Gender and ethnicity in intonation:

A case study of North Carolina English

Jennifer Cole[1], Erik R. Thomas[2], Erica Britt[1], and Elizabeth Coggshall[3]

[1]University of Illinois at Urbana-Champaign

[2] North Carolina State University

[3]New York University

Corresponding Author:
Jennifer Cole
Dept. of Linguistics, 4080 Foreign Languages Bldg.
University of Illinois
707 South Mathews
Urbana, IL 61801 USA

Email: jscole@uiuc.edu
Phone: 217-244-3057
Fax: 217-244-8430

[7,891 words, including figure captions]

Gender and ethnicity in intonation: A case study of North Carolina English

1. INTRODUCTION

Sociolinguistic factors condition a wide range of phonological and phonetic variation, and recent studies demonstrate that these effects extend to intonation. In English, intonation varies regionally in the British Isles, Australia and New Zealand (Fletcher et al. 2005; Grabe 2004), in the United States (Arvaniti and Garding 2007), and in Nigeria (Gut and Milde 2002). Gender effects on intonation are also reported for New Zealand English (Daly and Warren 2001) and American English (McConnell-Ginet 1978). Though studies of regional and gender factors in intonation variation are scarce, even less is known about ethnic variation in intonation. This paper examines intonation variation at the intersection of gender and ethnic identification among speakers of American English through a comparison of male and female speakers of African American English (AAE) and European American English varieties in North Carolina. Evidence from analyses of F_0 patterns that encode intonational features of an utterance, including pitch accents and phrase boundary tones, reveals significant differences between males and females for both African Americans and European Americans. Effects of ethnicity on intonation patterns are also apparent, but differ for males and females. Among males, the differences between African Americans and European Americans emerge in the shape of the F_0 contour, while among females, European Americans display an overall elevation in F_0 range relative to African Americans.

Previous research on African American intonation has been limited, albeit more extensive than intonational work on other North American dialects of English. Past researchers have listed numerous features that characterize African American intonation. However, these features fall into three categories: pitch range, boundary tones, and phrase-internal tones (i.e., pitch accents in current terminology). Several sources (Tarone 1973; Loman 1975; Hudson and Holbrook 1981, 1982; and Jun and Foreman, 1996) have reported that African Americans use a wider pitch range than corresponding European Americans. Their data come from widely scattered places—Seattle; Washington, D.C.; Florida; and Los Angeles, respectively—and thus cannot be discounted. Tarone (1973) and Loman (1975), in fact, report falsetto to occur in AAE. However, they each suggest that the wide pitch range characterizes only certain speech styles in AAE. Tarone (1973:32) associates the wide pitch ranges with "competitive...speech events" and Loman (1975:242) describes the highest pitch levels as "typically occurring in the excited speech of Negro men." Moreover, Hudson and Holbrook (1981, 1982) report that their African American subjects used a narrower pitch range in read speech than in spontaneous speech; they noted that previous studies had found the opposite for European Americans. Goodwin, Goodwin, and Yaeger-Dror (2002), nonetheless, find that African American girls use fewer instances of extremely high F_0 than Hispanic girls when playing hopscotch.

Most of the discussion regarding AAE boundary tones (or "final contours" in some earlier studies, e.g., Tarone 1973, that used older transcription systems) has concerned those in *yes/no* questions. As Jun and Foreman (1996) note, the ordinary IP

boundary tones in European American English are L-L% in declaratives and H-H% in *yes/no* questions. However, they find that African Americans commonly produce a variety of boundary tones—L-L%, H-L%, and H-H%—in *yes/no* questions. Green (2002), with data from Louisiana, likewise finds that African Americans frequently show level or falling boundary tones in *yes/no* questions. Earlier, Tarone (1973) had reported that African Americans showed falling final *yes/no* contours in formal situations but level or rising ones in informal situations. Tarone had also noted that African Americans were more prone to producing level and rising final contours in other types of utterances.

Phrase-internal (non-boundary) tones have proved harder to describe. Other than noting that conditional clauses could exhibit special contours in African American English when the word *if* was deleted, Tarone (1973) does not address them. Loman (1975), nevertheless, reaches two conclusions. He states that African American English shows a "high frequency of primary stresses" and that it exhibits a "constant and marked shift between pitch levels /3/ and /2/ [on a scale of 1 to 4] which basically is correlated with the shift between syllables with primary stress and with weaker stress..." (Loman 1975:242). Two other studies report other, apparently related, findings. Jun and Foreman (1996) state that "AAE tends to have a post-nuclear pitch accent and/or phrase boundary, especially when [the] focus is in sentence-initial position." Post-nuclear pitch accents ought to correspond to the greater number of primary stresses that Loman found. Wolfram and Thomas (2002), using data from Hyde County, North Carolina, examine whether successive stressed syllables would rise or fall in pitch. They find that African Americans showed a statistically higher frequency of rises than European Americans,

which is presumably related to the "constant and marked shift" of pitch levels that Loman notes. Wolfram and Thomas counted successive syllables regardless of intonational phrase or conversational turn boundaries in order to avoid subjectivity, even though this practice likely inflated the incidence of rises in European American speech.

Taken together, these previous findings indicate that the intonation features used by African American English speakers differ from those of European American speakers at both the phonetic and phonological levels. Phonologically, there are differences in the tone features that mark the IP boundary in *yes/no* questions. Another phonological difference is in the use of pitch accents or "primary" stress as marks of phonological prominence. Phonetically, there are differences in the implementation of tone features, observed primarily in differences in pitch register or in the scaling of pitch.

While the findings reviewed here point to dialectal variation in prosody, there are limitations in what can be concluded from them. One problem is that, with the exception of Jun and Forman (1996), the analyses are based on impressionistic data, without quantitative, objective measures of intonation. The studies differ in the speech materials they use. Tarone (1973) and Goodwin et al. (2002) used conversational speech among friends; Loman (1975) used spontaneous conversation by children with friends, adult relatives, and researchers; Wolfram and Thomas (2002) used conversational interview speech; Green (2002) used conversational speech but did not name the settings; Jun and Foreman (1996) used read dialogues designed to mimic vernacular conversations; and Hudson and Holbrook (1981, 1982) used both spontaneous and read speech. Findings

from read sentences may not be representative of the intonation patterns in speech addressed to an unfamiliar interviewer, which in turn may differ from conversational speech between familiars. Finally, except for Jun and Foreman (1996), the prior studies are not based on any explicit phonological model of prosody. The absence of such a model makes for uncertain comparisons with work on other dialects or languages. Furthermore, researchers operating under different assumptions about the basic units of prosody or operating with no clear definitions of those units may offer different descriptions for what are in fact similar intonation patterns, and at the same time, descriptions that appear to be similar may depict intonation patterns that differ phonologically and/or phonetically.

2. METHODS FOR ANALYZING INTONATION VARIATION THROUGH CORPUS ANALYSIS

In this study we investigate intonational differences between African American and European American English in a corpus study using both phonological and phonetic analysis. Capitalizing on developments in the theory of intonational phonology (e.g., Beckman and Pierrehumbert 1986; Ladd 1996), and in explicit models of the intonation of American English (Beckman and Ayers 1997), we look for differences in the phonological encoding of intonation and in the phonetic realization of tone features. The analysis is based on data from intonational transcriptions and measures of the acoustic correlates of the transcribed intonation features. The acoustic measures are the primary basis for comparison among speakers who differ in ethnicity and/or gender in this study,

and they establish a phonetic description of the intonation patterns for pursuing future phonological analysis. Our findings shed light on the nature of prosodic variation across dialects and invite immediate comparison with published analyses of other languages and dialects using the phonological framework of the ToBI system.

