

A COMPARISON OF THE ACOUSTIC CORRELATES OF FOCUS IN INDIAN
ENGLISH AND AMERICAN ENGLISH

By

RUSSELL MOON

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

UNIVERSITY OF FLORIDA

2002

Copyright 2002

by

Russell Moon

For Luther, who always had the time and patience to fix things.

ACKNOWLEDGMENTS

First and foremost, I really need to thank my advisors, Dr. Caroline Wiltshire and Dr. Ratreë Wayland, without whose very patient guidance I would still be editing data. I credit them with the pride I feel in this thesis. Any detectable trace of rigor or insight is their handiwork.

I also must thank everyone who helped and participated with data collection and my understanding of Indian English, especially Vijay Vijayakrisnan and Vuday Nandur. I would also like to thank Paul Boersman for creating the indispensable Praat.

Jimmy Harnsberger deserves his own paragraph because he helped me to figure out the statistics. He needs thanks in proportion to his help.

My friends (who could not all fit into this thesis) were a big source of strength and support. Cydney Alexis, Jodi Bray, Jonathan Frome, Luli Lopez-Merino, Karen Regn and Bill Ward deserve special awards for understanding and givingness. Jodi gave me the single best piece of advice I have ever heard. It got me through a lot.

Finally, I need to thank Mom, Dad, Toni, Chris, Matt, Granny, Ma, Pop, Annette, John, Suzanne and Gail for obvious reasons.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
ABSTRACT	x
CHAPTER	
1 INTRODUCTION	1
2 LITERATURE REVIEW	3
Stress Terminology	3
Stress and Focus	3
Stress	3
Focus	4
Pitch Accent	4
The Phonetic Analysis of Sentence-Level Focus	5
Fundamental Frequency	5
Duration	6
Perceptual Cues of Loudness	6
Focus Differences Between Indian English and American English	7
Acoustic Differences Between Focus in AE and IE	7
Wiltshire and Moon (2000)	7
Autosegmental Metrical Theory	8
Introduction	8
Background	8
Implications of the Current Study for AM Theory	10
3 METHODOLOGY	12
Participants	12
Data Collection	13
Speaker Preparation	13
Targets and Sentence Production	13

Recording	14
Acoustic Analysis	14
Fundamental Frequency	15
Amplitude	15
Duration	15
Measurements and Statistical Analyses	16
Cross-Dialectal Analysis	17
Ratio analysis	17
Difference analysis	18
Statistical analysis	18
Difference vs. ratio measurements	19
Intra-Dialectal Analysis	19
4 RESULTS	21
Cross-Dialectal Results	21
Intra-Dialectal Results	24
Focus Cue Results	24
F0 Contour Results	26
5 DISCUSSION	31
Implications for the Phonetic Analysis of Focus	31
Maximum Fundamental Frequency	31
Duration	32
RMS Amplitude	33
Summary	33
Implications for Cross-Dialectal Variations in Focus	34
Description of Cross-Dialectal Differences in Focus Cues	34
Fundamental frequency	34
Duration	35
RMS amplitude	36
Implications for Wiltshire and Moon 2000	37
RMS amplitude and max-intensity	37
Fundamental frequency	38
Duration	39
Summary and discussion	39
Implications for the Pitch Accent Types of Focus in IE and AE	40
Hindi L1 Pitch Accent Type	40
Telugu L1 Pitch Accent Type	41
American English Pitch Accent Type	41
Summary	41
Summary	42

APPENDIX

A LANGUAGE BACKGROUND QUESTIONNAIRE	44
B PITCH TRACK AND SPECTROGRAM EXAMPLES	46
LIST OF REFERENCES	48
BIOGRAPHICAL SKETCH	51

LIST OF TABLES

<u>Table</u>	<u>page</u>
4-1. Ratio test and difference test methodological summary.....	21
4-2. Summary of difference test and ratio test cross-dialectal results	23
4-3. Summary of intra-dialectal results, all target words.....	24
4-4. Summary of intra-dialectal results, <i>bait</i> only	25

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
4-1. <i>say bait</i> pitch contour in all dialects	27
4-2. <i>say bet</i> pitch contour in all dialects	28
4-3. <i>say boat</i> pitch contour in all dialects	28
4-4. <i>say bit</i> pitch contour in all dialects	29
4-5. <i>say</i> + token pitch contour in American English	29
4-6. <i>say</i> + token pitch contour in Hindi L1 English.....	30
4-7. <i>say</i> + token pitch contour in Telugu L1 English	30

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts

A COMPARISON OF THE ACOUSTIC CORRELATES OF FOCUS IN INDIAN
ENGLISH AND AMERICAN ENGLISH

By

Russell Moon

December 2002

Chair: Dr. Caroline Wiltshire
Cochair: Dr. Ratree Wayland
Major Department: Linguistics

This study's broad objective was to examine the acoustic cues of linguistic focus in three dialects of English: American English (AE), Hindi L1 English (HE) and Telugu L1 English (TE). The three specific goals of the study were to (1) describe how focused and unfocused words differ acoustically in each dialect, (2) to describe the differences in focus among these dialects and to discuss the ways in which differences in focus are relevant to cross-dialectal differences in word-level stress, and (3) to develop a partial model of pitch accent for focus, according to the principles of Autosegmental Metrical (AM) Theory.

In total, 4 AE, 3 HE and 4 TE speakers participated in the study. Participants read a list of sentences into a microphone and their voice data were recorded on DAT tape and analyzed using Praat speech analysis software. The vowels of focused and non-focused

words in each sentence were measured for seven types of focus cue: duration, RMS amplitude and five types of fundamental frequency (F0) measurements.

The results showed that the acoustic cues of focus for the three varieties of English differed significantly among each other in the measures of duration, RMS amplitude and two of the five F0 measurements. Within each variety of English, there were significant differences among the acoustic cues of focus for unfocused and focused words. AE speakers produced unfocused words which differed significantly from their focused words in RMS amplitude, duration and four of the F0 measures. HE speakers produced unfocused words which differed significantly from their focused words in RMS amplitude, duration and four of the F0 measures. TE speakers produced unfocused words which differed significantly from their focused words in RMS amplitude and three of the F0 measures. I also developed models of the pitch contours—or *pitch accents* according to AM Theory—associated with focused words in each dialect.

In conclusion, this study found that the acoustic correlates of focus of American English accord with other such studies except in the measure of RMS amplitude. It also found that cross-dialectal differences in focus are not equivalent to cross-dialectal differences in word-level stress, which indicates that the acoustic correlates of speech at different prosodic levels are at variance.

INTRODUCTION

The research conducted for this study is situated at the intersection of the acoustic study of semantic focus and the study of cross-dialectal variation between Indian English and American English. It follows closely on the heels of Wiltshire and Moon (2000), a study which examined the dialectal variation of word stress between Indian English and American English.

Focus is a prosodic feature associated with the introduction of new information or the highlighting of contrastive or important information in a sentence. This study examines the differences between acoustic correlates of focus in American English (AE) and two varieties of Indian English (IE)—the English dialect used by Hindi L1 speakers and the English dialect used by Telugu L1 speakers. The acoustic correlates of focus, such as F0, amplitude and vowel duration, have been studied with regard to American English, but have rarely been examined as a cross-dialectal phenomenon.

The broad goal of this research is to contribute to the phonetic and phonological analysis of prosody, to describe quantitatively the differences among Indian English and American English dialects, and to theoretically account for the differences among these different dialects of English. Specifically stated, these three goals are

- 1) To describe the acoustic correlates of focus in Hindi L1 English, Telugu L1 English, and American English.
- 2) To describe the differences in focus among these dialects and to discuss the ways in which differences in focus are relevant to cross-dialectal differences in word-level stress.

3) To develop a partial model of pitch accent for focus, according to the principles of Autosegmental Metrical Theory (Pierrehumbet 1980).

The results showed that the acoustic cues of focus for the three varieties of English differed significantly amongst each other in the measures of duration, RMS amplitude and two of the five F0 measurements. Within each variety of English, there were significant differences among the acoustic cues of focus for unfocused and focused words. AE speakers produced unfocused words which differed significantly from their focused words in RMS amplitude, duration and four of the F0 measures. Hindi L1 English speakers produced unfocused words which differed significantly from their focused words in RMS amplitude, duration and four of the F0 measures. TE speakers produced unfocused words which differed significantly from their focused words in RMS amplitude and three of the F0 measures. The results also produced models of the pitch contours—or *pitch accents* according to AM Theory—associated with focused words in each dialect.

