Pitch Accent Use in Appalachian English

I. Introduction

Researchers have been studying variation in the pronunciation of consonants and vowels for decades, but until recently they have given relatively little attention to variation in prosody.\(^1\) When Beckman and Ayers (1994, 1997) developed the MAE ToBI (Mainstream American English Tones and Break Indices) system for analyzing intonation and phrasing, linguists gained a powerful tool for exploring this aspect of variation in American English. However, most work on intonational variation to date has focused on British English (Fletcher, et al. 2004, Grabe, et al. 2001, Vizcaino-Ortega 2002, Grabe and Post 2002) and Australian and New Zealand English (Fletcher et al. 2004). The British studies used a different, more phonetically-based system known as Intonational Variation in English (IViE) (Grabe, et al. 2001), created to analyze the kind of variation that exists between the different regional varieties of English in the British Isles (e.g. scaling and alignment of pitch accents). This study explores the question of how well ToBI can be used to capture differences between regional varieties of American English, while focusing on an understudied variety of American English, Appalachian English (AE).

Several sociolinguists have analyzed the high-rising terminal (HRT) feature of American English (in ToBI, H-H\%), but they have often linked it with broad social meanings such as uncertainty (Lakoff 1973). This work can be done without ToBI, and

\(^1\) The British School (e.g. Halliday 1967) did study intonation, but not as comprehensively as newer, ToBI-based work. They did not discuss phrasing, and in their approach, meaning was not compositional.
links one easily analyzed feature with vague meanings. ToBI offers us the ability to look at intonational variation on several more fine-tuned levels, including the utterance-internal realization of pitch accents, intonational phrase (IP)-level tunes, and compositional meaning (Pierrehumbert and Hirschberg 1990).

Researchers have begun work on intonation in a few regional varieties of American English, i.e. Minnesota English and Southern California English (Arvaniti and Garding 2004), but this remains an area highly in need of study. I give here an analysis of intonation in Appalachian English (AE), one of the most divergent, yet least described, varieties of American English (Wolfram and Christian 1976). AE is spoken in the central part of the Appalachian mountain region, including my home region of Elliott County, Kentucky, where I conducted this field research (see Figure 1). It is closely related to Southern American English (SAE). There currently exists little research on Appalachian English, variation within it, or speech from Eastern Kentucky (Luhman 1990). There is also little work that explores its relationship with Southern American English (Hazen, in progress).
I used ToBI in my analysis, but this was not necessarily the obvious choice, because it was not developed or tested for use with nonstandard American English, and also, another system, IViE, exists, created for looking at variation in intonation in the British Isles (Grabe, et al. 2001). I explored the possibility that the ToBI labels are not suited for the description of AE and SAE, but I found that ToBI was in fact useful for both nonstandard varieties, at least to the extent that the pitch accents in the varieties seem to be realized phonetically in a similar way. However, pragmatic considerations do need to be addressed.

The developers describe ToBI as:
“a framework for developing community-wide conventions for transcribing the intonation and prosodic structure of spoken utterances in a language variety, […] grounded in careful research on the intonation system and the relationship between intonation and the prosodic structures of the language […].” (http://www.ling.ohio-state.edu/~tobi/).

It describes intonation as pitch accents, phrase accents, and boundary tones, which are seen as phonological units with certain pragmatic functions, part of the grammar of a language variety. Certain issues of detail such as alignment and pitch range are considered phonetic and not included.

Researchers have used ToBI to describe intonation patterns in Mainstream American English (MAE), British Received Pronunciation, and Australian varieties. Variations on ToBI have been created for analysis of Glasgow English, and Japanese, Korean, and German, among other languages. A separate ToBI system was deemed necessary to describe Glasgow English, but not for other varieties if English in the British Isles or Australia. As I mentioned above, only a few researchers have used ToBI to analyze regional varieties of American English (Arvaniti and Garding 2004). Therefore, I used the ToBI system to analyze conversational speech of eight speakers of Appalachian English, and compare it to the intonation of five MAE speakers and four Southern American English (SAE) speakers.

I set out to explore the following questions using ToBI:

1. Do Appalachian English (AE) speakers use the Mainstream American English (MAE) inventory of pitch accents and boundary tones, and combinations thereof (tunes)?

2. Do they phonetically implement those features differently from MAE or Southern American English (SAE) speakers, in terms of scaling and alignment?
3. Do AE speakers use those features with the same pragmatic functions as MAE or SAE speakers?