2.1 The corpus

Speech data are drawn from the North Carolina Language and Life Project (NCLLP) corpus (see, e.g., Wolfram and Thomas 2002). These corpus materials are ideally suited for the study of prosody in non-standard dialects because the speech is collected from unscripted, conversational interactions. The NCLLP interviews used here are of African Americans and European American English from Hyde, Robeson and Warren Counties of North Carolina, conducted during the period from 1994 to 1999. Hyde County is located along the coast of North Carolina, Robeson County along the South Carolina border, and Warren County along the Virginia border. Interviewees were accessed through personal contacts made within the communities. All interviews were conversational and most involve multiple participants (e.g., a group of three friends, or a mother and daughter). Interviews are around 45 minutes to an hour in length, and short excerpts of one to three minutes (with minimal pause and disfluency content) were extracted from each interview for this study. Topics of conversation ranged widely, but personal experiences and opinions about the community were emphasized. Interviews were conducted in the speakers' home communities, usually in their homes. Recordings were made with Marantz tape recorders, model PMD221. A variety of high-quality table-top microphones were used.

2.2 Intonational transcription

The speech materials were transcribed by a team of four trained transcribers using a simplified ToBI transcription., which adopts the criteria for pitch accent and boundary labeling from the original MAE_ToBI system, but with a restricted inventory of intonation labels. Our simplified transcription marks the location of pitch accented syllables and prosodic phrase boundaries with the labels "A" (accent) and "B" (boundary), using the ToBI labeling criteria, which draws on the transcriber's auditory impression in concert with evidence from visual inspection of the waveform, F_0 track and spectrogram display. The simplified transcription does not specify the tone features that distinguish among pitch accented syllables and boundaries; nor does it label the difference in boundary strength (intermediate phrase vs. intonational phrase). The simplified transcription is justified on the two grounds. First, because intonation analyses for other languages (e.g., Grice et al. 2005 for German) and for other varieties of English (e.g., Arvaniti and Garding 2007) vary in their inventories of pitch accents and phrase tones and in the number of phrase level distinctions recognized, we could not assume that the two North Carolina varieties of English considered here would exhibit the same inventory of pitch accents and phrase tones as in the ToBI system for Mainstream American English. Second, we needed an efficient and reliable transcription process. Prosodic transcription of spontaneous speech is notoriously challenging because of hesitations, disfluencies, cross-talk, non-speech sounds such as laughter and coughing and the greater incidence of reduced forms. By adopting a simplified ToBI transcription, we could obtain sufficient transcribed speech for eight speakers, with at least two independent

transcriptions for every sound file. The resulting transcriptions are admittedly coarse-grained, but at the same time they offer a first look into the intonational system by identifying the locations of pitch accented syllables and prosodic phrase boundaries and their acoustic correlates in F_0 contours. These transcriptions may also guide future investigations, leading to a more detailed phonological analysis.

A team of three transcribers labeled speech samples from eight speakers in the NCLLP corpus. The lead transcriber (author E.B.) is a linguist with phonetics experience, trained in the full ToBI system for Mainstream American English, and can command AAE (though not the North Carolina variety) as a stylistic option. The other two transcribers are undergraduate linguistics students who were trained only in the simplified ToBI system and who were supervised by the lead transcriber. They are not speakers of AAE. Transcribers worked independently but met regularly to resolve transcription questions and to arrive at a consensus transcription for such cases. A subset of the speech materials (360 words) was independently transcribed by a fourth transcriber who was trained in the full ToBI system and who is also a linguist with phonetics experience. Inter-transcriber agreement was calculated for each of the three primary transcribers paired with the fourth transcriber using Cohen's Kappa statistic (Cohen 1960). The mean agreement rate for pitch accented syllables is 82.3%, with a mean Kappa statistic of 0.65; for boundary labels the mean agreement rate is 92%, with a mean Kappa statistic of 0.84. Kappa statistics above 0.6 are accepted as indicating strong agreement.[1] After completing their independent transcriptions, the transcriber team met to

[1] Our agreement rates are comparable to or slightly better than agreement rates for simplified ToBI transcriptions of conversational speech from the Switchboard corpus of American English (Yoon et al. 2004). Our agreement rates for boundary are similar to agreement rates on the presence vs. absence of boundary tones in a full ToBI transcription of two professional speakers (radio news announcers) of

discuss problems and questions and arrived at a consensus transcription for a subset of the speech materials.

2.3 Hypotheses

The study reported here is informed by our earlier exploratory study of the NCLLP corpus. The intonational component of the distinctive prosody of the African American speakers impressionistically seems to consist of a prominent rising-falling F_0 contour on words with phrase-level or sentential stress (i.e., pitch accented words), which results in a dynamic pattern of F_0 modulation across the prosodic phrase. These rising-falling F_0 contours that mark pitch accented syllables also seem to enhance the rhythmic quality of the African American speech, perhaps through alignment of the F_0 contour with rhythmic structure (though rhythm analysis is not carried out in the present study). The European American speakers, in comparison, produced F_0 contours on pitch accented syllables that seemed flatter and/or broader in shape. Our preliminary impressions also suggested differences between African American and European American speakers in the F_0 patterns across successive pitch accented syllables within the prosodic phrase. The African American speakers seem to exhibit a shallower declination over the phrase, where declination was assessed visually by comparing the height of successive peaks or valleys in the F_0 track. Shallow declination over the prosodic phrase is also typical of male European American speakers, but with the striking difference that the European American males produce much smaller F_0 excursions overall. The European American females typically exhibited more F_0 modulation than their male

---

American English (Syrdal & McGory 2000), but not as high as agreement for presence vs. absence of pitch accent in that study.

counterparts, but with downward drift (or sometimes a full downstep) across successive high peaks, and a similar declining slope across F_0 valleys as well. Finally, we also observed ethnicity-based distinctions in F_0 range: African American male speakers use a larger F_0 range with more dynamic F_0 movement than European American males, as has been reported in earlier studies, and European American females use an elevated F_0 range relative to African American females.

Although these intonational features were not observed in every utterance, they were perceived to be characteristic of the salient distinctions between African American and European American speakers from this corpus. Figure 1 displays typical examples of the distinctive F_0 contours of four speakers, with one male and female from each of the African American and European American groups. Note the compressed F_0 range and lesser F_0 modulation for the European American male speaker, the elevated F_0 range of the European American female speaker, and the rise-fall F_0 contours associated with pitch accent for both the male and female African American speakers.

[FIGURE 1 ABOUT HERE]

In the present study we ask whether the distinctive properties of African American intonation qualitatively described in our exploratory study generalize to a larger set of utterances. We identify several acoustic correlates of the African American intonation contours previously observed and then test if these correlates reliably distinguish African American from European American speakers in a larger set of utterances. This is a quantitative analysis that measures every utterance in the corpus for

the acoustic correlates of steep phrase-final falling F_0 contours, of F_0 range, and for pronounced falling F_0 contour marking phrasal stress. Three experimental hypotheses are formulated below.