In conclusion, this study found that the acoustic correlates of focus of American English accord with other such studies except in the measure of RMS amplitude. It also found that cross-dialectal differences in focus are not equivalent to cross-dialectal differences in word-level stress, which indicates that the acoustic correlates of speech at different prosodic levels are at variance. Finally, this study found that the pitch accents associated with focused items differ cross-linguistically.

LITERATURE REVIEW

The literature review is subdivided in three main sections dealing with the phonetic analysis of sentence-level focus, prosodic comparisons of Indian English and American English, and Autosegmental Metrical Theory, respectively. Prior to the three main sections is a brief section on stress terminology, which has been included for clarification.

Stress Terminology

Many of the defined terms which follow were adopted from terminology presented in Ladd (1996).

Stress and Focus

Stress

Throughout this study, *stress*, *word-level stress* or *word-level prominence* refers to the perceptual prominence of one syllable over other syllables in a word. Stress is typically assigned at the lexical level and is often accompanied by an increase in perceptual loudness and vowel duration in American English (Bolinger, 1958; Fry, 1958; Crystal, 1969; Berinsein, 1979; Potisuk *et al.*, 1996—from Chen *et al.*, 2001). Stress is also sometimes accompanied by an increase in fundamental frequency (F0). According to AM theory (the theoretical model of intonation adopted in this study), F0 changes associated with stressed syllables are a result of the association of pitch-accents with stress syllables (more on pitch accents and AM theory below).

Focus

Focus, sentence-level focus or *sentence-level prominence* refers to a word that is made acoustically salient from surrounding words or phrases, usually in an attempt by the speaker to introduce new or contrastive information. In American English, the acoustic correlates of focused words have been identified as increases in vowel durations, maximum F0 and loudness.

Pitch accent

Pitch accent—the backbone of Autosegmental Metrical theory (discussed at length below)—refers to the phonological units of the intonation contour. They are the discrete tones which give shape to the pitch contour of an utterance at the sentence level. For example, the familiar intonational rise at the end of American English questions is due to the positioning of a pitch accent at the end of the question.

In the AM model, pitch accents are attached or associated to sites of stress. So, for example, the increased loudness and duration of a stressed syllable may also be accompanied by a rise in F0 due to pitch-accent association. However pitch accents are understood to be independent of lexical stress; the association of a pitch accent to a stress site occurs at the post-lexical level. While pitch accents “serve as an indirect cue to syllable prominence . . . they do not in and of themselves constitute the prominent syllable’s prominence” (Ladd 1996: 58-59). However, in contrast with stress-accent languages such as English, pitch-accent languages such as Japanese use pitch as a more direct cue to syllable prominence than cues such as duration and intensity (Beckman, 1986).

The Phonetic Analysis of Sentence-Level Focus

The first area of interest to this study involves the acoustic cues of focus.

Considerable experimental work done has been conducted in order to identify those cues which accompany the production of stress. However, studies which specifically study sentence focus are much fewer in number.

Since the earliest studies on stress, it has been reliably shown for English that word-level stress is marked by changes in F0, vowel duration and vowel intensity (Bolinger, 1958; Fry, 1958; Crystal, 1969; Berinstein, 1979; Potisuk *et al.*, 1996—from Chen *et al.*, 2001). Since the acoustic correlates of focus have been identified as being similar to those associated with stress (Folkins *et al.*, 1975; Weismer and Ingrisano, 1979; Eefting, 1990; Ferreira, 1993—from Pell 2001), methodologies and results from stress research are used justify some the methodological choices in the current study of focus. Also presented are findings from research that specifically examines focus; this study adds to the acoustic descriptions of focus developed in earlier research.

The three following sections discuss findings from phonetic studies concerning the focus cues of fundamental frequency, duration and the perceptual cues of loudness. One of the goals of the current study is to verify the various claims made about these acoustic parameters as acoustic correlates of focus.

Fundamental Frequency

Increase in F0 is a well attested cue for sentence-level focus in several studies of American English. Chen (2001) found higher F0 for vowels in focused contexts compared to the same vowels in unfocused contexts. Similarly, Cooper *et al.*, (1985) and Eady *et al.*, (1986) found that focused vowels have a higher F0 than the same vowels in unfocused contexts. In a more limited context, Eady and Cooper (1986) found F0 to be

significantly higher on sentence final focused words than non-focused words in same context.

The current study, however, does not compare focused words versus their non-focused counterparts in the same sentence environment. Rather, focused words are compared to the non-focused words which directly precede them in the sentence. There is also evidence from previous studies that focused words have higher F0 than unfocused adjacent items (Atkinson 1973 and O'Shaughnessy 1979). Another motivation for the adjacency style analysis used in this study is the observation from Ladd (1996) that the prosodic prominence of linguistic elements can only be defined in terms of their difference from other elements in the immediate environment. Throughout this paper, the comparison of focus cues between focused and non-focused adjacent words will be referred to as *adjacency analysis* or *adjacency testing*.

Duration

Durational increase is another well attested focus cue. Cooper *et al.* (1985), Eady *et al.* (1986), Eady and Cooper (1986), Pell (2001), and Chen *et al.* (2001) found significant durational increases for words in focused positions as compared to their non-focused counterparts in the same context. Weismer and Ingrisano (1979) found that focused words have greater duration than adjacent unfocused words. The current study takes the same approach as Weismer and Ingrisano (1979) and compares the duration of adjacent focused and non-focused words.

Perceptual Cues of Loudness

Chen *et al.* (2001) found that AE speakers produce focused words with greater intensity than the same word in an unfocused context, where intensity was calculated as the median of three intensity measurements taken at different locations along the vowel.

The current study measures RMS amplitude as a perceptual cue of loudness. Since RMS amplitude is proportional to intensity, intensity measurements from other studies are considered legitimate benchmarks for RMS amplitude measurements gathered in this study (Hayward 2000).

Focus Differences Between Indian English and American English

The second area of investigation is the acoustic differences in focus between Indian English and American English. Within this broad area of query, there are two specific goals for this study:

1) Describe the differences between the acoustic cues of focus in Indian English and American English.

2) To compare the results of the current study with the results of Wiltshire and Moon (2000). This will help us to determine if syllable stress and sentence-focus manifest in similar ways, across dialects. This analysis will help to answer questions about the nature of cross-dialectal prosody and is relevant for our discussion of Autosegmental Metrical theory (more on this in the following section).

Acoustic Differences Between Focus in AE and IE

To our best knowledge, there are no known studies which describe acoustic differences in focus between these two dialects. For descriptive reasons alone, this study examines the focus cues in Indian English and American English in order to offer an initial quantitative description of their differences in sentence-level focus.

Wiltshire and Moon (2000)

Wiltshire and Moon (2000) examined the acoustic correlates of stress differences in the prominent syllables of disyllabic words of both dialects, finding that AE speakers show significantly greater increases in maximum amplitude than IE speakers but did not show differences for other cues examined—maximum F0, minimum F0, medial F0 (at

the center of vowel), and duration. Results from Wiltshire and Moon (2000) study are compared to the results of the current study (see Discussion).

These comparisons allow for the description of similarities and differences between prominence manifestation at the word-level and the sentence-level. The conclusions drawn from these comparisons inform the later discussion about the AM theory for focus, with regard to pitch accent association at the site of focus.

Autosegmental Metrical Theory

Introduction

Moving away from empirical prosodic phonetics, the third area of investigation in this study is the more abstract field of prosodic phonology. The current findings have aided in the development of a partial model of pitch accent type for AM theory. The following section presents a brief overview of AM Theory and subsequent developments. The Discussion chapter presents results of the current study and their relevance to AM Theory.

Background

Many theories of intonation have been proposed during the past fifty years due in part to the difficulty in studying what appears to be a non-discrete system (Lieberman, 1967; Pike, 1945; ‘t Hart *et al.*, 1990—from Ladd 1996). Intonation does not lend itself easily to the type of discrete analysis that has been so successful in segmental phonology. For example, measuring phonetic differences in pitch between two sentences which differ only in the prominence relations of their content words (as in (1) and (2) below) has often been no more insightful than a continuous line drawn over the top of each sentence, showing a subjective interpretation of the pitch track:

(1) I shall say BAT now.

(2) I shall SAY bat now.

The work of Pierrehumbert has been a stabilizing force in the study of intonation; her dissertation (1980) and subsequent modifications (Gussenhoven, 1983; Beckman and Pierrehumbert, 1986) have produced a body of ideas about intonation that is collectively known as autosegmental-metrical (AM) theory. Central to AM is the idea that an intonational contour is comprised of a series of discrete tonal units. This notion of a sequence of tone units contrasts with earlier ideas about intonation. Before AM, many held that intonation contours could be described globally as, perhaps, a grid containing the contour, which could be superimposed wholesale on segmental features. Smaller, more localized pitch excursions—those pertaining to word prominence, for example—were seen to conform somehow to the overarching global contour.