I found that MAE ToBI can be applied to prosody in AE and SAE, and AE shows a significantly different distribution of pitch accent types than the other two varieties. Specifically, the L+H* pitch accent type is much more common in AE than in MAE and SAE (although there are no significant differences between the realization of L+H* pitch accents in terms of scaling or alignment). AE speakers appear to be using L+H* pitch accents in contexts where MAE and SAE speakers would use H*. Other speakers appear to use L* in this context, but to a far lesser extent. This suggests that L+H* and L* do not share the same pragmatic functions in AE that they do in MAE and SAE. I specifically found that in AE, L* is used for a wider set of information types than it is described to be used for in MAE by Pierrehumbert and Hirshberg (1990). Specifically, L* use is extended from use with non-predicated information to use with some predicated information as well. Overall, this research explores a new type of variable, specifically, greater or lesser regimentation of pitch accent choice across different varieties, demonstrating that AE is less regimented than MAE and SAE.

By using ToBI, researchers accept specific theoretical assumptions about which forms are represented distinctly in the mind of a speaker, how those forms will approximately look, and the pragmatic function of those forms. By loosening some of these assumptions that underlie it, however, ToBI can offer more insight into variation in intonation in at least one variety of American English, specifically AE. I suggest when using ToBI for nonstandard varieties of American English, researchers not assume that ToBI pitch accents will be realized in the same combinations, with the same pragmatic
functions, as in MAE. There is also reason to doubt that the pitch accent categories
themselves are the same across dialects (Arvaniti and Garding 2004).

The fact that these data come from conversational speech, whereas most phonetic
work on is based only on experimental or read speech, was another complicating factor in
my analysis. It is difficult to analyze natural speech, because it is full of disfluencies,
fillers, and interrupted thoughts, and conversational utterances do not chop up cleanly
into the meaningful chunks described in the literature. This paper therefore also gives
methodological notes for the study of conversational speech.

II. Prior research on prosodic variation

1. IViE

Most of the work to date on variation in intonation has been conducted by Grabe
and colleagues in the British Isles, using IViE, a transcription system that was developed
to investigate variation in intonation in the British Isles (Grabe, et al. 2001), where
regional dialects have been developing across a long history. American English on the
other hand is a much younger variety, and therefore conclusions about intonational
variation in the British Isles are not necessarily likely to appear in American English as
well. I will describe IViE and the findings that have come from its use, but explain how
they may or may not be relevant to intonational variation in American English. Overall,
it seems a phonetic system like IViE is not necessary for analyzing American English,
but it is necessary to loosen some of the assumptions associated with ToBI, as discussed
above.
ToBI contains four tiers, the orthographic tier (for the transcription of speech), the tone tier (for phonological labeling), the break index tier (indicating degree of juncture), and the miscellaneous tier (for noting disfluencies, laughter, etc.) On the other hand, IViE has five tiers, including a target tier and a prominence tier, in place of the ToBI break index tier. (The miscellaneous tier is optional.) IViE has much looser theoretical assumptions about pitch accent categories or phonetic implementation, and none about pragmatic function. Decisions about phonological categories are left until after rhythmic structure and pitch movement have been transcribed.

In Figure 2 is an example of IViE transcribed speech.

**Figure 2. IViE transcription**

The target tier, which is the middle gray tier in Figure 1, involves assigning low, medium, or high values to the pre-accented syllable, the accented syllable, the syllable following the accented syllable, and the final syllable in the domain of a pitch accent. The labeler also marks whether there is interpolation across the syllables following the accented syllable. This series of labels provides a clear picture of the shape of each pitch
accent. This added level of analysis allowed Grabe et al. (2001) to find that in Cambridge English, when the accented syllable is very short, phrase-final H*+L is realized as a steep fall. In Leeds English, it is realized as a very shallow fall. Due to this target tier, applying IViE to variation in American English could reveal pitch accent peak shape and height differences that are beyond the scope of this project, and have not yet been described. However, my findings indicate that three varieties of American English do not differ significantly in terms of pitch accent alignment, so this added aspect of IViE’s descriptive power would not be useful in their description.

The prominence tier, the lowest gray tier in Figure 1, labels ‘prominent’ syllables, and for Grabe and colleagues, “prominence” is different from “stress”. Prominence is defined by intensity and duration, but stress is defined by pitch. They found this was a useful variable for describing the intonation of some varieties of British English, including British Punjabi English (Grabe et al. 2001:4). See Figure 3 for an example of this difference represented in IViE transcription. I did not notice the presence of unaccented yet “prominent” syllables, so IViE likely does not improve over ToBI for describing American English in that dimension either.

\[ \text{Cambridge English} \]

<table>
<thead>
<tr>
<th>H*+L</th>
<th>0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>mH-l</td>
<td>%</td>
</tr>
<tr>
<td>P</td>
<td>%</td>
</tr>
<tr>
<td>on</td>
<td></td>
</tr>
<tr>
<td>the</td>
<td></td>
</tr>
</tbody>
</table>

\[ \text{British Punjabi English} \]