Hypothesis 1—Rising-Falling F_0 contours for pitch accent: African American speakers exhibit pitch accented syllables with more pronounced rising-falling F_0 contours than European American speakers, as measured by (i) higher F_0 maximum values, (ii) lower F_0 minimum values, and (iii) a larger fall in F_0 from the start of the accented syllable to the end of the word.

We consider the F_0 movement over entire word, rather than looking only at the pitch accented syllable, because in some instances a rise-fall F_0 pattern extends beyond the stressed syllable. While there are also cases where the falling F_0 pattern extends beyond the accented word,, in many other instances there is an abrupt fall from the F_0 peak of the pitch accent that is complete at or before the end of the word. If pitch accents with prominent F_0 rising-falling contours in fact characterize African Americans in our corpus, in contrast to flatter and/or broader pitch accent contours for European Americans, then measurement of the F_0 contour in the interval from the accented syllable to the word edge should produce the distinctive F_0 patterns posited in Hypothesis 1.

Hypothesis 2— F_0 **Range:** African Americans use a wider F_0 range than European Americans.

Hypothesis 3— Phrasal F_0 contours: African Americans show a pattern of dynamic F_0 rises and falls against a less steep declination across the prosodic phrase compared to European American speakers. That is, phrases produced by African Americans exhibit a shallower slope across successive F_0 peaks, or across successive valleys, than phrases produced by European Americans.

3. METHODS

Three trained transcribers labeled about fifteen minutes total of continuous speech excerpted from the recorded dialogues of eight speakers: two male and two female African Americans and two male and two female European Americans. We acknowledge here that the group size is small, with just two speakers in each ethnicity-gender group (e.g., two European American males, etc.). The limited scale of this study reflects the labor-intensive nature of intonation transcription, coupled with reliability testing, which necessitates multiple transcribers working on the same data. Results obtained from this study must be considered preliminary and require confirmation with additional speakers, though results that confirm findings from earlier studies may be considered reliable.

Transcribers assigned the label "A" to any word that was perceived as bearing a pitch accent, using the ToBI criteria for the assignment of pitch accent. In the ToBI system, pitch accent marks words that bear phrasal stress, and which also often (or typically) are focused. The "A" label was used for both nuclear (rightmost) and pre-nuclear pitch accented words. The A label was aligned at the peak of the F_0 contour in the accented

syllable, or in the center of the syllable if there was a sustained F_0 plateau across the syllable. The pitch accent domain was also labeled, beginning with the onset of the vowel in the accented syllable (always identified with the primary stressed syllable of the word) and ending with the end of the accented word. Transcribers marked prosodic phrase boundaries ("B"), again based on the perceptual criteria of the ToBI transcription standard (Beckman and Ayers 1997), which includes phrase-final lengthening, pause, and F_0 contours (falling, rising or plateau) that do not follow the declination slope, based on visual inspection of the F_0 track. Table 1 reports the total word count for each speaker and the number of words transcribed with pitch accent marking phrasal stress (labeled "Accent") and the number of words in phrase-final position (labeled "Boundary").

Speaker	Speaker	# Words	# Accents	# Boundaries
African American Female	AAF1	481	208	105
	AAF2	297	133	49
	Total AAF	778	341	154
African American Male	AAM1	267	98	70
	AAM2	412	156	92
	Total AAM	679	254	162
European American Female	EAF1	310	128	75
	EAF2	234	91	63
	Total EAF	544	219	138
European American Male	EAM1	371	135	74
	EAM2	379	178	100
	Total EAM	750	313	174
TOTAL (all speakers)		2751	1127	628

Table 1. Frequency of accents and boundaries and the total word count for each speaker's transcribed data. Speakers are identified as African American female or male (AAF, AAM) or European American female or male (EAF, EAM).

F_0 values were extracted from each pitch accent domain and filtered to remove F_0 doubling and halving due to autocorrelation error. The filtered data were then transformed into units on the ERB (equivalent-rectangular-bandwidth-rate) scale using the formula in (1), from Hermes and van Gestel (1991).[2]

(1) $\text{ERB} = 16.7 \log_{10} (1+f/165.4)$, where f is frequency in Hertz

Several F_0 measures were taken from the filtered F_0 data. F_0 measures were analyzed separately for phrase-medial and phrase-final pitch accents because the F_0

[2] The ERB scale is based on the human auditory frequency response. Critical bands are narrower than for the Bark scale, especially at lower frequencies, and while the Bark scale is linear for frequencies below 400 Hz, the ERB-rate scale is intermediate between linear and logarithmic. Psychoacoustic tests conducted by Hermes and van Gestel (1991) suggest that ERB is more appropriate than Bark or semitone scales for the frequencies that F_0 values cover.

contours of phrase-final pitch accents are often influenced by tone features marking the end of the prosodic phrase. Measures used to test Hypothesis 1 assess the scaling and shape of the pitch accent F_0 contour and its location on the frequency scale and include:

- Mean, Maximum, and Minimum F_0 in the pitch accent domain
- Start F_0, taken at the onset of the accented vowel
- End F_0, taken at the right edge of the sonorant portion of the final syllable rhyme (i.e., at the end of the final sonorant rhyme consonant if there is one, and otherwise at the end of the vowel)
- F_0 Range measured over the pitch accent domain (Max minus Min)
- F_0 Fall: the drop in F_0 over the pitch accent domain.

The Max and Min values are also used to test Hypothesis 2 (F_0 range). To test Hypothesis 3 (phrasal F_0 contours), we used linear regression analysis and fit F_0 values in the interval of the prosodic phrase to a regression line, comparing measurements of the slope of the regression line for African American and European American female speakers. In addition to these measures, we also report the frequency of pitch accent and boundary as a function of word count.

Statistical analyses were conducted using non-parametric tests to look for differences in F_0 measures between speakers who belong to different ethnicity or gender groups. As is typical for analyses of corpus data, where the experimenter cannot control for sources of acoustic variation, the F_0 measures within comparison groups do not display a normal distribution and the comparison groups (measures grouped according the ethnicity and gender of the speaker) are not of equal size or homogenous variance. Thus, the assumptions of ANOVA are not met for these data. To test the null hypothesis

for each F_0 measure, namely that there are no differences between values from the four ethnicity-gender groups (i.e., they come from a single distribution), the non-parametric Kruskal-Wallis test is used and the Kolmogorov-Smirnov test is used for comparison between two groups (e.g., African American males vs. European American males, or all females vs. all males). Kruskal-Wallis looks for differences in the distribution of three or more groups by comparing the population medians among groups, and the Kolmogorov-Smirnov test seeks differences in both the locations and shapes of the distributions for each group. Using these tests, each F_0 measure is tested individually under three analyses. The first analysis makes comparisons across four ethnicity-gender groups—European American male, European American female, African American male, and African American female—to find any intonational distinctions that index speakers' identities as they relate to ethnicity and gender. The second analysis looks separately at the effects of ethnicity and gender on F_0 measures over the entire group of speakers, at ethnicity effects within each gender group, and at gender effects within each ethnicity group.

4. RESULTS

4.1 Frequency of pitch accent and prosodic phrase boundaries

The frequency of pitch accented words (words transcribed with an "A" label), calculated as the number of words per accent, varied somewhat by speaker (Figure 2), but there are no consistent differences between African Americans and European Americans or between males and females. All speakers have a mean frequency of between two and

three words per accent. The frequency of transcribed prosodic phrase boundaries varied across speakers even more than pitch accent (Figure 2), but no patterns of difference according to the ethnicity or gender of the speaker emerged. Speakers produced prosodic phrase boundaries roughly every four to six words, which is thus the range for mean length of prosodic phrase for these speakers.