Pierrehumbert's analysis, however, reformulates contour as the interaction between a given tone unit and its adjacent units. The maximum pitch height of a particular section of the contour, for example, is defined by the contour which precedes it and itself defines the pitch ceiling of the following section.

Pierrehumbert (1980) also laid to rest earlier disputes over the proper representation of pitch phenomena. The American structuralists (cf. Pike, 1945; Wells, 1945; and Trager and Smith, 1951—from Ladd, 1996) held that English tones could be analyzed as four phonemes: Low, Mid, High and Overhigh. Other theories, such as The Institute for Perception Research (IPO) approach, contended that the best way to refer to functional tone units was in terms of pitch movements from relatively high to relatively low positions, and *vice versa* (Ladd 1996). Working from earlier models, Pierrehumbert identified two tonal phonemes for English: H(igh) and L(ow). She further showed that H

and L tones combine in different ways to form the discrete tone units which together form the intonation contour. Tones can attach themselves to sites of word-level stress or align themselves with phrasal boundaries but are independent of lexical stress assignment rules. For example, a syllable may be stressed but not aligned with a pitch accent because the global intonation contour of the utterance in which the stressed syllable sits may not call for a pitch excursion at the site of the stressed syllable.

Tone units fall into one of two broad categories, and it is this category that determines the structure of the unit. The first category, pitch accents, are tones that are associated with a stressed syllable or a prominent word in an utterance and may take the form of H^* , L^* , H^*+L , L^*+H , $H+L^*$ and $L+H^*$. The monotonal pitch accents (H^* , L^*) are aligned with the stressed syllable. In bitonal pitch accents (H^*+L , L^*+H , $H+L^*$ and $L+H^*$), it is the starred tone which is directly aligned with the stressed syllable. The unmarked tone indicates a rapid pitch change before or after the starred tone.

The second category of tone units are known as edge tones. Edge tones are associated with phrase boundaries in the intonation contour of an utterance. Each utterance is comprised of one intonation phrase and one or more intermediate phrases. The end of each intermediate phrase is associated with a phrase accent (notated as H^- or L^-). The end of each intonational phrase is marked with a boundary tone (notated as $H\%$ or $L\%$). Since intonational phrases are exhaustively comprised of intermediate phrases, the end of every utterance is marked by an a phrase tone (H^- or L^-) followed by a boundary tone ($H\%$ or $L\%$).

Implications of the Current Study for AM Theory

An addition to AM theory relevant for the current study is Focus-to-Accent (FTA) theory, developed by Gussenhoven (1983). The theory proposes that focused words and

phrases are marked with a pitch accent. The current study adopts these ideas about the association of pitch accent with sites of focus.

Using a pitch contour model that developed from F0 measurements, the current study suggests several pitch accent types which may be associated with sites of focus in different English dialects. An examination of boundary tones or other pitch accents that may surround the pitch accent of focus, however, is out of the scope of the present study due to measurement and design limitations. It is hoped that future research can address this issue more thoroughly by creating a complete taxonomy of tone types which occur at the site and at the periphery of focus events.

METHODOLOGY

Participants

In this study, we gathered speech data from seven male Indian English speakers and four male American English speakers. All participants filled out a background questionnaire pertaining to L1 background and present use, English language background and use (when different from L1), language education, parent's language, regions of long-term residence and language use in different environments. (See Appendix A for a sample questionnaire.)

Participants were volunteers from a population of University of Florida graduate and undergraduate students. Indian English speakers were all graduate students. All had lived in India prior to arrival in the U.S. and all had resided in the U.S. less than 3 years. American English speakers grew up primarily in Florida and were between the ages of 19 and 22. None of the AE speakers had experience with an Indian language. They all used English 100% of the time, at work and at home.

The seven Indian English speakers chosen for analysis were divided up according to their L1 into sets of speakers. Set I consisted of 3 native Hindi L1 speakers; set II consisted of 4 native Telugu L1 speakers. The IE speakers ranged in age from 22 to 25 years. Their length of stay in the US averaged 19 months. The average number of years they had studied English was 19, and they used English an average of 90% at work and 36% at home.

Data Collection

Speaker Preparation

Prior to each recording session, the speakers were instructed how to read scripted sentences which showed focus in different positions according to text bolding. The example sentences were preceded by a short story which would prompt the reader to focus a particular item. For example, after a short story about a woman named **Pria** who made sandwiches and then took them out of the kitchen, the readers were asked the following question:

- (1) Who took the sandwiches? (among a given cast of characters).

Their scripted responses, which they read aloud, was presented as

- (2) **Pria** took the sandwiches.

Readers were instructed to interpret the bold word in each sentence as the sentential focus and to say the sentence as they would in natural speech. They were not instructed how to produce the focused item.

Targets and Sentence Production

Targets were placed in carrier phrases having the form

- (3) I should say **X** now.

where **X** was one of the following 11 target words: *bait* [bejt], *bat* [bæet], *beet* [bijt], *bet* [bet], *bit* [bit], *boat* [bowt], *book* [buk], *boot* [buwt], *bot* [bət], *bought* [bɔt], *but* [bʌt].

note that these are transcriptions for American English; Hindi L1 and Telugu L1 English have the same transcriptions except, in most cases, vowels that are glided—or diphthongized—in American English are unglided in the Indian English varieties.

Additionally, Hindi L1 and Telugu L1 speakers did not produce the vocal distinction between *bot* and *bought*, pronouncing both words as [bɔt].

Each target word appeared focused in two sentences for 22 (2 sentences X 11 words) carrier phrases. Further, each carrier phrase was preceded by a priming sentence and word designed to help the reader interpret the focus in each target:

(4) Should you say pig now or should you say bet now? (bet) I should say **bet** now.

(5) Should you say dog now or should you say boat now? (boat) I should say **boat** now.

The 22 priming phrase and carrier phrase pairs were randomized and printed out.

Participants were instructed to read through the list of 22 sentence pairs, first reading the priming sentence silently to themselves and then reading the carrier phrase aloud.

Recording

Participants sat in a quiet room and read the randomized list of 22 sentences into a Shure (SM 10A) Professional Unidirectional head-word dynamic microphone with the microphone at a distance of approximately 2 inches from the mouth. They were instructed to read at a natural pace and loudness. Recordings were made on a Sony TCD-D8 DAT recorder. Data were then edited and transferred to .wav files using the Cool Edit Pro software package, using a 25,000 Hz sample rate with 16-bit resolution. Each edited segment contained one carrier phrase.

Acoustic Analysis

Data were analyzed using Praat v 3.9.11 (www.praat.org), a freeware speech analysis program. The vowels of three words in each edited segments—*say*, the target word, and *now*—were measured for F0 (in mels), RMS amplitude (in dB), and duration (in

seconds). See Appendix B for some example screen captures of the Praat display of spectrogram and F0 track. For the following measurements, *vowel onset* is defined as the beginning of a well-defined F2 formant and *vowel end* is defined as the end of a visible F2 formant.

Fundamental Frequency

F0 has been used as a cue for focus in several studies (Cooper and Sorensen, 1981; Bruce, 1982; Eady, 1983; Kutik, *et al.*, 1983; Liberman and Pierrehumbert, 1984). The present study made the following F0 measurements in mels (which correspond more closely to the human perception than Hertz):

- i) 10 ms after the vowel onset (Onset-F0)
- ii) half-way between the vowel onset and end as measured for duration (Mid-F0)
- iii) 10 ms before vowel end (Final-F0)
- iv) the highest F0 in the vowel (Max-F0)
- v) the lowest F0 in the vowel (Min-F0)

In addition to these measurements, the pitch contour of each vowel was descriptively recorded. Each pitch contour was defined by three points—the beginning, middle and end of F0 for each vowel—as given by the measurements recorded in (i-iii).

Amplitude

Amplitude (in decibels) was measured as RMS amplitude over the entire vowel of each word, from vowel onset to vowel end.

Duration

Both Cooper, *et al.*, (1985) and Eady, *et al.*, (1986) found significant durational increases for words in focused positions as compared to their non-focused counterparts in

the same context. In this study, duration was measured over the entire vowel of each word, again from vowel onset to vowel end.