<table>
<thead>
<tr>
<th>H*+L</th>
<th>0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>mH-l</td>
<td>%</td>
</tr>
<tr>
<td>hM-l</td>
<td>%</td>
</tr>
<tr>
<td>P</td>
<td>%</td>
</tr>
<tr>
<td>on</td>
<td></td>
</tr>
<tr>
<td>the</td>
<td></td>
</tr>
</tbody>
</table>

\[ \text{Figure 3. Cross-varietal difference in rhythmic structure represented in IViE} \]
On the phonological tier, the inventory of pitch accents used in IViE is also significantly broader than that in ToBI. This allows transcribers to label intonation without having to fit it into ToBI’s narrower phonological categories, so any variation in this realm becomes observable. Neither AE nor SAE appear to contain pitch accent types that do not exist in MAE, so this added level of detail also does seem useful in analyzing varieties of nonstandard American English.

In comparing my results to those from the IViE researchers, it seems American English dialects do not differ in intonation as dramatically as those in the British Isles. A ToBI with loosened assumptions, inspired by IViE’s description of phonetic detail, seems sufficient for describing variation in American English intonation.

2. Glasgow (Gla)ToBI

Researchers are in the process of developing a ToBI-based system for the English spoken in Glasgow, Scotland (Mayo, et al. 1997), because they found several key differences between Glasgow English intonation and the varieties of English covered by MAE ToBI. It is useful to examine the reasoning supporting the development of a separate GlaToBI to see whether AE or other varieties of American English could also require a whole new system, as GlaE was deemed to. Ladd and colleagues discovered that the rising pitch accents in Glasgow English (GlaE) do not align like L+H* or L*+H do in standard British and American English varieties, and described the GlaE rising pitch accent as L*H. My results however indicate that AE and SAE do not differ significantly from MAE in terms of pitch accent alignment, but differ mainly in terms of the
proportion and function of the same accent inventory. Therefore, a whole new ToBI system is probably not necessary for varieties of American English.

GlaToBI also changes one aspect of MAE ToBI’s treatment of the effects of phrase accents. In MAE ToBI, a high phrase accent (H-) triggers upstep on a following boundary tone, but this is not the case for GlaE, so this factor has been changed in labeling guidelines for GlaToBI. I did not look explicitly at upstep in this project, so it is possible that American English varieties like AE could differ from MAE in this way, and that is a possible direction for future research. If researchers were to find such differences among varieties of American English, it could raise questions about whether new ToBI systems could be warranted.

3. Arvaniti and Garding

The salience of patterns of L+H or “rising” type pitch accents in distinguishing varieties of American English is supported both by my work and that of Arviniti and Garding (2004), who are among the first researchers to discuss regional variation in intonation in American English. They examine Minnesotan English and Southern California English for the presence of the phonologically distinct rising pitch accents H*, L*+H, and L+H* that Pierrehumbert (1980) posited for MAE. They also compare phonetic realizations of rising pitch accents in those two regional varieties of American English. They found that both Minnesota English and Southern California English contain the distinction between L*+H and L+H*. They also found that Minnesotan English may not have a contrast between H* and L+H*, and that emphasis affects rising
pitch accents differently in the two varieties. They conclude that researchers need to look more closely at intonation in different dialects.

Arvaniti and Garding’s methodology involved a production experiment in which subjects hear and are familiarized with two dialogs repeated under increasing degrees of emphasis (for them, emphasis means increased “insistence,” and brings about greater pitch variability and to some degree, duration.) The subjects are then asked to repeat the dialogs.

One dialog was designed to elicit an H* L-L% tune without emphasis and an L+H* L-L% tune with emphasis (see below).

(1)
Investigator: Who was on the phone?
Speaker: Raymona.
Investigator: Who?
Speaker: Raymona.
Investigator: It was who?
Speaker: Raymona.
Investigator: Raymond?
Speaker: Raymona!

The other was designed to elicit an L*+H L-H% tune under all levels of emphasis (see below).

(2)
Investigator: I need someone who would hold her own.
Speaker: Raymona would.
Investigator: But someone who would understand politics.
Speaker: Raymona would.
Investigator: But someone who would be really funny, too.
Speaker: Raymona would!
Investigator: But someone who would look totally hot!
Speaker: Raymona WOULD!