[FIGURE 2 ABOUT HERE]

4.2 ERB measures of F_0 contours of pitch accents

The F_0 contour for pitch accents is very similar in shape for each of the four ethnicity-gender groups. Figure 3 shows plots that model the F_0 contour space of phrase-medial and phrase-final pitch accents for each group using error bar plots of the 95% confidence intervals for the F_0 measures taken within the pitch accent domain: Start, Max and Min. The plots position the Max F_0 value between the Start and Min values, as that is the contour pattern observed in the majority of pitch accents (rising-falling) for all speakers. The F_0 contours in medial and final position in the prosodic phrase are similar, but the phrase-final contours have lower mean Min F_0, most likely influenced by the prevalence of a low F_0 target at the end of the phrase (i.e., L- or L-L% in the ToBI system).

[FIGURE 3 ABOUT HERE]

Comparison across four speaker groups. Results from Kruskal-Wallis tests show significant differences among the four speaker groups for all F_0 measures for both phrase-medial and phrase-final pitch accents. Trends as they relate to the ethnicity and gender factors are detailed in the results for ethnicity and gender effects below. For phrase-medial pitch accents, Start [χ^2=374.7, p<.001], Max [χ^2=361.2, p<.001], Min [χ^2=348.0 p<.001], Range [χ^2=53.4, p<.001], and Fall [χ^2=19.6, p<.001]. For phrase-final pitch accents, Start [χ^2=214.6, p<.001], Max [χ^2=210.0, p<.001], Min [χ^2=205.8, p<.001], Range [χ^2=20.8, p<.001], and Fall [χ^2=21.6, p<.001].

Effects due to ethnicity. When male and female speakers are pooled, Kolmogorov-Smirnoff tests show significant differences between African Americans and European Americans for every F_0 measure in phrase-medial position and for every measure except Max F_0 in phrase-final position. Statistical analyses within gender groups find significant effects of ethnicity for both males and females. In phrase-final position the pattern is fairly consistent: African American males and females have lower Min values than European Americans of the same gender, and these low F_0 values also contribute to a higher mean Range and F_0 Fall values for African Americans compared to European Americans (though the distinction in F_0 fall only approaches significance for the male speakers). For Max F_0 , African American males show non-significant tendency towards higher peaks than European American males, but the opposite holds for females; European American females have significantly higher F_0 peaks than other speaker groups, as is evident in Figure 3. In phrase-medial position, the ethnicity effect goes in the opposite direction for males and females for many F_0 measures. For example, European

American males have lower F_0 values for phrase-medial pitch accents than African American males for every measure except Min F_0, while for females it is the African Americans who have lower F_0. Results from significance testing with data pooled for both gender groups and from testing within each gender group are shown in Table 2.

	Pooled (M&F)		Males		Females	
F0 (ERB)	phrase-med	phrase-fin	phrase-med	phrase-fin	phrase-med	phrase-fin
Start	*** z =2.46	** z =1.89	*n.s.*	*n.s.*	*** z =3.95	*** z = 3.47
Max	*** z =2.66	** z =1.59	*n.s.*	*n.s.*	*** z =3.81	*** z = 3.08
Min	*** z =2.67	** z =2.55	* z =1.43	* z = 1.49	*** z =5.46	*** z = 4.36
Range	*** z=2.79	** z=2.31	** z = 2.01	** z = 1.75	* z =1.55	* z = 1.42
Fall	*** z=2.09	** z=2.31	*n.s.*	*n.s.* (p = .058)	* z =1.56	** z = 1.83

Table 2. Results of non-parametric tests (Kolmogorov-Smirnov) showing significant effects of ethnicity on pooled male and female data and effects of ethnicity within each gender group. Separate results are shown for F_0 data from phrase-medial and phrase-final pitch accents. Significance levels are $p < .001$ (***), $p < .01$ (**), and $p < .05$ (*).

Plots of ethnicity effects on mean Start, Max and Min F_0 values are shown in Figure 4, and F_0 Range data are plotted (95% CI) in Figure 5.

[FIGURE 4 ABOUT HERE]

[FIGURE 5 ABOUT HERE]

Effects due to gender. For European American and African American speakers combined, Kolgomorov-Smirnov tests show significant effects of gender for every F_0 measure in phrase-medial and phrase-final position. As would be expected, females exhibit higher F_0 values for every measure than males. The gender effect is notably larger for European Americans than it is for African Americans for every measure, reflecting both the highly elevated F_0 of the European American females and the relatively low F_0 of European American males. Gender does not affect all measures uniformly, and the biggest gender difference occurs for the Max F_0 measure—again reflecting the extremely high F_0 peaks prevalent among European American females, in contrast to the low peaks and compressed F_0 range of their male counterparts. These effects of gender can be seen in the F_0 mean plots in Figure 3. Results from significance testing with data pooled from both ethnicity groups, and from testing within each ethnicity group are shown in Table 3.

	Pooled (EA&AA)		European American		African American	
F0 (ERB)	phrase-med	phrase-fin	phrase-med	phrase-fin	phrase-med	phrase-fin
Start	*** z =7.09	** z =7.09	*** z =7.29	*** z = 6.03	*** z =6.92	*** z = 5.05
Max	*** z =6.65	** z =6.65	*** z = 7.29	*** z = 6.14	*** z = 6.01	*** z = 4.56
Min	*** z =7.22	*** z =7.22	*** z =7.06	*** z = 5.97	*** z = 7.68	*** z = 5.78
Range	*** z=1.91	*** z=1.91	*** z = 2.78	* z = 1.41	** z = 1.75	*n.s.*
Fall	*** z=1.51	* z=1.51	***. z = 1.82	*n.s.*	** z = 1.69	*n.s.* (p = .056)

Table 3. Results of non-parametric tests (Kolmogorov-Smirnov) showing significant effects of gender on data pooled over European Americans and African Americans and effects of gender within each ethnicity group. Separate results are shown for F_0 data from

phrase-medial and phrase-final pitch accents. Significance levels are $p < .001$ (***), $p < .01$ (**), and $p < .05$ (*).

4.3. Regression modeling of F_0 phrasal contours

In order to provide some confirmation of the observed ethnic differentiation in F_0 fall patterns, we examined declination for a larger sample of young female speakers. For this purpose, we analyzed conversational interviews of six young African American females and six young European American females from the communities noted earlier. The speech excerpts analyzed in this part of our study were not the same set of materials that were transcribed with Accent and Boundary labels as described above, though they are from the same corpus and include speech from one of the female speakers analyzed here (though not exactly the same utterances from her speech). To investigate the fall patterns in phrasal F0 patterns, we marked the location of Intonational Phrase junctures based on examination of spectrograms with F_0 tracks and impressionistic listening. Note that lower-level prosodic phrase junctures (e.g., intermediate phrases) are not marked in the transcription for the data set used in the F_0 fall analysis, as they were for the data set described above and used for analysis of the F_0 contour on the accented word. Questions and imperatives—identified on the basis of syntactic structure—were excluded, leaving only declaratives. We also excluded any utterances that had declarative syntax but, based on context, could have been interpreted as questions, but we note in passing that questions of any sort are not frequent in the speech of the interviewee in this corpus. Then, using the F_0 tracking function in Praat (Boersma and Weenink 2007), we

determined the trajectory of F_0 through the course of each IP by means of linear regression. Ordinarily, of course, F_0 fell during an IP, and the rate of fall was measured in Hz/second. For each speaker, IPs were categorized according to their duration. That is, IPs of 0.5 second or less were placed in one category, those of 0.51-1.0 sec in another category, those of 1.01-1.5 sec in a third category, and so on. Within each durational category, mean values of the F_0 slopes were calculated for each speaker. These values were then plotted against their IP durations, as shown in Figure 6.