Measurements and Statistical Analyses

Two types of analyses were used in this study to compare focus cues across dialects and within dialects, respectively. The Cross-Dialectal analysis compared the ratios and differences of focus cues in target words and preceding unfocused words to the cue ratios and cue differences of similar focused-unfocused pairs in other dialects. The Intra-Dialectal analysis compared the focus cues of focused words and unfocused words within the same sentence, for the same speaker group. The goal of the intra-dialectal analysis was to determine if there were significant differences in cue measurements between unfocused and focused words in the same dialect. For example, this analysis asked the question, *Does the length of focused vowels differ significantly from the length of unfocused vowels for American English speakers?*

The goal of the Cross-Dialectal analyses was to determine cross-dialectal differences in focus cues. For example, this analysis asked the question, *Is the change in max-F0 from unfocused to focused vowels in American English significantly different from the change in max-F0 from unfocused to focused vowels in Hindi L1 English?*

In order to analyze focus cues, many studies have compared the same word in both unfocused and focused contexts. For example, this type of analysis would compare the focus cues for the word *bait* in the following sentences:

(6) I should say **BAIT** now.

(7) I should say bait now.

Sentence (6) contains a focused version of *bait* and (7) contains an unfocused version.

This study, however, compares unfocused and focused words within the same sentence:

(8) I should say **BAIT** now

The acoustic cues of unfocused *say* are compared to the corresponding cues in focused *bait* (the target word). Unlike some types of focus analysis, these words are dissimilar and they appear adjacent to each other in the same sentence. The unfocused word in this study is always the word *say*. The focused word is one of the 11 target words which directly follows *say* in the sentence.

The main drawback to this type of adjacency analysis is the difference between the words under comparison. The vowels in *say* and the target word are different in most cases (the [ej] or [e] in *say* is different from the vowels in the target words, with the exception of [bejt] or [bet]). Since different vowels have different intrinsic acoustic qualities, there is potential for confounding effects here. Also, the consonant environments of *say* and the target are different—[s_ j] or [s_ #] for *say* and [b_ t] for the targets—which could give rise to differences in acoustic cues not caused by differences in focus.

Cross-Dialectal Analysis

Ratio analysis

The ratio analysis calculated the ratio of the acoustic cues for the focused target in each sentence (*bait, bat, beet, bet, bit, boat, book, boot, bot, bought, and but*) to the word *say*. For example, the duration of the vowel in *say* and the vowel in *bat* in the following sentence were measured:

(9) I should say **BAT** now.

Measurements were made for 6 of the 7 cues in each word. To find the ratio, the cue measurement for the target was divided by its corresponding cue in the word *say* ($Cue_{\text{target}}/Cue_{\text{say}}$). Note that ratio measurements were not made for the cue of RMS amplitude, since results given in decibels are already in a logarithmic ratio scale. For RMS, the difference between values is equivalent to their ratio. For example, the difference between 60dB and 50dB is 10dB, and this 10dB difference represents a 100-fold difference in intensity. The ratio of 60dB to 50dB is 1.2, which is meaningless in terms of decibels, since any 10dB increase signifies a 100-fold increase in intensity (regardless of the initial dB).

Difference analysis

In the difference analysis, the same measurements used in the ratio analysis were used to find the numerical differences between the cues for the focused target and *say* ($Cue_{\text{target}} - Cue_{\text{say}}$).

Statistical analysis

Once these calculations were complete, the ratios and the differences were compared across language groups for Hindi L1, Telugu L1 and AE speakers. A univariate analysis (making pairwise comparisons among estimated marginal means with a Bonferroni confidence interval adjustment) was used to determine significant differences ($p < .05$) in focus cues between the following three groupings of speakers:

- i) Hindi L1 and AE
- ii) Dravidian L1 and AE
- iii) Hindi L1 and Dravidian L1

Sets of ratios and differences were divided into groups according to the target word. Statistical comparisons were made such that only sentences having the same target word were tested for significant difference.

Difference vs. ratio measurements

The Cross-Dialectal analysis tests for changes of both subtractive difference and ratio between stress cues of adjacent words. Which is the superior approach? If one assumes that stress perception is based on the detection of difference between adjacent words, regardless of the average F0, amplitude or duration of a whole utterance, then ratio would probably be the best measure of stress. However, if there is a minimum threshold for stress change that is based on absolute value, then a difference measurement might be most appropriate.

For example, imagine two speakers at opposite end of the F0 spectrum. Speaker 1 produces, on average, F0 values in the range of 100Hz and speaker 2 produces F0 values around 200Hz. If we assume that a 10% ratio of change is most important for the production and perception of focus cues, then speaker 1 must raise his F0 an average of 10Hz to produce an perceptible stress shift and speaker 2 must raise her F0 by 20Hz.

Now assume that absolute difference is most crucial for communicating a stress shift and that the minimum threshold for F0 increase is 20Hz. Relying on the ratio formula, speaker 1 cannot meet the 20Hz threshold with his 10HZ increase. Clearly, both difference and ratio tests of significant difference are needed to capture not only the degree, but the types of changes which may occur among cue differences. Later sections will discuss the variation in significant differences between these two types of tests and their implication for the study of focus.

Intra-Dialectal Analysis

The intra-dialectal analysis examined the differences between the same cues in adjacent words for the same sentence. The adjacent words examined were *say* and the

corresponding target word; the target word was always focused and *say* was always unfocused. For example, in sentences such as:

(10) I shall say **BAT** now.

the cues in *say* were compared to the corresponding cues in *bat*, to determine if the speaker produced these words with any significant differences between corresponding cues.

First measurements were taken for each of the seven focus cues in the *say*-target pairs. Then, the measurements were normalized with corresponding cue measurements of *now*, which appeared at the end of each sentence. For the cues of duration, onset-F0, mid-F0, final-F0, max-F0 and min-F0, the normalization formula was $(\text{Cue}_{\text{target}}/\text{Cue}_{\text{say}})$. For the cue of RMS amplitude, the normalization formula was $(\text{Cue}_{\text{target}}-\text{Cue}_{\text{say}})$, since the difference of two decibel values is tantamount to their ratios. Then, using one-tailed, paired t-tests ($p < .05$), the focus cues for stressed target words were compared with the same cues in unstressed *say*, which appeared in the same sentence as the target. As in (10) above, *say* and the target word were always adjacent.

RESULTS

Cross-Dialectal Results

The following table contains a summary of the words measured, and the formulas and cues used in calculating the cross-dialectal results. Recall that all sentences took the form: *I should say **TARGET** now.*

Table 4-1. Ratio test and difference test methodological summary

	RATIO TEST	DIFFERENCE TEST
WORDS	<p>Targets measured: <i>bait, bat, beet, bet, bit, boat, bok, boot, bot, bought, but</i></p> <p>Non-target words measured: <i>say</i></p>	<p>Targets measured: <i>bait, bat, beet, bet, bit, boat, bok, boot, bot, bought, but</i></p> <p>Non-target words measured: <i>say</i></p>
FORMULAS	$\left(\frac{\text{Cue}_{\text{target}}}{\text{Cue}_{\text{say}}}\right)$ <i>for all cues except RMS amplitude</i>	$(\text{Cue}_{\text{target}} - \text{Cue}_{\text{say}})$
CUES	<ul style="list-style-type: none"> • onset-F0 • mid-F0 • final-F0 • max-F0 • min-F0 • RMS amplitude • duration 	<ul style="list-style-type: none"> • onset-F0 • mid-F0 • final-F0 • max-F0 • min-F0 • RMS amplitude • duration

The following table summarizes those cue measures that differed significantly among the three L1 groups (American English L1, Hindi L1, and Telugu L1). Blank cells indicate that a particular cue did not differ significantly between two of the speaker

groups. For those cues which did differ significantly, the table shows the p value for all target-*say* formulas compared across speaker groups. Appearing underneath the p -values is the average of the formula results for the target word *bait* and its preceding word, *say*, for each speaker group. *Bait* was used as a representative token since (in American English) it contains the same diphthong found in *say*: [ej] (in Hindi L1 and Telugu L1 English, both *bait* and *say* contain the vowel [e]). The vocalic similarity between *bait* and *say* acts as a control for differences in vowel quality, which may give rise to cue differences unrelated to focus. For example, when prosodic features are held constant, the default F0 of [ij] is still greater than that of [ej] (Hayward, 2000). It is this type of intrinsic F0 difference that the *bait*-only analysis is meant to control.