If with increasing emphasis, speakers produce two phonoetically distinct types of high rising pitch accents, in terms of alignment, shape, and scaling, it is considered
evidence that two distinct phonological categories exist. Arvaniti and Garding indeed found that there is a continuum of emphasis from H* to L+H* to L*+H for both dialect groups. They also found that both L*+H and L+H* exist separately for Southern California speakers, but that for Minnesota speakers, there may be no distinction between H* and L+H*. They also saw that the California speakers consistently aligned the L and H tones later than the Minnesota speakers. They also noted that the L of L+H* was scaled higher than that of L*+H, and inversely, the H of L+H* was scaled lower than that of L*+H. These results suggest that the pragmatics of L+H accents in American English are variable, and could also be in AE, but my results suggest that this is not the case for AE.

The type of production experiment done by Arvaniti and Garding could determine whether AE contains phonologically distinct H* and L+H* accents, and I suggest further work in that direction. I suspect that H* and L+H* are not phonologically distinct categories in AE, because if they were, and L+H* had the same meaning in AE as described for MAE, Minnesota English, and Southern California English, then it would mean that the AE speakers are using increased and unlikely amounts of emphasis (Arvaniti and Garding 2004, Ladd and Schepman 2003), or salience (as judged against some sort of scale) (Pierrehumbert and Hirshberg 1991). I also did not notice the kind of emphatic pragmatic function of L+H* accents in AE that is described for other varieties of American English. Rather, for some speakers, L+H* seems to appear nearly as often in standard declarative tunes as H*. For further discussion, see Section V. 1. below.
III. Appalachian English

Appalachian English is the regional variety of English spoken in part of the Appalachian Mountain range including parts of West Virginia, Eastern Kentucky, Eastern Tennessee, Western North Carolina, and Western Virginia (see Figure 1). AE has numerous unique features derived from its Scots-Irish heritage. For example, AE speakers sometimes pronounce *it* as *hit* and *ain’t* as *hain’t* (Wolfram and Christian 1976). Grammatically, for example, AE speakers might use a completive done (e.g. *I done forgot.*) (Wolfram and Christian 1976). AE also shares features with the Midlands dialect area, such as the ‘cot-caught’ merger (Hazen unpublished) and r-fulness (Wolfram and Christian 1976).

However, AE shares many features with Southern American English (SAE), and in fact many Americans do not distinguish perceptually or ideologically between AE and SAE (Greene 2002). Further, Appalachians do not identify themselves positively or negatively as such, and “Appalachian” is primarily a scholarly label, not part of the general American consciousness (Hazen unpublished, Greene 2002). Both varieties use monophthongal /ai/ (Greene 2002, Luhman 1990), participate in the Southern Vowel Shift (Thomas 2003, Fridland 2003, 2001, 2000, Feagin 2000), and use “broken” vowels (Greene 2005, Wells 1982). Both varieties are the butt of jokes that portray speakers as dumb or unsophisticated in popular American culture (Greene 2002).

A speaker of AE myself, I am never asked if I am from the “Appalachian” area, but rather, where in the South I am from. However, people generally identify me as sounding “really Southern,” from the “Deep South,” or Texas. These areas appear to be the ideological heart of the South for Americans. AE, though observably different from
the varieties spoken in the Deep South or Texas, appears to carry the stereotypes associated with the South but in an amplified way. In particular, it most strongly carries the negative “country” or “hillbilly” stereotypes of Southern speech. It does not carry the prestigious, gentrified stereotypes that Americans often associate with the old, moneyed plantation South (Greene 2002). This is likely due to the fact that the central Appalachian area where AE is spoken is rural and highly economically depressed.

I suspect that variation in intonation is one way in which Americans might be led to differ their impressions of AE speakers from those of SAE speakers, despite the fact that they cannot accurately attribute a geographic region to each variety. Perhaps because intonation is not the topic of explicit societal discourse on American regional dialects, speakers are less aware of the geographic distribution of intonational features. Further, I have observed a ‘standardization’ of AE, as younger AE speakers are using fewer of the traditional grammatical features described for AE, such as hit for it and possessive –n (e.g. yourn). However, other, phonological types of AE features, like vowel quality and intonation, seem more resistant to change, and may be the main way today in which young AE speakers recognize and signal solidarity with each other. The results of my study support this theory, with distribution of pitch accents in SAE patterning with MAE, and separate from AE.

IV. Research Method

I analyzed conversational speech from three different varieties of American English, AE, SAE, and MAE. The AE speakers consisted of eight speakers from Elliott County, including five women and three men, ranging in age from 27 to 77. (See Figure
4 for a description of each AE respondent.) I selected an approximately sixty-second segment of speech from the middle of each hour-long interview. I selected speech segments from the second half of each interview, since by that point the respondents would have become more relaxed and their speech less self-conscious. I labeled those sixty-second segments and found each to contain around seventy pitch accents. I then counted exactly seventy pitch accents for each speaker and used these as the basis for my analysis.

