+H* were both closer to each other in phonetic space than in the other two varieties, and the L+H* looked less like the descriptions in the standard materials. This finding dovetails intriguingly with other work suggesting greater use of L+H* in AAE compared to other varieties of English, as other work finding the use of L+H* in broad focus in AAE (Holliday, 2016). These findings may be more indicative of a single H*/L+H* category, which may not show as much variation based on the focal status of the item being marked. Annotating this distinction may still be useful, but it may be indicative more of phonetic variation within the same category rather than two distinct phonological categories.

Looking at JE and ApE, we see greater phonetic differences between the H* and L+H*, and may be more justified in thinking of these as indicative of two different phonological categories. We also see differences in how they're used: Jewish English has a tendency for L+H* pitch accents to be in nuclear positions. This finding lines up with previous work noting both greater amounts of rising-falling contours in JE, as well as claims that these are also used in different pragmatic contexts from MAE (Burdin, 2020; Weinreich, 1956). With greater confidence in a phonological distinction here, we can more carefully look at semantic variation using more carefully controlled materials.

These results suggest some potential pitfalls for annotation of varieties that one is not familiar with and/or that differ significantly from standardized varieties of English. And, in fact, we did see some issues in the initial round of annotations. For example, AAE L+H*s might be heard as H*s, and H*s as !H*, L* or even unaccented, particularly by someone more familiar with JE. As part of our inter-annotator reliability check, we made confusion matrices to look for areas of disagreement between the coded files[5]. The first author (a JE speaker) and the second author (an AAE speaker) differed on the annotations of the AAE speakers of what the second author had labeled L+H*: the first author tended to label them as H*s where there

[5] It should be noted that these were somewhat rough passes through the data, which was still in a somewhat early stage (with, e.g., occasional typos in the annotations); as such, we don't report any exact numbers here.