What follows is an example walk-through of how to interpret the cross-dialectal results in Table 2. Note that Hindi L1 speakers and AE speakers differed significantly in the onset-F0 ratio analysis. The corresponding box in the table shows the p value (in this case, $p = .000$) and the averages of the *bait*-only formulas for each speaker group. The AE speakers measured .977 in the ratio test for *bait*, meaning that the onset-F0 of *bait* was, on average, 97.7% of the onset-F0 of *say*. The Hindi L1 speakers, however, measured .903 in the ratio test for *bait*, meaning that the onset-F0 of *bait* was, on average, 90.3% of the onset-F0 of *say*. From these formula averages for *bait*, we can tentatively conclude that both speaker groups (AE and Hindi L1) slightly lower their onset-F0 from unfocused *say* to focused *bait*. Further extrapolating, we might assume that all the *say*-target word pairs undergo such a shift, if we ignore the effects of intrinsic F0 for different vowels. Finally, we note that, while both speaker groups lower their onset-F0, Hindi L1 speakers lower it to a greater degree than American English speakers. Looking at the

intra-dialectal which follows, this observation takes on greater significance since Hindi L1 speaker show a significant difference between onset-F0 in adjacent words while AE speakers do not. This accords with the above observation that Hindi speakers lower their onset-F0 to a greater degree than AE speakers.

Note that there was also a significant difference in the difference tests of onset-F0 between Hindi L1 and AE speakers. The AE speakers averaged -2.60 mels in the difference test of onset-F0 for *bait*, and the Hindi L1 speakers averaged -21.26 mels. Again, we can interpret these averages as meaning that both speaker groups lower their onset-F0 from unfocused to focused words, but that Hindi L1 speakers do so to a greater degree than AE speakers.

Table 4-2. Summary of difference test and ratio test cross-dialectal results

	Hindi L1 vs. AE		Telugu L1 vs. AE		Hindi L1 vs. Telugu L1	
	Diff target/say	Ratio target/say	Diff target/say	Ratio target/say	Diff target/say	Ratio target/say
Onset-F0	$p = .000$ HI = -21.26 mels AE = -2.60 mels	$p = .000$ HI = .903 AE = .977			$p = .000$ HI = -21.26 mels TE = 25.73 mels	$p = .000$ HI = .903 TE = 1.02
Mid-F0	$p = .012$ HI = 4.20 mels AE = 18.84 mels	$p = .000$ HI = 1.02 AE = 1.15		$p = .030$ TE = 1.13 AE = 1.15	$p = .003$ HI = 4.20 mels TE = 44.50 mels	$p = .004$ HI = 1.02 TE = 1.13
Final-F0						
Max-F0						
Min-F0						
RMS			$p = .018$ TE = 5.51dB AE = 2.70 dB		$p = .001$ HI = 3.06 dB TE = 5.51 dB	
Duration	$p = .000$ HI = .035 s AE = .052 s	$p = .000$ HI = 1.33 AE = 1.49	$p = .000$ TE = -.028 s AE = -.052 s	$p = .000$ TE = 1.52 AE = 1.49		

Hindi L1 speakers differ significantly from American English speakers and Telugu L1 speakers in both the onset-F0 and mid-F0 cues. AE speakers differ from both Hindi L1 and Telugu L1 speakers in the duration cue. Telugu L1 speakers differ from AE and Hindi L1 speakers in the RMS amplitude cue and differ from American English speakers in the mid-F0 cue.

Intra-Dialectal Results

Focus Cue Results

Using one-tailed, paired t-tests ($p < .05$), we determined if there were significant differences in cues for the target word and *say* within the same sentence. The purpose of this analysis was to determine if the three English varieties used cue shift in adjacent words to signal focus. Focus cues in the words *say* and the target word were normalized with the final word in the sentence—*now*—before statistical comparison. This normalization was performed to control for F0 variation within the same sentence. The following table contains the p -values for tests of significant difference between cues in target words and *say* for the same speaker groups:

Table 4-3. Summary of intra-dialectal results, all target words

<i>P</i> -values for cue measurements in unfocused <i>say</i> vs. focused target words for all target words using one tailed, paired t-test							
	RMS	DUR	ONSET-F0	MID-F0	FINAL-F0	MAX-F0	MIN-F0
Hindi (<i>p</i> -values)	0.000	0.462	0.000	0.023	0.000	0.400	0.109
Telugu (<i>p</i> -values)	0.000	0.072	0.446	0.000	0.000	0.058	0.000
AE (<i>p</i> -values)	0.000	0.000	0.241	0.000	0.000	0.000	0.000

Bolded values indicate significant difference ($p < .05$).

One confounding factor in these tests is the difference in vowel quality and environments of the two vowels being compared. In order to offer a more controlled set

of results, the following table contains the results for only the word *bait*. Again *bait* was chosen as the control for vowel quality because of the vocalic similarity between *say* and *bait*:

Table 4-4. Summary of intra-dialectal results, *bait* only

<i>P</i> -values for cue measurements in unfocused <i>say</i> vs. focused <i>bait</i> using one tailed, paired t-test							
	RMS	DUR	ONSET-F0	MID-F0	FINAL-F0	MAX-F0	MIN-F0
Hindi (<i>p</i>-values)	0.001	0.096	0.019	0.385	0.062	0.291	0.303
avg. (cue _{say} /cue _{now})	5.45	0.63	1.27	1.30	1.25	1.26	1.17
avg. (cue _{bait} /cue _{now})	8.51	0.80	1.14	1.32	1.44	1.28	1.26
Telugu (<i>p</i>-values)	0.000	0.077	0.154	0.043	0.075	0.149	0.176
avg. (cue _{say} /cue _{now})	3.79	0.73	0.96	1.00	1.04	1.13	1.19
avg. (cue _{bait} /cue _{now})	9.30	0.90	1.12	1.27	1.28	1.18	1.22
AE (<i>p</i>-values)	0.007	0.001	0.372	0.014	0.409	0.023	0.329
avg. (cue _{say} /cue _{now})	8.04	0.44	1.06	0.97	0.98	0.93	1.06
avg. (cue _{bait} /cue _{now})	10.74	0.64	1.05	1.11	0.99	1.04	1.04

Bolded values indicate significant difference ($p < .05$).

According to the *bait*-only analysis, speakers of all three dialects show significant rise in RMS amplitude. Hindi speakers show a significant drop in onset-F0. Telugu L1 speakers show a significant rise in mid-F0. American English speakers show a significant rise in duration, a significant rise in mid-F0 and a significant rise in max-F0. While these measurements show us the differences between acoustic cues of focus of adjacent words, they do not give us a complete picture of the pitch contours—in AM theory, pitch accents—that may be associated with focused and unfocused items. For this type of analysis, it is necessary to look at the pitch contour as defined by various measurements along the F0 track of each vowel.

F0 Contour Results

In order to develop a model of F0 contour for focused items, we plotted the onset-F0, mid-F0, and final-F0 of *say* and its succeeding target. Instead of using raw F0 values for the plots, we used the F0 values normalized with *now*, the final word in the carrier phrase. The target words *bait*, *bet*, *boat*, and *bit* were chosen for analysis in order to show the F0 contours of two representative tense vowels: [ej] (*bait*) and [ow] (*boat*), and two lax vowels: [ɛ] (*bet*) and [ɪ] (*bit*). These particular words were also chosen so that we would have examples of both front and back vowels and also vowels which are similarly articulated except for the feature of +/- tense: [ɛ] and [ej].

In the following F0 Figures, the abbreviations P1, P2 and P3 stand for onset-F0, mid-F0, and final-F0, respectively. On each x-axis, the first set of P1-P3 corresponds to the F0 points in *say* and the second set corresponds to the F0 points in the target. The first set of four charts (Figures 4-1 to 4-4) show the variations among speaker groups for each *say*-target combination, and the second set of three charts (Figures 4-5 to 4-7) show the variations among the *say*-target combinations for each speaker group.

Figure 4-1 shows the pitch contour variation for *say* and *bait* among AE, Hindi L1 and Telugu L1 speakers. The AE contour dips in the middle of *say* and rises to a peak in the middle of *bait*, whereafter it drops off. The Hindi L1 contour drops sharply, then begins to rise sharply at the onset of *bait*. The Telugu L1 contour rises steadily throughout and peaks in the middle of *bait*. The same or similar pattern is seen across the other *say* + target word pitch contour charts (Figures 4-1 to 4-4).

Figures 4-4 to 4-7, which show the combined contours of one dialect at a time, shows that the AE dialect pitch contours (Figure 4-5) rises steadily (with a possible intermediate

trough) until the midpoint of the target word, after which the pitch contour drops for all target words. Hindi L1 pitch contours (Figure 4-6) show a fairly sharp drop (with a possible intermediate rise) from the beginning of the contour to the onset of the target, followed a sharp rise through the end of the target. Telugu L1 pitch contours (Figure 4-7) show a fairly shallow climb (with possible intermediate troughs and peaks) to the midpoint of the target, after which there is a leveling-off or a continued rise of the contour.