<table>
<thead>
<tr>
<th>Respondent</th>
<th>Sex</th>
<th>Age</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Troy</td>
<td>M</td>
<td>77</td>
<td>Retired dentist</td>
</tr>
<tr>
<td>Sharon</td>
<td>F</td>
<td>75</td>
<td>Homemaker</td>
</tr>
<tr>
<td>Margaret</td>
<td>F</td>
<td>55</td>
<td>Retired bus driver</td>
</tr>
<tr>
<td>Thomas</td>
<td>M</td>
<td>53</td>
<td>High school teacher</td>
</tr>
<tr>
<td>Elaine</td>
<td>F</td>
<td>53</td>
<td>Elementary school teacher</td>
</tr>
<tr>
<td>James</td>
<td>M</td>
<td>41</td>
<td>High school principal</td>
</tr>
<tr>
<td>Amanda</td>
<td>F</td>
<td>41</td>
<td>Homemaker</td>
</tr>
<tr>
<td>Jeannette</td>
<td>F</td>
<td>28</td>
<td>Education resource director</td>
</tr>
</tbody>
</table>

**Figure 4. Characteristics of AE respondents**

I contacted interviewees through family connections. The interviews lasted approximately one hour each, and were conducted in a quiet area of respondents’ homes. The conversations were recorded on a digital voice recorder using a lapel microphone. I discussed with them issues designed to make them feel comfortable, such as their life history and hobbies.

I also included a segment in which I asked respondents to read aloud a list of sentences of various types and a reading passage. This part of the interview was designed to produce a set of data that would be easy to compare across speakers, and contained mostly voiced consonants for easy pitch track reading. However, the read data was highly unnatural, both in intonation and phrasing. I eliminated this data from my main
analysis, because it does not represent the natural, spontaneous speech patterns of AE. However, I used it to help discern whether certain features I saw in the conversational speech were indeed AE features, or register features that disappear in read speech.

I analyzed data from only eight of ten interviews I actually collected, because the remaining two had unusual voice qualities, one wavering and one lisping. I used Praat™ (Boersma 2002) to analyze data from the interviews. I labeled each interview segment according to ToBI conventions. After doing so, I calculated the proportion of all pitch accent types in the speech of each interviewee.

A crucial decision in the process was the difference between H* and L+H*, because I focus on their distribution in my conclusions. In accordance with ToBI conventions, I labeled an L+H* when a given rising tone had a low origin or dip that could not be potentially explained by a sonorant, which are known to decrease f0 (see the L+H* in Figure 5, ten, for which f0 drops quickly already on the prior word). Some cases were clearly H*, for example Figure 5, Mabel’ll, eight, week). In other cases it was difficult to decide between H* and L+H* but I eventually chose L+H* (see Figure 6, n’he, and we’d, in which a sonorant could potentially be influencing the dip in f0 but is probably not responsible for the entire dip, or fishin’, in which f0 is partially untracked due to voiceless segments). Other times, I was tempted to label L+H* but I chose not to do so (for example, see Figure 7, switch, in which the percept roughly matched other tokens which had been previously labeled L+H* but too much f0 track was missing for me to be confident). Overall I was conservative when assigning L+H*, which strengthens my claims about unexpectedly high frequencies of L+H* in AE.
Figure 5. Clear examples of $H^*$ and $L+H^*$, James

Figure 6. Less clear examples of $H^*$ and $L+H^*$, Thomas
Figure 7. Less clear examples of $H^*$ and $L+H^*$, Margaret

The MAE and SAE speakers’ data came from the Switchboard corpus (Godfrey et al. 1992) of conversational speech. The Switchboard data was collected by recording subjects, who did not know each other, having telephone conversations about a variety of randomly assigned topics. I received selections of conversations partially ToBI labeled by Jason Brenier and colleagues at Stanford. Each selection had been ToBI labeled except for the pitch accents, and I labeled the pitch accents for enough speech to collect seventy pitch accent tokens from each of the five MAE speakers and four SAE speakers. To select the MAE speakers, I chose people who were from the “North” or “West” region and who did not sound perceptibly nonstandard to me. For SAE speakers, I chose individuals who were identified as being from the South\(^2\). The characteristics of these speakers are listed in Figure 8.