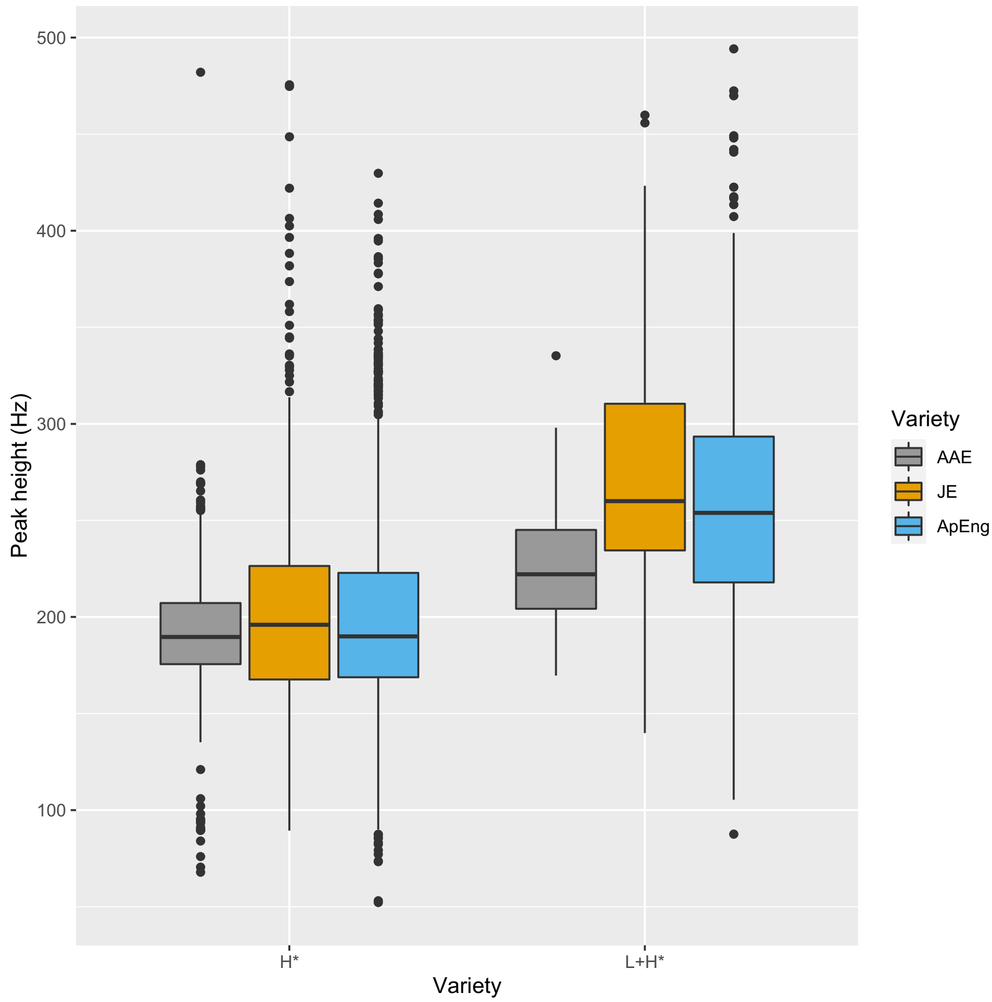

Fig. 14. Peak height by varieties and pitch accent.

was disagreement; likewise, H*s were labeled as being unaccented when they were in disagreement in the initial annotation. Interestingly, we didn't see as many "errors" in the other direction—e.g., things that the first author had labeled as H*s in JE being labeled as L+H*, or unaccented words being labelled with H* by the second author. This suggests that annotators using MAE-ToBI should pay attention to such differences as part of the annotation process, but it also presents challenges for measures of interrater reliability in contexts where coders may command different varieties of English.

The consistency following training and discussion presents a positive aspect of the phonological nature of ToBI annotation: having agreed that there were likely two pitch accent categories here (either phonetic or phonological), we were able to apply these labels to different phonetic realizations in a consistent way. Our own experience thus suggests both caution and optimism for the use of MAE-ToBI on these varieties. The three authors showed consistency in how they approached the data, and did some adaptation to the very different H* and L+H*s in all three varieties; however, all three authors also specifically work on intonational variation and went into the task knowing that L+H*s might sound different in all of the varieties. This data also shows that just because pitch accents are labeled the same across varieties (or, potentially, across different types of reading tasks), they may not sound the same, or be used in the same way. Introducing a wider range of audio files and examples in ToBI training materials, as well as simply making annotators aware of potential differences in both semantics and phonetic realization, would be useful for future annotation guides and training.

This work also shows some of the drawbacks of ToBI annotation, in that no concrete mechanism exists to indicate these phonetic differences in early stages of annotation. More phonetically driven labeling systems under development such as PoLaR (Ahn, Veilleux, Shattuck-Hufnagel, & Brugos, 2020), and for British English, IViE (Grabe, Kochanski, & Coleman, 2005) may help bridge this gap and allow for initial exploration of differences in phonetic realization without the multi-step process outlined here (ToBI annotation, marking of phonetic landmarks, extraction, etc.). This type of annotation will also allow for more neutral exploration of potential variation therein without *a priori* assumptions of phonological categories.

Overall, this work emphasizes the warnings from the original creators of MAE-ToBI: that it was designed for use on particular varieties of English, and may not work as well for others without modifications. As such, annotators should be cautious about using the system, unaltered, on other varieties. However, this work also shows that the underspecification of the system works to its advantage, in that researchers working on American English may not need to start from scratch.

CRediT authorship contribution statement

Rachel Steindel Burdin: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft. **Nicole R. Holliday:** Conceptualization, Methodology, Data curation, Writing – review & editing. **Paul E. Reed:** Conceptualization, Methodology, Data curation, Writing – review & editing.

Acknowledgements

The authors would like to thank Emelia Benson Meyer for help with data processing, as well as audience members at ETAP-4 and Speech Prosody 2018. An initial analysis of the L+H* data appeared in the conference proceedings of the latter.

Appendix A. Reading passages

Highlighted words are content words, as defined in Corver and van Riemsdijk (2001).

Comma Gets a Cure (Jewish English speakers); 199 content words.

Well, here's a **story** for you: **Sarah Perry** was a **veterinary nurse who** had been **working daily** at an **old zoo** in a **deserted district** of the **territory**, so she was **very happy** to **start** a **new job** at a **superb private practice** in **north square** near the **Duke Street Tower**. **That area** was **much nearer** for her and **more** to her **liking**. **Even** so, on her **first morning**, she **felt stressed**. She **ate** a **bowl** of **porridge**, **checked** herself in the **mirror** and **washed** her **face** in a **hurry**. Then she **put** on a **plain yellow dress** and a **fleece jacket**, **picked** up her **kit** and **headed** for **work**.