Figure 4-1. *say bait* pitch contour in all dialects

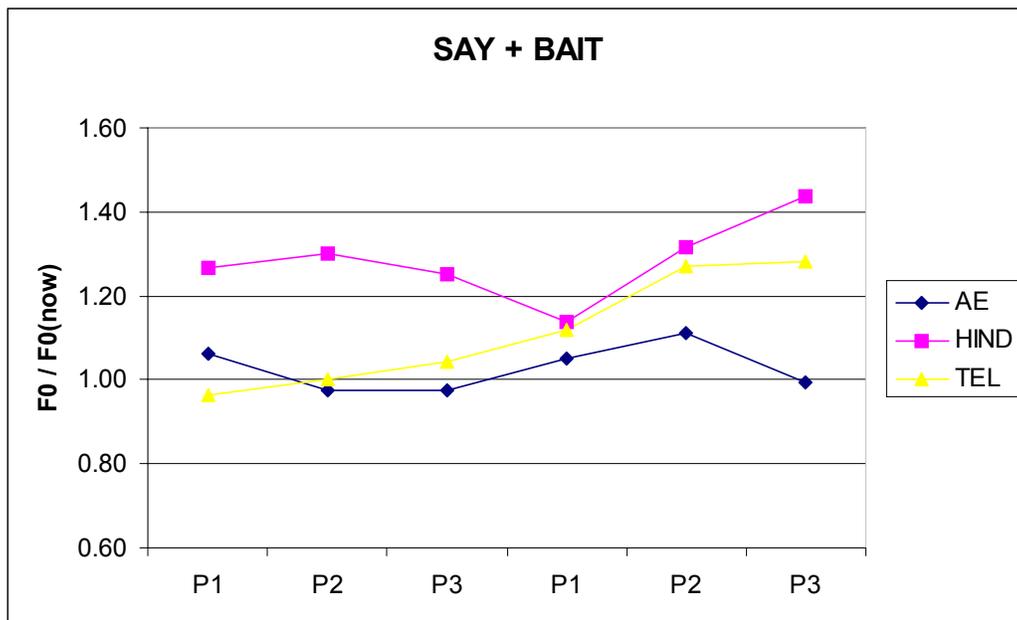


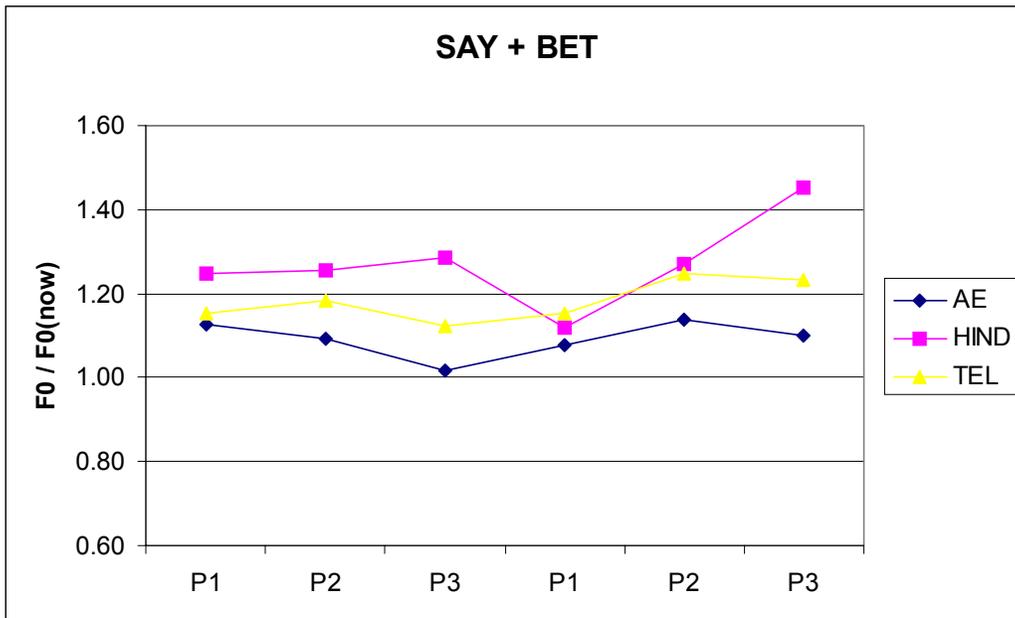
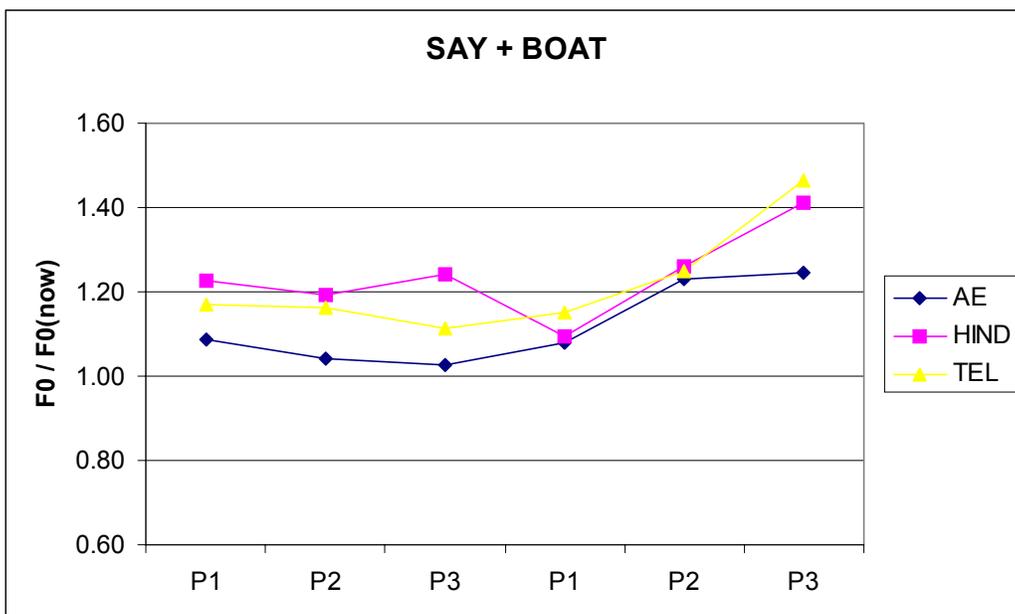
Figure 4-2. *say bet* pitch contour in all dialectsFigure 4-3. *say boat* pitch contour in all dialects