[FIGURE 6 ABOUT HERE]

F_0 slopes for the shortest category of IPs shown in Figure 6 vary widely, which is unsurprising. Such short IPs often consist of truncated phrases in which there is time only for a single rise or fall in F_0. The longest IPs—those with a duration over 3.0 sec—show no obvious ethnic differentiation. In long IPs, declination is drawn out to the point that the slope inevitably approaches zero. The ethnic differentiation appears for IPs of intermediate duration, especially in the range of 0.51-1.5 sec. F_0 slopes for all the African Americans lie close to zero, indicating that F_0 is relatively flat or that it shows peaks of approximately equal magnitude throughout the IP. European Americans show considerable variation, however. Some of them show fairly flat F_0 slopes similar to those of the African Americans. Others, however, show strongly negative slopes, reflecting the presence of an especially high, prominent peak near the beginning of the IP.

5. DISCUSSION

The measures of F_0 contours marking pitch accents show both similarities and differences in the intonation patterns of European American and African American speakers. Speakers from all four ethnicity-gender groups show pitch accent F_0 contours with a rising-falling pattern over the portion of the word from the accented syllable to the minimum F_0 value at or near the word end. The similarity in F_0 contours among the speakers in this study and their similarity to the rising-falling F_0 movements on accented syllables as described for Mainstream American English (e.g., Beckman and Ayers 1997) suggests a common inventory of phonological pitch accents with a High tone: H*, L+H* and L*+H. In other words, in our data as in other descriptions of American English intonation, we find that syllables perceived as pitch accented are marked with an F_0 peak in the vicinity of the accented syllable (specifically, towards the end or even immediately following), that is followed by a fall on the following syllable. Our study does not consider how the turning points in the F_0 contour are aligned with respect to landmarks of the accented syllable. Thus, it does not indicate whether all three of these pitch accent types are available for each speaker group or whether groups differ in the frequency or distribution of these pitch accent types. The findings from our study establish the basic rise-fall shape of the contour, which extends over the pitch accented syllable and in some cases into a following syllable up to the word boundary, and hence lay an empirical foundation for subsequent research on phonological pitch accent types in each variety.

It is noteworthy that the expanded and elevated F_0 range of the European American females comes with a higher variance as well, evident from the Start, Max and Min F_0 distributions plotted in Figure 3. This pattern may indicate that female speakers

operate in or near a falsetto F_0 range and with less precise control over F_0 modulation. The finding of greater F_0 range for African American males relative to European American males partially confirms Hypothesis 2, but given the absence of an ethnicity-based F_0 range distinction for female speakers, we cannot conclude that an expanded F_0 range is a general feature of AAE.

Besides the differences in F_0 range that distinguish the European American speakers, there is a subtler distinction. The African American males exhibit lower Min F_0 values and larger F_0 Fall values compared to their European American male counterparts in both phrase-medial and phrase-final position and exhibit higher Max F_0 values in phrase-final position. These findings confirm Hypothesis 1 and corroborate empirically our impression of a more dynamic F_0 pitch accent contour for male African American speakers from this community.

The more prominent fall in African American males' pitch accent contours suggests a phonological distinction. The low F_0 value at the end of the pitch accent domain (i.e., taken here to be the end of the accented word) is plausibly the expression of a Low tone feature, e.g., as the second tone in a bitonal pitch accent: H*+L. This pitch accent type is not included in the inventory for Mainstream American English,[3] so an analysis that invokes a H*+L pitch accent posits a difference in the phonological pitch accent inventories for African American and European American male speakers. An alternative account, which does not invoke a H*+L pitch accent, is to attribute the

[3] Malcah Yaeger-Dror (p.c.) reports that a falling pitch accent, H*+L, does occur in mainstream varieties of American English in the production of disagreement utterances, based on her work with the CallHome corpus. If pitch accents with falling F_0 are indeed part of the inventory for Mainstream American English, then the distinction we observe between African American and European American speakers from these North Carolina communities may be a distinction in the pragmatic conditions under which the falling pitch accent is used, rather than a distinction in the phonological inventory of pitch accents *per se.* We thank Yaeger-Dror for illuminating discussion on this point.

prominent F_0 falls of African American males to the presence of a Low tone marking the end of a prosodic phrase. However, this low falling contour typifies pitch accents that occur medial to the prosodic phrases as labeled in the intonation transcriptions, so this alternative analysis amounts to a claim that the pitch accent domain is in fact a minor phonological phrase—perhaps a case of a true level 2 prosodic phrase in a ToBI analysis. An analysis based on the minor phonological phrase as the domain for pitch accent also posits a phonological distinction between African American and European American males, but locates the distinction in the prosodic phrase structure rather than the pitch accent inventory. Our findings do not resolve these competing hypotheses, but future research on timing of the turning points in rise-fall contours should shed light on this matter.

Female speakers also show differences in their pitch accent F_0 contours, the most obvious being the upward-shifted F_0 register of European American females. However, beyond this difference, the pitch accents produced by African American females in phrase-final position show the same pattern of a lower F_0 minimum and greater F_0 fall relative to their European American female counterparts that we observed for the African American males. This suggests that the female speakers may share the phonological distinction of a H*+L pitch accent, or alternatively of a Low tone marking a minor phrase boundary, as the African American males, in which case this phonological feature may be an ethnicity-based dialect feature of AAE in North Carolina.

At the level of the phrasal F_0 contour, the results from regression modeling of data from female speakers confirm Hypothesis 3, that African Americans produce phrasal

contours with more dynamic F_0 contours for pitch accents against a less steep declination across the phrase than European Americans.[4]

The results of F_0 analysis provide evidence for a gender-based distinction as well for both ethnicity groups. As expected, males have overall lower F_0 than females for all F_0 measures. In fact, the difference is manifested in Hz as well as in the ERB measurements reported here. Moreover, as already noted, females have a larger dynamic F_0 range than males (again, whether measured in in Hz or in ERB), which matches similar findings of gender-based distinctions, based on ERB analysis, in New Zealand English (Daly and Warren 2001), and in Dutch (Haan and van Heuven 1999). In addition, the larger fall excursions for females in our study corroborate similar ERB data findings of larger rise excursions for female speakers of New Zealand English (Warren and Daly 2000; Daly and Warren 2001).

The intonational features that distinguish African American speech as an ethnicity-based dialect in our North Carolina data resemble some features reported in earlier, largely impressionistic, studies. We observe the wider F_0 range for AAE speakers, albeit only for the males in our corpus, though we do not observe the use of falsetto for African American males. We also find a pattern of dynamic rising-falling F_0 contours marking pitch accent for African Americans, especially males, which may contribute to the perception of the speech as more emphatic or as expressing more numerous and/or more prominent phrasal stress.