When she **got** there, there was a **woman** with a **goose waiting** for her. The **woman gave Sarah** an **official letter** from the **vet**. The **letter implied** that the **animal could** be **suffering** from a **rare form** of **foot** and **mouth disease**, which was **surprising**, **because normally** you **would only expect** to **see** it in a **dog** or a **goat**. **Sarah** was **sentimental**, so **this made** her **feel sorry** for the **beautiful bird**.

Before **long**, that **itchy goose began** to **strut** around the **office** like a **lunatic**, **which made** an **unsanitary mess**. The **goose's owner**, **Mary Harrison**, **kept calling**, "**Comma**, **Comma**," **which Sarah thought** was an **odd choice** for a **name**. **Comma** was **strong** and **huge**, so it **would take** some **force** to **trap** her, but **Sarah** had a **different idea**. **First** she **tried gently stroking** the **goose's lower back** with her **palm**, then **singing** a **tune** to her. Finally, she **administered ether**. Her **efforts** were **not futile**. In **no time**, the **goose began** to **tire**, so **Sarah** was able to **hold** onto **Comma** and **give** her a **relaxing bath**.

Once **Sarah** had **managed** to **bathe** the **goose**, she **wiped** her off with a **cloth** and **laid** her on her **right side**. Then **Sarah confirmed** the **vet's diagnosis**. Almost **immediately**, she **remembered** an **effective treatment** that **required** her to **measure** out a **lot** of **medicine**. **Sarah warned** that **this course** of **treatment might** be **expensive**—either **five** or **six times** the **cost** of **penicillin**. I **can't imagine paying** so **much**, but **Mrs**. **Harrison**—a **millionaire lawyer**—**thought** it was a **fair price** for **a cure**.

The Rainbow Passage (African American English speakers); 164 content words.

When the **sunlight strikes raindrops** in the **air**, they **act** as a **prism** and **form** a **rainbow**. The **rainbow** is a **division** of **white light** into **many beautiful colors**. **These take** the **shape** of a **long round arch**, with **its path high above**, and **its two ends apparently** beyond the **horizon**. There is, **according** to **legend**, a **boiling pot** of **gold** at **one end**. **People look**, but **no one** ever **finds** it. **When** a **man looks** for **something** beyond his **reach**, his **friends say** he is **looking** for the **pot** of **gold** at the **end** of the **rainbow**. Throughout the **centuries people** have **explained** the **rainbow** in **various ways**. **Some** have **accepted** it as a **miracle without physical explanation**. To the **Hebrews** it was a **token** that there would be **no more universal floods**. The **Greeks** used to **imagine** that it was a **sign** from the **gods** to **foretell war** or **heavy rain**. The **Norsemen considered** the **rainbow** as a **bridge** over which the **gods passed** from **earth** to their **home** in the **sky**. **Others** have **tried** to **explain** the **phenomenon physically**. **Aristotle thought** that the **rainbow** was **caused** by **reflection** of the **sun's rays** by the **rain**. Since then **physicists** have **found** that it is **not reflection**, but **refraction** by the **raindrops** which **causes** the **rainbows**. **Many complicated ideas** about the **rainbow** have been **formed**. The **difference** in the **rainbow depends considerably** upon the **size** of the **drops**, and the **width** of the **colored band increases** as the **size** of the **drops increases**. The **actual primary rainbow observed** is **said** to be the **effect** of **super-imposition** of a **number** of **bows**. If the **red** of the **second bow falls** upon the **green** of the **first**, the **result** is to **give** a **bow** with an **abnormally wide yellow band**, **since red** and **green light when mixed form yellow**. **This** is a **very common type** of **bow**, **one showing mainly red** and **yellow**, with **little** or **no green** or **blue**.

Arthur the Rat (Appalachian English speakers). 297 content words.

Once upon a **time** there was a **rat who couldn't make** up his **mind**. **Whenever** the **other rats asked** him if he **would like** to **come** out **hunting** with them, he would **answer** in a **hoarse voice**, "I **don't know**." And **when** they **said**, "**Would** you **rather stay** inside?" he **wouldn't say yes**, or **no** either. He'd **always shirk making** a **choice**.

One fine day his **aunt Josephine said** to him, "**Now look here**! **No one** will **ever care** for you if you **carry** on **like this**. You have **no more mind** of your own than a **greasy old blade** of **grass**!"

The **young rat coughed** and **looked wise**, as **usual**, but **said nothing**.

"**Don't** you **think** so?" **said** his **aunt stamping** with her **foot**, for she **couldn't bear** to **see** the **young rat** so **coldblooded**.