Figure 4-4. *say bit* pitch contour in all dialects

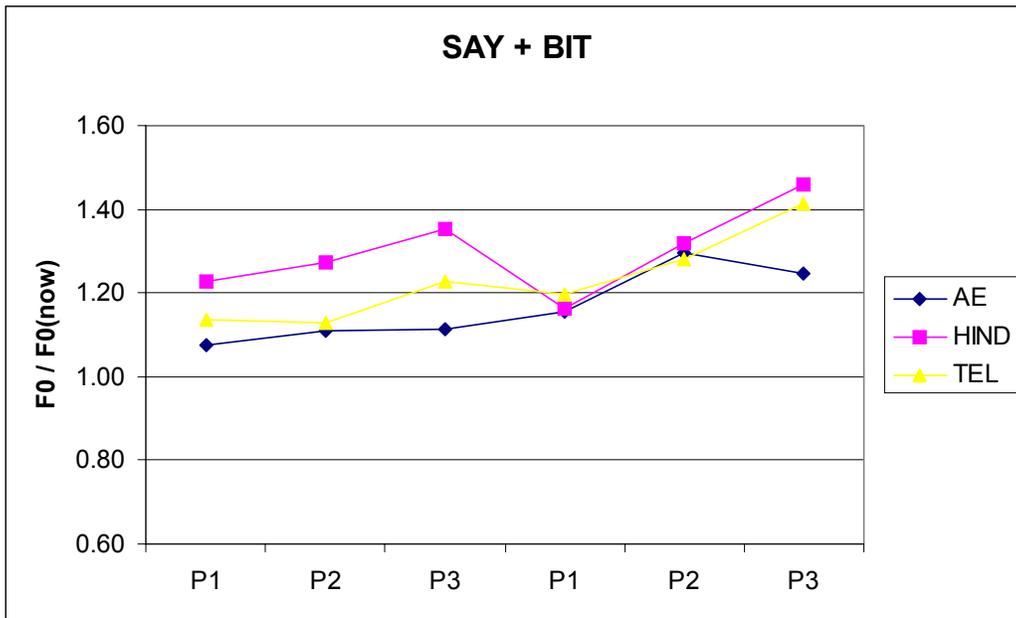


Figure 4-5. *say + token* pitch contour in American English

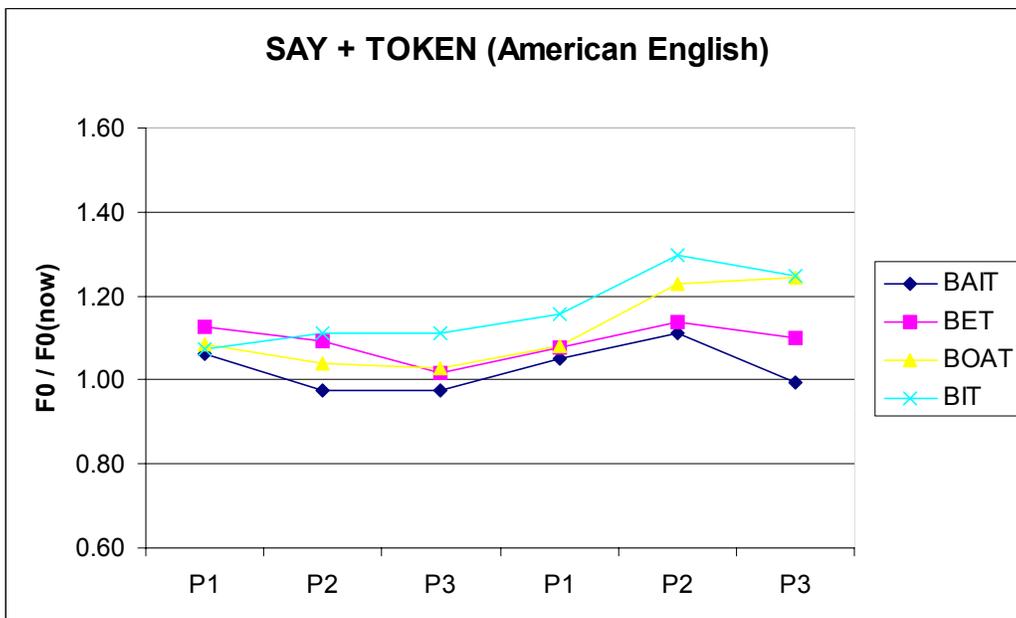
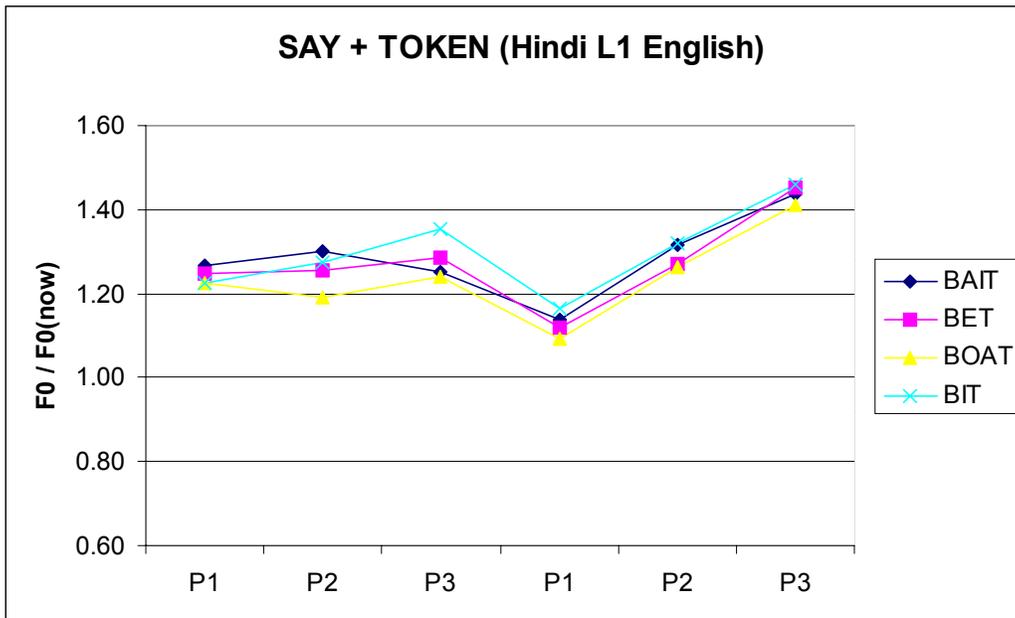
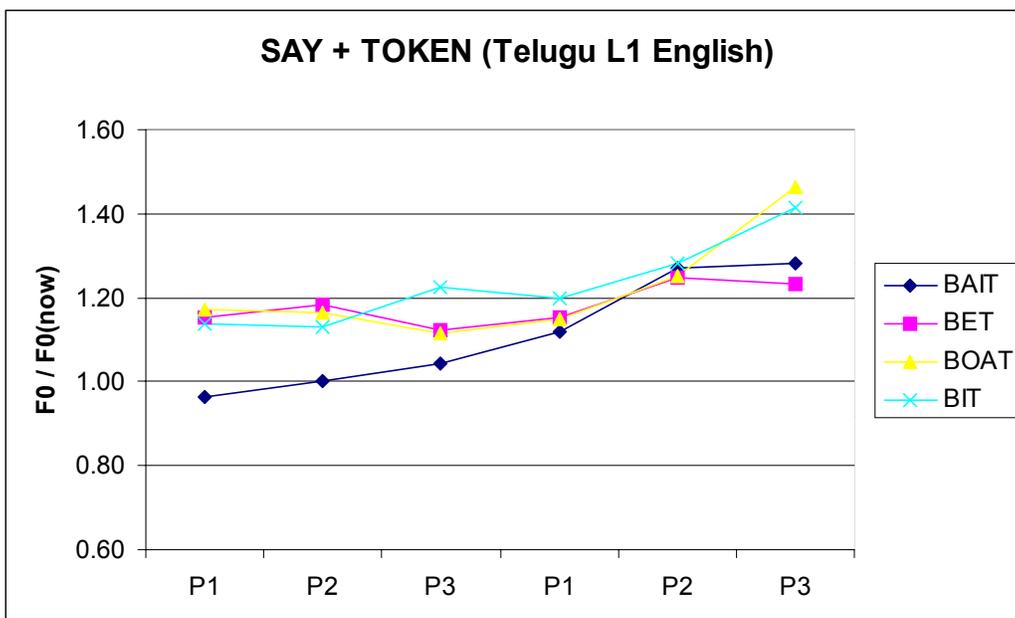


Figure 4-6. *say* + token pitch contour in Hindi L1 EnglishFigure 4-7. *say* + token pitch contour in Telugu L1 English

DISCUSSION

Implications for the Phonetic Analysis of Focus

This section of the discussion reports phonetic observations about focus without regard to AM theory or any other theory of focus or prosody. It is limited in this way to facilitate comparison with similar studies of phonetic focus which solely examine phonetic cues without theoretical analysis. The two sections following the current section will discuss the theoretical implications of the results.

Maximum Fundamental Frequency

In the intra-dialectal analysis, we find that maximum F0 differs significantly from unfocused to focused words for AE speakers but not Hindi L1 or Telugu L1 speakers. Since F0 measurements were taken directly over the vowel of the words under analysis, this significant difference between unstressed and stressed words appear to suggest some max-F0 increase directly over the focused word. For American English, this seems to agree with the well-attested findings that focus is accompanied by a rise in F0.

The intra-dialectal results for the word *bait* alone confirms our findings in the across-target test, which compared all target utterances to their preceding words. Again, AE speakers show a significant difference for max-F0 between *bait* and adjacent *say*, while the two IE varieties do not. The intra-dialectal test using *bait* only also includes the averages for the cue ratios ($\mathbf{cue}_{bait}/\mathbf{cue}_{now}$ and $\mathbf{cue}_{say}/\mathbf{cue}_{now}$) in order to determine whether *say* or *bait* had the greater cue value. Descriptively, this comparison of cue values will help determine if certain cues increase or decrease in value from non-focused to focused

words. Note that averages were not determined for the across-target tests because the variation in vowel quality between target and *say* rendered the formula averages difficult to analyze.

For the test of max-F0, we find that max-F0 is greater in AE focused words than in AE preceding adjacent non-focused words (i.e., *bait* has a max-F0 cue ratio of 1.04 and *say* has a max-F0 cue ratio of .93, meaning that max-F0 for *bait* is proportionally greater than the max-F0 of *say* when both are normalized with *now*).

Once more, this analysis controls for the fundamental differences in F0 that may be found between different vowels in *say* and the target word. Since *say* and *bait* both contain the diphthong [ej], we should expect that any differences in F0 to result from prosodic differences. Still, there is a confounding factor in that the consonantal environments of the diphthong in *bait* and *say* are different. This difference in consonantal environment may result in F0 discrepancies between the two words which are unmotivated by differences in focus.