---

\(^2\) I am not skilled at determining whether a speaker has a slight Southern accent or not, so I did not use my perception as a criterion.
V. Discussion of results

1. $L+H^*$

In the MAE ToBI labeling system, there is a three-way contrast between $H^*$, $L+H^*$, and $L^*+H$, also described as “rising” pitch accent types. The canonical shapes can be seen in Figure 9. This contrast is widely accepted though little experimental support for it exists (Arvaniti and Garding 2004).
In L+H* and L*+H, as opposed to the H*, there is said to be a rise from a low point. In L+H*, the L tone is aligned before the accented syllable, and in L*+H, it is on the accented syllable. Also, H* would show interpolation with another H*, but H* L+H* or H* L*+H would not. As mentioned above, Arvaniti and Garding (2004) determined experimentally that in Minnesota and Southern California English, L*+H conveys more emphasis than L+H*, both of which convey more emphasis than H*. My data for AE, MAE, and SAE show no evidence for an L*+H pitch accent type, but this is possibly because my data are conversational, and speakers are not performing extreme emphasis regularly. On the other hand, emphasis can be manipulated easily in experimental data. Therefore if the data for this project were designed to analyze emphatic pitch accents, it is possible that some L*+H would emerge. However, I only found strong evidence for H* and L+H* in my data.

I found that four of the eight AE speakers used L+H* pitch accents at a rate significantly greater than the MAE or SAE speakers (p ≤ 0.001 according to a chi-square test3), and none of the MAE or SAE speakers used a high proportion of L+H* (see Figures 10 and 11).

3 www.georgetown.edu/cball/webtools/web_chi.html
 Speakers Margaret, Thomas, Sharon, and Troy all use a high rate of L+H*. These speakers’ ages are 55, 53, 75, and 77, respectively. The other four AE speakers are 28, 41, 41, and 53 years old. Perhaps this is a feature that is more common among older speakers of AE, and is disappearing among younger people. Further research should
include more young speakers to check this hypothesis. Note that the differences in rising pitch accent form do not appear to correspond to gender, since of the four speakers who use high rates of L+H*, two are male and two are female.

Figure 12 shows examples of L+H* as realized by the speakers who use it often.

*Figure 12. Examples of AE speakers using high rates of L+H*
B. Thomas, L+H* use

C. Sharon, L+H* use
D. Troy, L+H* use

Note that this pattern of greater L+H* use can be phonetically implemented in different ways. In A., Margaret uses very sharp rises and falls. In C also, Sharon’s peaks are rather sharp, but in most of the other examples, the speakers use fairly gradual rises and falls on the L+H* tokens. It would be useful in further research to compare the pitch range of each speaker with the average pitch excursion of their rising tones, to find out if AE speakers who use more L+H* are also using more acoustically extreme versions of it, or perhaps if AE L+H* in general is more acoustically extreme than MAE or SAE L+H*. Or, AE L+H* could be phonetically similar to the L+H* in other varieties. My impression is that AE, MAE, and SAE L+H* do not differ acoustically, and since AE is one of the most divergent dialects of American English (Wolfram and Christian 1976), perhaps L+H* similar wherever it appears in American English. For comparison, see
below some examples of AE, MAE, and SAE speech without a large proportion of L+H* (Figure 13). Note that the speech of Margaret, even without L+H*, still exhibits more pitch movement on average than the other speakers below who do not have high levels of L+H* in their speech. This supports the hypothesis that the AE speakers who use more L+H* use greater pitch movement in general than other speakers.

*Figure 13. Speech with low level of L+H* use in AE, MAE, SAE*

A. Margaret (AE speaker), uncharacteristically low L+H* use
B. Elaine (AE speaker), characteristically low L+H* use

C. George (MAE speaker), characteristically low L+H* use
D. Ginger (SAE speaker), characteristically low L+H* use

Pierrehumbert (1980) and work based on her intonational model, including ToBI-based analyses, consider L+H* more emphatic than H*, and this is the impression I also had when labeling MAE and SAE data. However, this pragmatic function would seem unusual if applied to L+H* in AE, because half of the speakers use such high rates of it. We would have to assume that these speakers are being more emphatic than MAE or SAE speakers, and this is not the impression that I have from labeling these data. Such a theory could possibly be influenced by and reinforce stereotypes of Appalachians as animated and skillful orators. However, this is a possibility, since the AE respondents are being interviewed about their life by another person from the community, and in the Switchboard recordings, strangers are speaking to each other about given topics. I doubt this possibility, because I did not notice impressionistically any increased emphasis in the
AE speakers, but further research should ideally control for such differences, and use exactly the same type of conversational data from each variety under study.

In order to see whether L+H* is realized phonetically in AE and SAE as it is in MAE, I checked the average scaling of the AE L+H* peaks (normalized by dividing by the pitch range of the speaker, i.e. max f0-min f0) and found no patterned difference among the three groups (see Figure 14).

<table>
<thead>
<tr>
<th>Variety</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE</td>
<td>1.36</td>
<td>1.54</td>
</tr>
<tr>
<td>MAE</td>
<td>1.09</td>
<td>2.26</td>
</tr>
<tr>
<td>SAE</td>
<td>0.93</td>
<td>1.14</td>
</tr>
</tbody>
</table>

Figure 14. Avg. max f0 of L+H* / (max f0 - min f0)

This finding demonstrates that overall the L+H* pitch accent, speculation as to pragmatic function aside, is not realized with more or less extreme pitch excursion in AE than in MAE or SAE.