"I **don't know**," was **all** he ever **answered**, and then he'd **walk** off to **think** for an **hour** or **more**, whether he **would stay** in his **hole** in the **ground** or go out into the **loft**.

One night the **rats heard** a **loud noise** in the **loft**. It was a **very dreary old place**. The **roof let** the **rain come washing** in, the **beams** and **rafters** had all **rotted** through, so that the **whole thing** was **quite unsafe**.

At **last one** of the **joists gave way**, and the **beams fell** with **one edge** on the **floor**. The **walls shook**, and the **cupola fell** off, and **all** the **rats' hair stood** on **end** with **fear** and **horror**.

"**This won't do**," **said** their **leader**. "We **can't stay cooped** up **here any longer**." So they **sent** out **scouts** to **search** for a **new home**.

A **little later** on that **evening** the **scouts came back** and **said** they had **found** an **old-fashioned horse-barn where** there would be **room** and **board** for **all** of them.

The **leader gave** the **order** at **once**, "**Company fall** in!" and the **rats crawled** out of their **holes right away** and **stood** on the **floor** in a **long line**.

Just then the **old rat caught sight** of **young Arthur** — **that** was the **name** of the **shirker**. He **wasn't** in the **line**, and he **wasn't exactly** outside it–he **stood** just by it.

"**Come** on, **get** in **line**!" **growled** the **old rat coarsely**. "**Of course** you're **coming** too?"

"I **don't know**," **said Arthur calmly**.

"Why, the idea of it! You don't think it's safe here any more, do you?"

"I'm not certain," said Arthur undaunted. "The roof may not fall down yet."

"Well," **said** the **old rat**, "we **can't wait** for you to **join** us." Then he **turned** to the **others** and **shouted**, "**Right** about **face**! **March**!" and the **long line marched** out of the **barn** while the **young rat watched** them.

"I **think** I'll go **tomorrow**," he **said** to himself, "but then **again**, **perhaps** I **won't** — it's so **nice** and **snug** here. I **guess** I'll go **back** to my **hole** under the **log** for a while just to **make** up my **mind**."

But **during** the **night** there was a **big crash**. Down **came beams**, **rafters**, **joists** — the **whole business**.

Next **morning** — it was a **foggy day** — some **men came** to **look** over the **damage**. It **seemed odd** that the **old building** was **not haunted** by **rats**. But at last **one** of them **happened** to **move** a **board**, and he **caught sight** of a **young rat**, **quite dead**, **half** in and **half** out of his **hole**.

Thus the **shirker got** his **due**, and there was **no mourning** for him.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wocn.2022.101163.

References

Ahn, B., Veilleux, N., Shattuck-Hufnagel, S., & Brugos, A. (2020). *PoLaR Annotation Guidelines v0.9.5*. Retrieved from https://osf.io/a6rcm/.

Arvaniti, A., & Garding, G. (2007). Dialectal variation in the rising accents of American English. In J. Cole & J. I. Hualde (Eds.), *Phonology and phonetics 4–3* (pp. 547–575). Mouton de Gruyter.

Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German. *Journal of Phonetics, 32*, 177–197.

Barnes, J., Veilleux, N., Brugos, A., & Shattuck-Hufnagel, S. (2010). The effect of global F0 shape on the perception of tonal timing contrasts in American English intonation. *Speech Prosody 2010.*

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Becker, K. (2014). (r) we there yet? The change to rhoticity in New York City English. *Language Variation and Change, 26*(2), 141–168. https://doi.org/10.1017/S0954394514000064.

Beckman, M. E., & Elam, G. A. (1997). *Guidelines for ToBI labelling, version 3.0*.

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 9–54). Oxford University Press.

Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer. Version 6.0.12. Retrieved from http://www.praat.org/.

Bucholtz, M. (2003). Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics, 7*(3), 398–416.

Burdin, R. S. (2016). *Variation in form and function in Jewish English intonation*. The Ohio State University.

Burdin, R. S. (2017). New notes on the rise-fall contour. *Journal of Jewish Languages, 5* (2), 145–173.

Burdin, R. S. (2020). The perception of macro-rhythm in Jewish English. *American Speech, 95*(3), 263–296.

Burdin, R. S., Phillips-Bourass, S., Turnbull, R., Yasavul, M., Clopper, C. G., & Tonhauser, J. (2015). Variation in the prosody of contrastive focus in head-edge and edge-marking languages. *Lingua, 165*, 254–276.

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: A probablistic perspective. *Language, 86*(1), 1–42.