[4] Malcah Yaeger-Dror points out (p.c.) that declination patterns may vary with discourse context, even within a dialect or speaker. Since the discourse context is similar for both African American and European American speakers in this North Carolina corpus (all speakers were participating in ethnographic interviews), the difference in declination patterns we observe may reflect differences between the speakers in how they associate declination with discourse context.

To summarize, our study of gender and ethnicity effects in the speech of African Americans and European Americans from North Carolina yields three primary findings. First, this study produces strong evidence that females have a larger dynamic F_0 range than male speakers, extending the finding of Daly and Warren (2001) to American English varieties defined ethnically. Second, we find that a major difference between intonation patterns of African Americans and European Americans is in the encoding of gender. Gender distinctions are similar in type but greater in magnitude for European American speakers, indicating that the role of intonation as an index of gender varies across speech communities. The elevated F_0 register for European American females and low and compressed F_0 range for European American males may reflect no more than a phonetic encoding of gender identity. Conversely, it is possible that these F_0 range properties reflect the speaker's attitude towards the propositional content of an utterance, as Clopper and Smiljanić (this volume) suggest, which would point to encoding of genderlect distinctions at the level of pragmatics in the intonational phonology.

Our third finding is a partial confirmation of our hypothesis from our pilot study that African Americans produce more prominently falling F_0 contours over the pitch accent domain than European Americans. Males provide the strongest supportive evidence. The pronounced F_0 falling contour across the pitch accent domain may reflect a H*+L pitch accent as a distinctive phonological feature of AAE or it may arise from a Low tone assigned to a minor phonological phrase, which would point to a dialectal difference in prosodic phrasing. These differences in phonological encoding of intonation must be investigated in future work and call for more detailed phonological transcription of the tonal elements of intonation.

ACKNOWLEDGMENTS

This research is supported by NSF grant IIS 04-14117 and a University of Illinois Research Board award to Jennifer Cole. NCLLP recordings used here were conducted under NSF grants SBR-9319577 and SBR-9619331; analysis was supported by NSF grant BCS-0542139. Special thanks go to our research assistants in prosody transcription: Zachary Hulstrom , Ayesha Saied and Young-il Oh. We appreciate Chilin Shih, José I. Hualde, members of the UIUC Prosody & ASR research group, and participants in the NWAV-34 special panel on intonation for providing helpful feedback and discussion.

REFERENCES

Arvaniti, Amalia, and Gina Garding. 2007. Dialectal variation in the rising accents of American English. In Jennifer Cole and José I. Hualde (eds.), *Laboratory Phonology 9*. Berlin: Mouton de Gruyter, 547-576.

Beckman, Mary, and Gayle Ayers Elam. 1997. Guidelines for ToBI labeling. Downloaded from http://www.ling.ohio-state.edu/research/phonetics/E_ToBI

Beckman, Mary E. and Janet B. Pierrehumbert. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook* 3: 255-309.

Boersma, Paul, and David Weenink. 2007. Praat: doing phonetics by computer (Version 4.6.36) [Computer program].Retrieved from /www.praat.org/.

Cohen, Jacob. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20: 37–46.

Clopper, Cynthia, and Rajka Smiljanić. (this volume). Effects of Gender and Regional Dialect on Prosodic Patterns in American English.

Daly, Nicola, and Paul Warren. 2001. Pitching it differently in New Zealand English: Speaker sex and intonation. *Journal of Sociolinguistics* 5: 85-96.

Fletcher, J., Esther Grabe, and Paul Warren. 2005. Intonational variation in four dialects of English: the high rising tune. In Sun-Ah Jun (ed.) *Prosodic typology. The Phonology of Intonation and Phrasing*. Oxford, OUP.

Grabe, Esther. 2004. Intonational variation in urban dialects of English spoken in the British Isles. In Peter Gilles and Jörg Peters (eds.) *Regional Variation in Intonation*. Linguistische Arbeiten, Tübingen: Niemeyer, 9-31.

Green, Lisa J. 2002. *African American English: A Linguistic Introduction*. Cambridge, U.K., and New York: Cambridge University Press.

Grice, Martine, Stefan Baumann and Ralf Benzmüller 2005. German Intonation in Autosegmental-Metrical Phonology. In Jun, Sun-Ah (ed.) Prosodic Typology: The Phonology of Intonation and Phrasing. Oxford University Press.

Goodwin, Marjorie Harness, Charles Goodwin, and Malcah Yaeger-Dror. 2002. Multi-modality in girls' game disputes. *Journal of Pragmatics* 34:1621-49.

Gut, Ulrike, and J. T. Milde. 2002. The prosody of Nigerian English. In *Proceedings of Speech Prosody 2002.*

Haan, Judith, and Vincent J. van Heuven. 1999. Male vs. female pitch range in Dutch questions. *Proceedings of the Thirteenth International Congress of Phonetic Sciences,* San Francisco. 1581-1584.

Hermes, Dik J., and Joost C. van Gestel. 1991. The frequency scale of speech intonation. *Journal of the Acoustical Society of America* 90: 97-103.

Hudson, Amelia I., and Anthony Holbrook. 1981. A study of reading fundamental vocal frequency of young black adults. *Journal of Speech and Hearing Research* 24: 197-201.

Hudson, Amelia I., and Anthony Holbrook. 1982. Fundamental frequency characteristics of young black adults: Spontaneous speaking and oral reading. *Journal of Speech and Hearing Research* 25: 25-28.

Jun, Sun-Ah, and Christina Foreman. 1996. Boundary tones and focus realization in African-American intonation. Paper presented at the 3rd joint meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii, 6 December.

Ladd, D. R. 1996. *Intonational Phonology.* Cambridge: Cambridge University Press.

Loman, Bengt. 1975. Prosodic patterns in a Negro American dialect. In Håkan Ringbom, Alfhild Ingberg, Ralf Norrman, Kurt Nyholm, Rolf Westman, and Kay Wikberg (eds.), *Style and Text: Studies Presented to Nils Erik Enkvist*. Stockholm: Språkförlaget Skriptor AB, 219-42.

McConnell-Ginet, Sally. 1978. Intonation in a man's world. *Signs: Journal of Women in Culture and Society* 3: 541-559.

Syrdal, Ann K., and Julia McGory. 2000. Inter-transcriber reliability of toBI prosodic labeling. In *Proceedings of the International Conference on Spoken Language Processing 2000*, vol.3, 235-238, Beijing, China.

Tarone, Elaine E. 1973. Aspects of intonation in Black English. *American Speech* 48: 29-36.

Warren, Paul, and Nicola Daly. 2000. Sex as a factor in rises in New Zealand English. In John Holmes (ed.) _Gendered Speech in Social Context: Perspectives from Town and Gown_. Wellington: Victoria University Press, 99-115.

Wolfram, Walt, and Erik R. Thomas. 2002. *The Development of African American English*. Language and Society 31. Oxford, U.K., and Malden, MA: Blackwell.

Yoon, Tae-Jin, Sandra Chavarría, Jennifer Cole, and Mark Hasegawa-Johnson. 2004. Intertranscriber reliability of prosodic labeling on telephone conversation using ToBI. In *Proceedings of the ISCA International Conference on Spoken Language Processing* (Interspeech 2004), Jeju, Korea, pp. 2729-2732.