2. L*

In Figure 10 above and in Figure 15 below, it is clear that AE speakers overall use significantly more L* on average than MAE or SAE speakers (p≤0.001), though this is not true of all AE speakers individually.
This pattern is far weaker than that seen for L+H* though, because out of 70 total pitch accents per speaker, high L+H* use appears to range between 20 and 30 tokens, but high L* use appears to range between perhaps five and ten tokens. Therefore any pragmatic conclusions are more suggestive for L* than for L+H*, since L* has less impact on the basic declarative tune than L+H* in AE. Pierrehumbert and Hirshberg (1990) describe L* as used to indicate information that is intended to be salient but not to form part of the predication of the utterance, or in other words, information that is not to be added to the hearer’s mutual beliefs. They posit four contexts for a piece of information that might cause it to be highlighted but not predicated: 1) questioning (e.g. *Do you like ICE CREAM?*), 2) denying (e.g. *I don’t like MEAT*), 3) conveying something
that is already mutually believed (e.g. "This is a great MOVIE."), or 4) outside of the predication of the utterance (e.g. discourse particles such as well or like). Since AE speakers use L* more than MAE or SAE speakers, perhaps L* serves a different pragmatic function for AE speakers than for MAE or SAE speakers. To explore this point, I labeled each AE token for predication or non-predication, and I found that of 56 tokens across the 8 speakers, L* was used for non-predicated information 37 times (67%), predicated information 13 times (23%), and indeterminate\textsuperscript{4} information 6 times (11%). It seems that AE speakers are using in L* for both predicated and non-predicated information, but there is a tendency toward the standard pattern of using L* for non-predicated information. However, the data point to the conclusion that a L* has a wider set of pragmatic functions in AE than in MAE.

Overall, it appears that AE speakers are drawing on a different combination of intonational resources for pragmatic expression than are MAE and SAE speakers. On the other hand, SAE speakers appear to use pitch accents in the same proportion and with the same pragmatic function as MAE speakers. Perhaps this is one way in which listeners differentially evaluate AE and SAE, and SAE is afforded higher global prestige.

3. Implications for the simple declarative tune

Pierrehumbert and Hirschberg (1990) describe H* (H*) L-L% as the simple declarative tune in American English. My data discussed above suggest that this is indeed true for the MAE and SAE speakers but not for the AE speakers. Four of the AE speakers commonly use L+H* where others use H*, and several AE speakers use

\textsuperscript{4}These were cases in which I could not confidently determine whether the token was predicated or not in the discourse. Often this was the case if the stressed word is part of a phrase with a more important, content word that receives the high pitch accent of the predicated unit.
increased numbers of L* in that position. Therefore I suggest that the simple declarative
tune in AE may be more nuanced and variable than in MAE or SAE, and needs further
analysis.

Also, interspeaker variation appears to be greater in AE than in MAE or SAE,
since the latter groups seem to use H* so frequently, at the expense of other pitch accent
types. If this finding holds across a large number of speakers, it could mean that AE is
less regimented in terms of pitch accent use than other varieties of American English.
This would be a new of dimension of variation in intonation. It will be a sort of meta-
variable, because it means using a wider of variety of intonational variables overall,
rather than specific intonational variables like increased L+H* or L*. This is similar to
describing women’s speech as overall more variable than that of men, and therefore not
an unprecedented type of description in sociolinguistics more generally (Eckert and

4. The use of conversational speech data

Most researchers of intonation have used experimentally elicited speech as the
basis for their conclusions. This is understandable, since most of that work has examined
how English intonation changes under different contexts such as focus and emphasis.
However, I used spontaneous, conversational speech for my data. This was a theoretical
move, based on the common sociolinguistics finding that people use more “standard”
speech features in more formal situations like reading (cf. Labov 1972, Schilling-Estes
2002). If the goal is to describe variation among American English dialects, one should
look at the speech context where that variation should be greatest.
Spontaneous, conversational speech offers theoretical advantages, because it is more naturalistic than lab speech or read speech. It also offers practical advantages, because for researchers who want to describe variation in intonation, one would expect more variation in the more casual styles of speech. Therefore less speech data would have to be collected in a given fieldwork community to obtain a desired amount of tokens of nonstandard features.

For the most common dialect types, e.g. MAE, SAE, corpora also offer great sources of data for researchers. Those such as Switchboard (Godfrey et al. 1992) provide large amounts of conversational data, along with social information about the speakers such as age, region, and gender. The theoretical advantages discussed above and the relatively new availability of conversational data through corpora makes it an exciting line of research.