Chodroff, E., & Cole, J. (2018). Information structure, affect, and prenuclear prominence in American English. In *Proceedings of Interspeech 2018* (pp. 1848–1852).

Clopper, C. G., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics, 39*, 237–245.

Cole, J., Britt, E. R., Thomas, E. R., & Coggshall, E. L. (2005). Intonational Distinctiveness of African American English. Presented at *New Ways of Analysing Variation 34, New York*.

Corver, N., & van Riemsdijk, H. (2001). *Semi-lexical categories: The function of content words and the content of function words*. De Gruyter Mouton.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge University Press.

Dainora, A. (2002). Does intonational meaning come from tones or tunes? Evidence against a compositional approach. In: B. Bel & I. Marline (Eds.), *Proceedings of the Speech Prosody 2002 Conference* (pp. 325–238). Aix-en-Provence: Laboratoire Parole et Language.

Dainora, A. (2006). Modelling intonation in English: A probabilistic approach to phonological competence. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory Phonology 8* (pp. 107–132). Mouton de Gruyter.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology, 41*, 87–100.

Farinella, A., Yu, K. M., Brugos, A., & Green, L. (2021). Biases from MAE-ToBI intonational transcription conventions in the intonational analysis of African American English. *Presented at NWAV 49 (University of Texas at Austin/Virtual)*.

Fletcher, J., Grabe, E., & Warren, P. (2006). Intonational variation in four dialects of English: The High Rising Tune. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 390–409). Oxford University Press.

Fletcher, J., & Stirling, L. (2014). Prosody and discourse in the Australian English Map Task. In J. Durand, U. Gut, & G. Kristoffersen (Eds.), *The Oxford Handbook of Corpus Phonology* (pp. 562–575). Oxford University Press.

Gold, D. L. (1985). Jewish English. In J. Fishman (Ed.), *Readings in the sociology of Jewish languages* (pp. 280–298). E. J. Brill.

Goldsmith, J. (1976). An overview of autosegmental phonology. *Linguistic Analysis, 2*, 23–68.

Gooden, S. (2009). Authentically Black, bona fide Pittsburgher: A first look in African American women's language in Pittsburgh. In S. L. Lanehart (Ed.), *American Women's language: Discouse, Education, and Identity* (pp. 142–164). Newcastle-upon-Tyne, U. K.: Cambridge Scholars Publishing.

Gooden, S., Drayton, K.-A., & Beckman, M. (2009). Tone inventories and tune-text alignments. *Studies in Language, 33*(2), 396–436.

Grabe, E., Kochanski, G., & Coleman, J. (2005). The intonation of native accent varieties in the British Isles: potential for miscommunication? *English Pronunciation Models: A Changing Scene*, 311–337.

Grabe, E., Post, B., Nolan, F., & Farrar, K. (2000). Pitch accent realization in four varieties of British English. *Journal of Phonetics, 28*, 161–185.

Greene, R. (2006). *Pitch accents in Appalachian English* Doctoral Dissertation. Stanford University.

Holliday, N. R. (2016). Identity Performance among Black / Biracial Men through Intonation : Examining Pitch Accents and Peak Delay. *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV 44*, *22* (2), 71–80.

Holt, Y. F., & Rangarathnam, B. (2018). F0 declination and reset in read speech of African American and White American women. *Speech Communication*, *97* (December 2017), 43–50. https://doi.org/10.1016/j.specom.2018.01.001.

Hudson, A., & Holbrook, A. (1981). A Study of reading fundamental vocal frequency of young black adults. *Journal of Speech and Hearing Research, 24*, 197–201.

Jun, S.-A. (Ed.). (2005). *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press.

Jun, S.-A. (2014). Prosodic typology: by prominence type, word prosody, and macro-rhythm. In *Prosodic typology II: The Phonology of Intonation and Phrasing*. Oxford University Press.

Kurath, H., & McDavid, R. I. (1961). *The pronunciation of English in the Atlantic states: Based upon the collections of the linguistic atlas of the eastern United States*. University of Michigan Press.

Ladd, D. R. (1996). *Intonational phonology (First edition)*. Cambridge University Press.

Ladd, D. R. (2008). *Intonational phonology (Second edition)*. Cambridge University Press.

Ladd, D. R., & Schepman, A. (2003). “Sagging transitions” between high accent peaks in English: Experimental evidence. *Journal of Phonetics, 31*, 81–112.