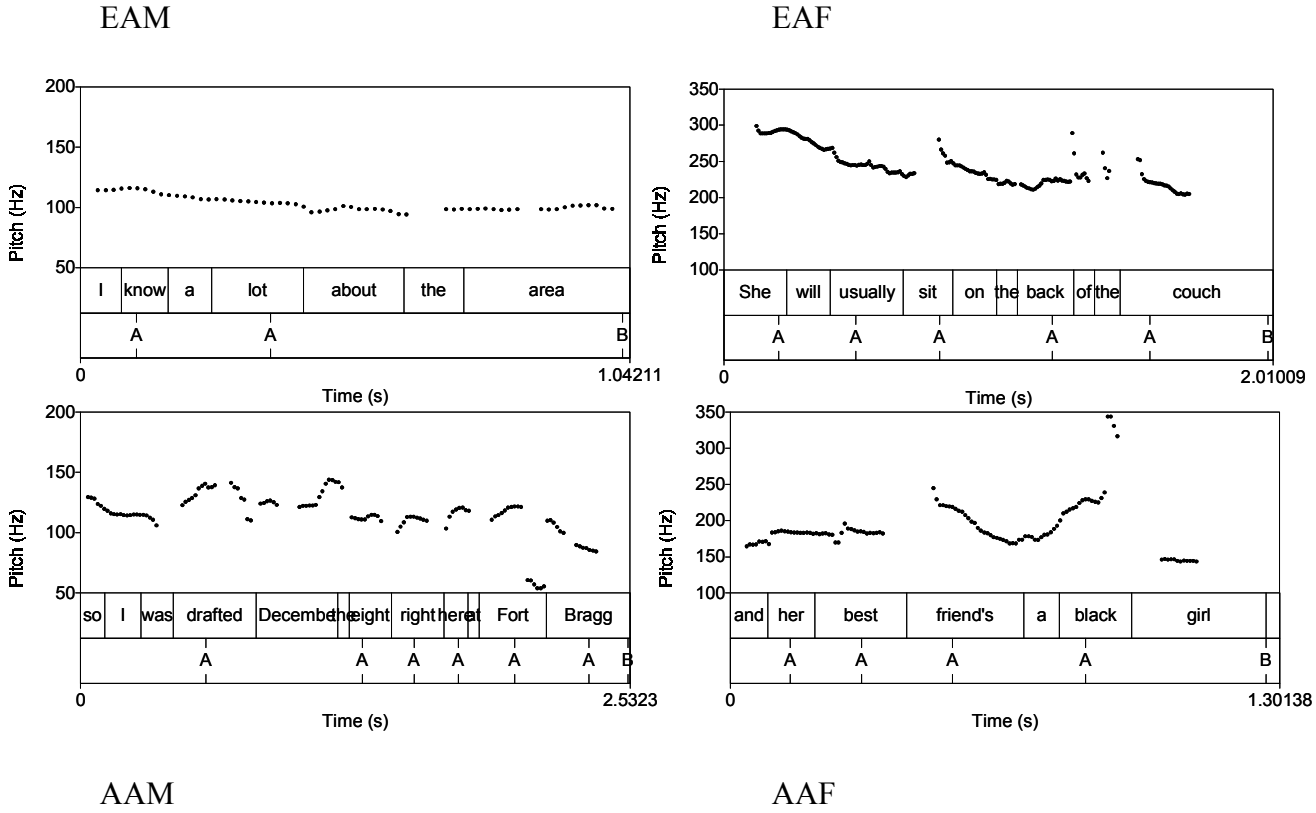

Figure 1. Sample F_0 tracks.Upper left: European American Male "*I know a lot about the area*;*"* upper right: European American Female *"She will usually sit on the back of the couch*;*"* lower left: African American Male *"so I was drafted December the eight right here at Fort Bragg*;*"* lower right: African American Female *"and her best friend's a black girl."*

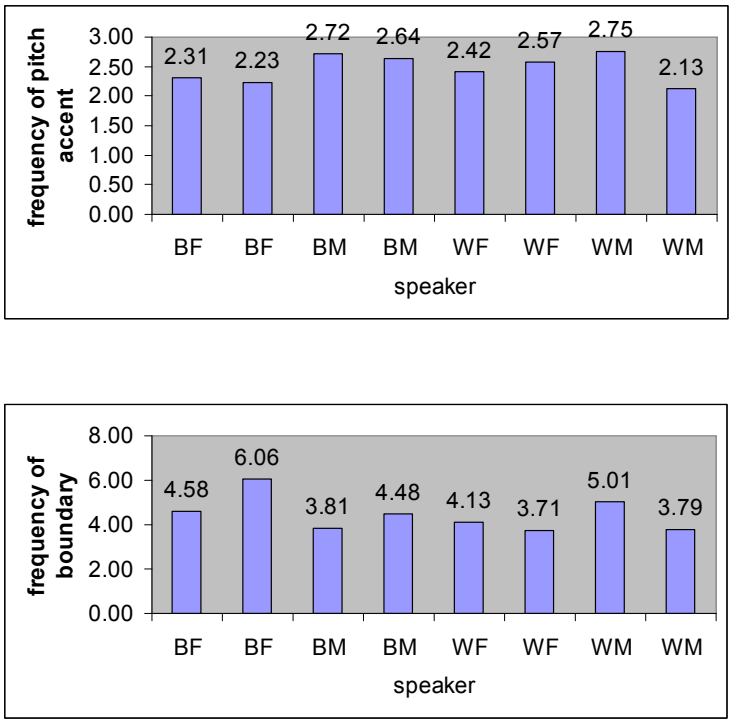

Figure 2. Frequency of accent (top panel) and prosodic phrase boundary (bottom panel) for each speaker as a function of the number of words for that speaker (*number of words / number of accents).* Speakers identified as African American female or male (AAF, AAM) or European American female or male (EAF, EAM).

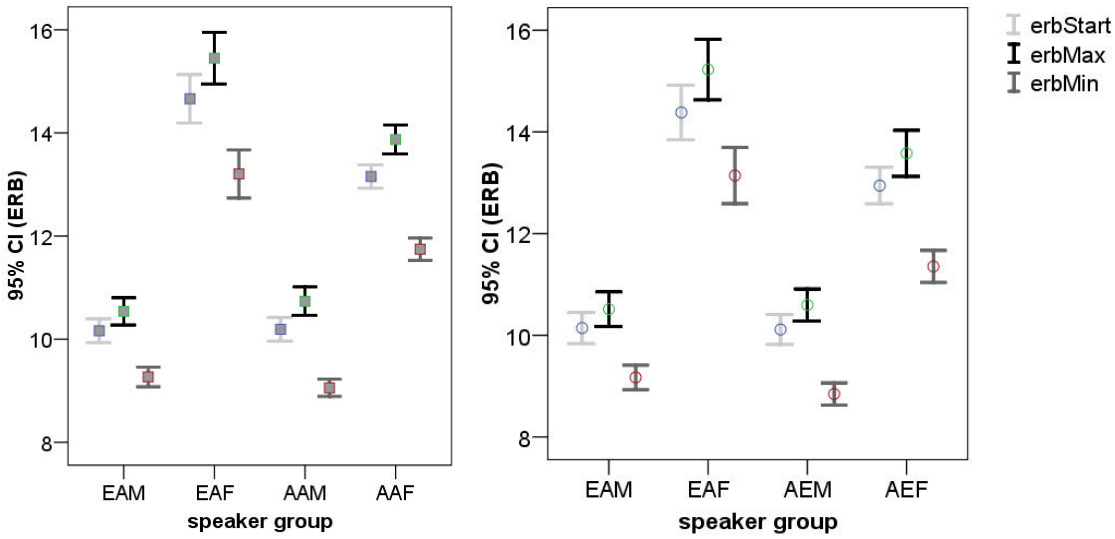

Figure 3. Start, Max and Min F_0 values (ERB) for phrase-medial (left panel) and phrase-final (right panel) pitch accents, 95% Confidence Intervals.