Despite these advantages, the fact that this is a new approach lends it certain difficulties. First of all, it is challenging to ToBI label conversational speech. Such speech is halting and full of disfluencies, and speakers often change direction in the middle of their train of thought. Also, there is no control for voicing of consonants, so the f0 track contains many gaps. Finally, I found that speakers of all three varieties use a different pattern of rising pitch accent alignment than in read speech. In conversational speech, people align a significant proportion of pitch accents early in the accented syllable, but in read speech, late alignment is far more common. This means that transcribers accustomed to labeling read or experimental speech could be confused by the different appearance of pitch accents in conversational speech, which look more like they are “falling” than “rising.” Also, because consonants raise f0 immediately after them but
lower f0 immediately before them (Gussenhoven 2002:7), early alignment makes high pitch accents easier to confuse with microprosodic effects of surrounding consonants, potentially confusing transcribers further (see Figure 16).

![Figure 16. Microprosodic effects of consonants on surrounding f0](image)

Greenberg et al. (1996), who also examined data from the Switchboard corpus, note the large number of differences between conversational and formal speech styles: “Detailed phonetic transcription of a spontaneous-speech corpus indicates that the spectral properties of many phonetic elements deviate significantly from their canonical form (4).”

I suggest that for those wishing to study variation in American English, conversational speech itself deserves greater recognition as an object of study in the field of phonetics, because it is the locus of maximal dialect differentiation. Also, researchers should not rely strictly on ToBI labeling conventions when examining conversational speech, because it consistently exhibits its own unique patterns. Specifically they should take into consideration alignment differences and the effects of microprosody. They should also make no assumptions about the pragmatic function of intonational features such as specific pitch accent types.
VI. Conclusion

In closely examining the intonational patterns of AE and comparing them with those of MAE and SAE, I determined that AE overall uses a different proportion of pitch accent types than the other two varieties. Half of the AE speakers in this study used L+H* and L* in significantly greater proportions than the other varieties. AE speakers overall also used significantly more L* than MAE and SAE speakers did. This suggests that those pitch accents serve a different pragmatic function in AE than in MAE and SAE. It therefore also suggests that AE speakers have a set of expectations about what type of pitch accents their fellow AE speaker is going to use that is different from the set of expectations held by MAE and SAE speakers. Researchers should not assume American English varieties all use pitch accents equivalently, and should be prepared to recognize and describe new pragmatic functions for pitch events in “nonstandard” dialects. I suspect that in AE, these two pitch accent types are more pragmatically similar to less "emphatic" ones (including H*) than they are in MAE and SAE.

These results also suggest that some types of variation captured by IViE but not ToBI, i.e. alignment and prominence, are not present in American English, and therefore ToBI (without its associated assumptions about pragmatic function) is appropriate for its analysis. A more phonetically based system like IViE is not necessary for describing these aspects of American English intonation. However, there could be peak shape differences across varieties, and IViE would be a useful tool to determine that. There is some evidence that AE speakers use greater pitch movement overall, and IViE would capture that difference, because it captures the interpolation or non-interpolation between high pitch points in its targets tier.
Also, because SAE uses intonation in the same way as MAE, and AE diverges from the two, it is possible that intonation is one dimension of language that allows interlocutors to distinguish between the two very similar varieties (in the form of an intensification of negative stereotypes).

I used conversational speech as the basis of this analysis and this brought theoretical advantages and practical difficulties along with it. However, it is necessary to examine conversational speech in order to understand the scope of differences between dialects of American English (Labov 1972, Schilling-Estes 2002), and so any method for researching intonation, like ToBI, should include explicit treatment of the differences between spontaneous and non-spontaneous speech.

Further research into variation in American English intonation could look more closely at the pragmatic functions of different pitch accents, phrase accents, boundary tones, and overall tunes in AE and other dialects. Specifically, there seems to be significant variation in the use of rising pitch accents in American English, as found here and in Arviniti and Garding (2004). The basic declarative tune may also vary across dialects, and should be analyzed critically in each variety, and perhaps even in each speaker. A production experiment that determines how categorically people realize different levels of emphasis on a given pitch accent could be used to determine whether AE actually has a distinction between L+H*, H*, and L*.

I recommend future research control avoid comparing speech from different types of conversational contexts, because it is possible that there are intonational differences between speech on different kinds of topics. I also suggest further research use as many
subjects as possible, in order to explore the issue of age that arose here, and any potential effects of gender or social class, etc.
References


Beckman, Mary and Ayers G.M. 1994. Guidelines for ToBI transcription, version 2.0, Ohio State University.