Ladd, D. R., Schepman, A., White, L., Quarmby, L. M., & Stackhouse, R. (2009). Structural and dialectal effects on pitch peak alignment in two varieties of British English. *Journal of Phonetics, 37*, 145–161.

Loman, B. (1975). Prosodic patterns in a Negro American dialect. In H. Ringborn, A. Ingberg, R. Norrman, K. Nyholm, R. Westman, & K. Wikberg (Eds.), *Style and text: Studies presented to Nils Erik Enkvist* (pp. 219–242). Språkförlaget Skriptor AB.

Mayo, C., Aylett, M., & Ladd, D. R. (1997). Prosodic transcription of Glasgow English: an evaluation study of GloTaBI. *Proceedings of the ESCA Workshop on Intonation: Theory, Models and Applications, Athens, Greece*, 231–234.

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal forced aligner: Trainable text-speech alignment using Kaldi. *Proceedings of the 18th conference of the international speech communication association*.

McLarty, J. (2018). African American Language and European American English Intonation Variation over time in the American South. *American Speech, 93*(1), 32–78.

Nagy, N., & Irwin, P. (2010). Boston (r): Neighbo(r)s nea(r) and fa(r). *Language Variation and Change, 22*(2), 241–278. https://doi.org/10.1017/S0954394510000062.

Newmark, K., Walker, N., & Stanford, J. (2016). ‘The rez accent knows no borders’: Native American ethnic identity expressed through English prosody. *Language in Society, 45*, 633–664. https://doi.org/10.1017/S0047404516000592.

Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.

Podesva, R. J. (2011). Salience and the social meaning of declarative contours: Three case studies of gay professionals. *Journal of English Linguistics, 39*(3), 233–264. https://doi.org/10.1177/0075424211405161.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna: Austria.

Reddy, S., & Stanford, J. (2015). A web application for automated dialect analysis. In *Proceedings of NAACL-HLT 2015*.

Reed, P. (2016). *The sociophonetics of Appalachian English: Identity, intonation, and monophthongization* Doctoral Dissertation. University of South Carolina.

Reed, P. E. (2020). The importance of rootedness in the study of Appalachian English: Case study evidence for a proposed rootedness metric. *American Speech, 95*(2), 203–226. https://doi.org/10.1215/00031283-7706532.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication, 52*, 638–651.

Ritchart, A., & Arvaniti, A. (2014). The use of high rising terminals in Southern Californian English. *POMA*.

Rosenfelder, I., Fruehwald, J., Evanini, K., & Yuan, J. (2011). FAVE (Forced alignment and vowl extraction) program suite.

Schmid, M., & Bradley, E. (2019). Vocal pitch and intonation characteristics of those who are gender non-binary. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, 2685–2689. Retrieved from https://assta.org/proceedings/ICPhS2019/papers/ICPhS_2734.pdf%0Ahttps://icphs2019.org/icphs2019-fullpapers/pdf/full-paper_178.pdf

Tarone, E. E. (1973). Aspects of intonation in Black English. *American Speech, 48*(1/2), 29–36.

Thomas, E. R. (2007). Phonological and Phonetic Characteristics of African American Vernacular English. *Language and Linguistics Compass, 1*(5), 450–475. https://doi.org/10.1111/j.1749-818x.2007.00029.x.

Thomas, E. R. (2015). Prosodic Features of African American English. In S. L. Lanehart (Ed.), *The Oxford Handbook of African American English* (pp. 420–435). Oxford University Press.

Turnbull, R., Burdin, R. S., Clopper, C. G., & Tonhauser, J. (2015). Contextual information and the prosodic realization of focus: A cross linguistic comparison. *Language, Cognition and Neuroscience, 30*(9), 1016–1076.

Veilleux, N., Shattuck-Hufnagel, S., & Brugos, A. (2006). Transcribing prosodic structure of spoken utterances with ToBI. Retrieved from https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/

Warren, P. (2005). Issues in the study of intonation in language varieties. *Language and Speech, 48*(4), 345–358. https://doi.org/10.1177/00238309050480040201.

Warren, P. (2016). *Uptalk: The phenomenon of rising intonation*. Cambridge University Press.

Weinreich, U. (1956). Notes on the Yiddish Rise-Fall Contour. In M. Halle (Ed.), *For Roman Jakobson* (pp. 633–643). Mouton de Gruyter.

Zimman, L. (2013). Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language and Sexuality, 2*(1), 1–39. https://doi.org/10.1075/jls.2.1.01zim.