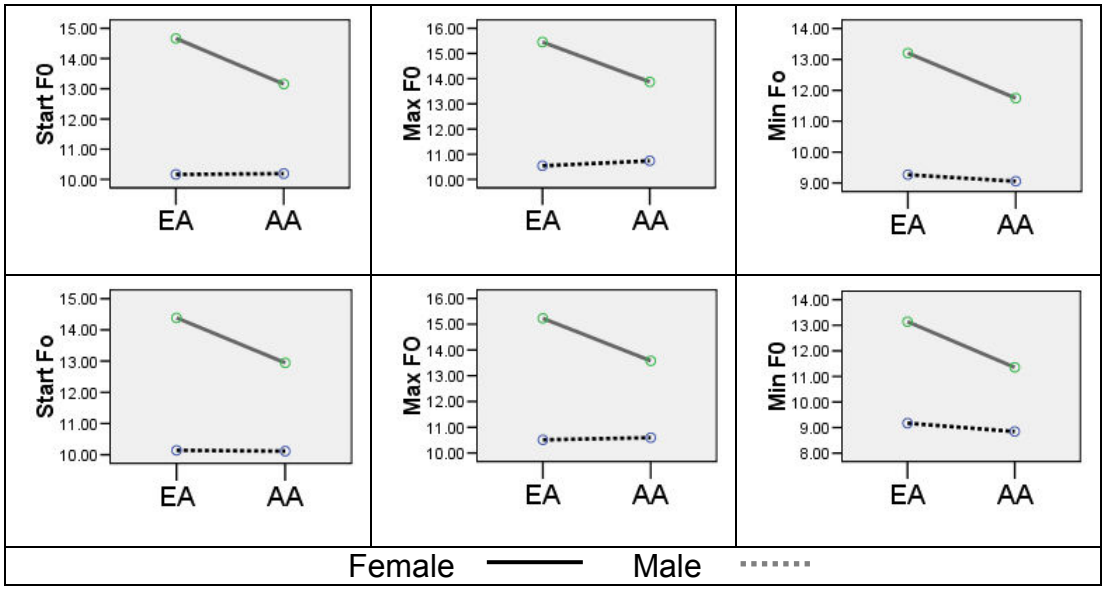

Figure 4. Mean plots showing effects of ethnicity (European American=EA and African American=AA) on Max F_0 and F_0 range, for males and females. Phrase-medial F_0 means in top panels; phrase-final F_0 means in bottom panels. All F_0 values in ERB units.

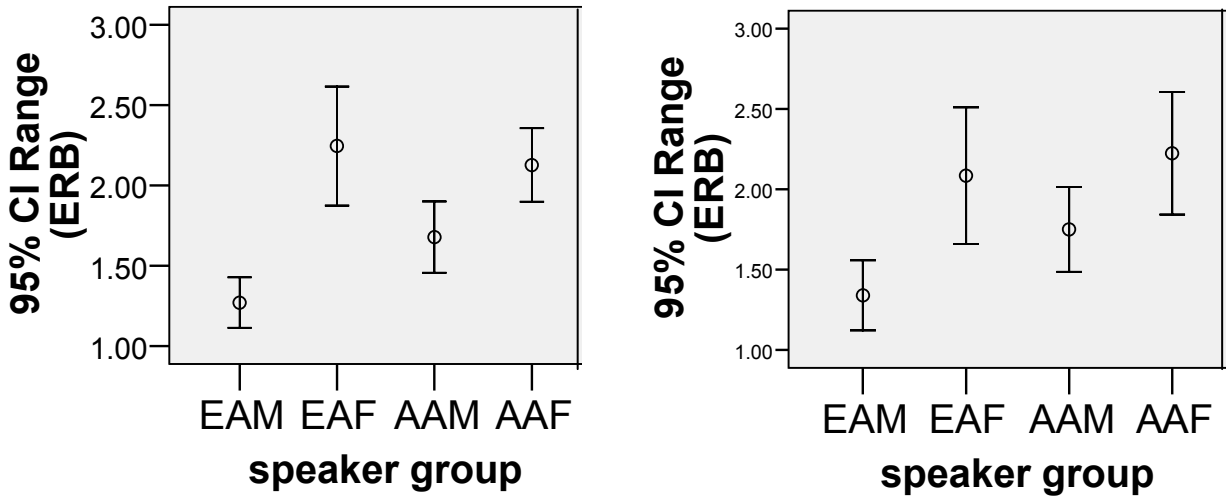

Figure 5. F_0 range (ERB) for speakers grouped by ethnicity-gender: European American males (EAM), European American females (EAF), African American males (AAM) and African American females (AAF). Phrase-medial (left panel) and phrase-final (right panel) pitch accents.

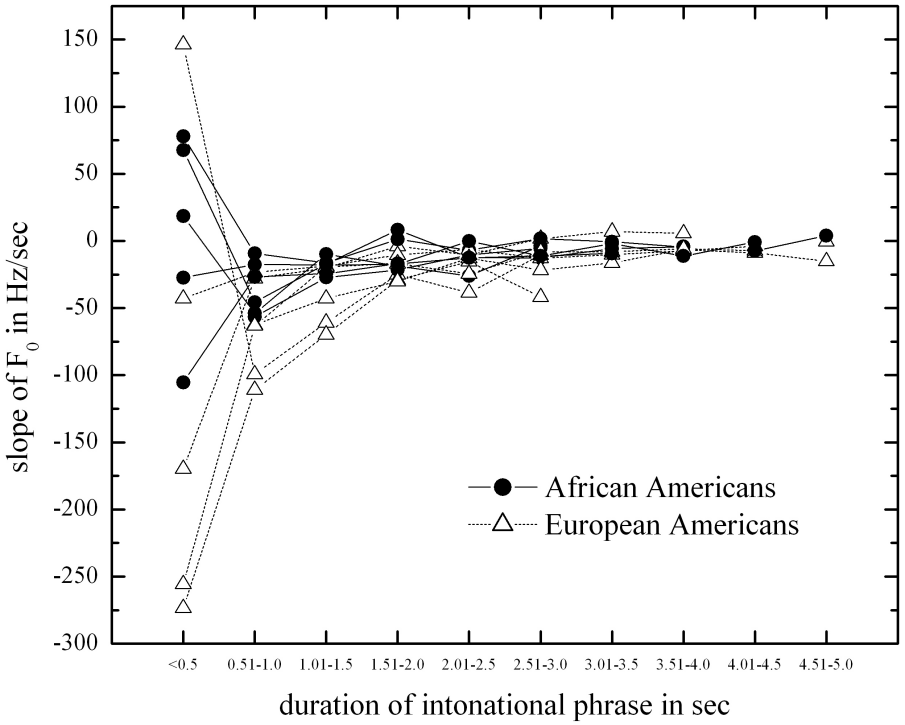

Figure 6. Variation in F_0 slope over the Intonational Phrase as a function of phrase duration for African American speakers ● and European American speakers △.