Duration

The intra-dialectal analysis supports earlier findings that increased duration is an acoustic cue for focus in American English. The *bait* test shows that AE speakers produce focused vowels that are longer than preceding unfocused vowels. We further find that duration does not appear to be a cue for focus in the English of Telugu L1 speakers and Hindi L1 speakers. Again, the results of this analysis are somewhat confounded by the difference between the vowel in the stressed target word and its unstressed benchmark, *say*. However, the results for *bait* alone confirm the results as measured for all target words: namely, that AE shows significant durational change

between non-focused and focused words when vowel quality is kept constant, while both Hindi L1 and Telugu L1 varieties do not.

RMS Amplitude

In the intra-dialectal analysis, Hindi L1, Telugu L1 and American English speakers show a significant difference for RMS amplitude between adjacent non-focused and focused items. The findings for AE speakers support the findings of Robb *et al.*, (2001), which found a significant difference in median intensity between focused words and their unfocused counterparts in the same sentential context. Note that the different measurement types used in each study attenuates the worth of the comparison: Robb *et al.*, (2001) uses median intensity as a correlate for perceptual loudness and the present study uses RMS amplitude.

Our analysis of the target *bait* confirms this analysis: speakers of all three dialects show significant difference for RMS amplitude. Moreover, the average cue ratios for RMS amplitude in *say* and *bait* show that speakers produce focused vowels with greater RMS amplitude than preceding unfocused vowels.

Summary

Not surprisingly, when the cue values for max-F0, duration and RMS amplitude differ significantly from unfocused to focused word, the cue value for the focused word is always greater than its corresponding value in the unfocused word. This supports the well-attested observation that pitch, duration and loudness increase at sites of prosodic emphasis in English. These results further support the adjacency model, since many of the results of the generated in the adjacency model match results found in other models of focus measurement.

Implications for Cross-Dialectal Variations in Focus

The following two sections present a descriptive account of cross-dialectal cue differences for focus and a discussion of the cross-dialectal findings of Wiltshire and Moon (2000) with regards to the current findings.

Description of Cross-Dialectal Differences in Focus Cues

Fundamental frequency

In terms of F0, Hindi L1 speakers have the most distinctive pitch measurements of the three speaker groups. Hindi L1 speakers differ significantly from AE and Telugu L1 speakers in the onset-pitch and mid-pitch measures for both the difference and ratio tests. According to the formula averages for *bait*, Hindi L1 speakers produced a much lower onset-pitch and mid-pitch (as compared to preceding unfocused words). Moving from unfocused *say* to focused *bait*, Hindi speakers showed an average decrease of 21.26 mels in onset-F0 values, while AE speakers showed an average decrease of 2.6 mels and Telugu speakers an average increase of 25.73 mels in onset-F0. The mid-F0 values show a similar relationship among speaker groups.

The only other significant difference in F0 was found between Telugu L1 speakers and AE speakers in the mid-pitch ratio test. An account of why these two speaker groups differ significantly in the ratio test but not the difference test was presented earlier: although the difference between the target and *say* measurements in each group may not be significantly different, the raw values of each F0 measurement may be quite different between groups, leading to significantly different ratio measurements. For example, imagine two sets of numbers: set A (100, 120) and set B (200, 220). While the difference between the numbers in each set is 20, the ratio of the numbers in set A is .83, but the ratio of the numbers in set B is .91.

Notably, the value of max-F0 did not differ significantly among these three English varieties (however, the position of max-F0 does differ across dialects—this will be shown in the following discussion of pitch accents). One possible explanation for this is that the change in max-F0 that has been observed over focused items in English does not vary cross-dialectally. This explanation seems sound until we consider the results of the intra-dialectal tests, which reveal that only AE speakers (and not the other speaker groups) show a significant change in max-F0 from non-focused to focused items. This discrepancy between cross-dialectal and intra-dialectal tests merits further analysis, such as further testing to determine if Indian English varieties do use max-F0 increases to signal focus. The discrepancy between tests could also point to a weakness in the adjacency analysis used in this study to detect changes in perceptual cues of focus. Other types of analyses, such as those which compare the same word in focused and non-focused contexts, might be better adapted for showing cue changes associated with focus.

Although max-F0 does not differ among groups, the onset-F0 and mid-F0 do differ significantly. This seems a fairly clear indication that the pitch contours of these three varieties do differ, even if their maximum pitch values do not. I discuss these differences below in terms of Autosegmental Metrical theory.

Duration

Differences in duration across speaker groups are fairly clear cut: AE speakers produce focused vowels with greater duration (as compared to preceding unfocused vowels) than both Hindi L1 and Telugu L1 speakers. This observation is supported by both the difference test and ratio tests for both speaker pairings (Hindi L1 vs. AE *and* Telugu L1 vs. AE). Formula averages for *bait* support the assertion that AE speakers have greater increase (rather than decrease) in duration. AE speakers show an average

increase of .052 s from *say* to *bait*, while Hindi L1 and Telugu L1 speakers show an average increase of .035s and .038 s, respectively. There were no significant differences between the Indian English varieties for duration.

The duration results support two non-exclusive claims. The first claim is that American English uses greater duration increases to signal focus than other English varieties. The second claim is that Indian English may indeed constitute a unified variety of English-- notwithstanding the L1 of the Indian English speaker—which has less durational increase at the site of focus than other varieties. Such a variety would of course have regional variations based on L1 (as these results show), but also share certain commonalities, as in the case of duration. The only way to determine if either claim or both claims are correct is to examine more English dialects for durational correlates of focus.

RMS amplitude

The RMS amplitude of Telugu L1 speakers differs significantly from both Hindi L1 and AE speakers in the difference test but not the ratio test. According to the formula averages for *bait*, Telugu L1 speakers produce focused vowels with greater RMS amplitude (relative to preceding non-focused vowels) than either Hindi L1 or AE speakers. Telugu L1 speakers show a 5.51 dB increase over focused vowels, while Hindi L1 and AE speakers show a 3.06 dB increase and a 2.70 dB increase, respectively.

Judging from these results, it may be safe to claim that the Telugu L1 variety of English marks focus with a greater increase in RMS amplitude from non-focused to focused vowels. Since there is significant difference in the difference test but not the ratio test, we can suppose that while the ratios of the cue values (target/*say*) are similar

among speaker groups, the differences between these values (*target-say*) are significantly different for Telugu L1 speakers.

Implications for Wiltshire and Moon 2000

What follows is a comparison of the present results to the findings of Wiltshire and Moon (2000), which examined the difference between IE and AE stressed syllables. This section begins with a descriptive account of the similarities and differences between the two studies and then attempts to interpret the comparison. Note that the participants in Wiltshire and Moon (2000) spoke many different L1's. Our comparisons, therefore, are not exact, but Northern Indian Indo-Aryan languages (such as Hindi) and South Indian Dravidian languages (such as Telugu) were well represented as participants' L1's in Wiltshire and Moon (2000). Since Indo-Aryan and Dravidian families constituted most of the Indian languages in Wiltshire and Moon (2000), the comparisons with the current study will be profitable if there are many similarities between the varieties of English which spring from these two language families, but less so if it turns out that these varieties English differ sharply in terms of prosody.

RMS amplitude and max-intensity

The current results show that Telugu L1 speakers produce greater RMS amplitude in focused words than unfocused words when compared to AE speakers and Hindi L1 speakers. However, in a study of change in stress between unstressed and stressed syllables *within the same word*, Wiltshire and Moon (2000) found that AE speakers produce significantly greater increases in maximum intensity from unstressed to stressed syllables than IE speakers with both North and South Indian L1's.

Both RMS amplitude and max intensity were intended to capture the acoustic cue associated with perceptual loudness. However, RMS amplitude corresponds better to

perceptual loudness than peak to peak type measurements (such as maximum intensity) because perception is better aligned to the amplitude measurement across a complex waveform (RMS) than to a discrete intensity maximum (Ladefoged, 1996).

In summary, while our study shows that only Telugu speakers are significantly different from other speaker groups in terms of RMS amplitude increase, Wiltshire and Moon (2000) finds that AE speakers have greater maximum intensity increases than IE speakers. These differences in measures and speaker groups renders the comparison between studies less than ideal. Additionally, the many Indo-Aryan (Northern Indian) speakers in the Wiltshire and Moon (2000) participant group may have obscured any significant differences created shown by the Dravidian family speakers. However, it is possible that American English speakers and Indian English speakers produce acoustic cues of loudness differently for stress than for focus.

Fundamental frequency

In terms of F0 measurements, Wiltshire and Moon (2000) reports that changes in max-F0 and min-F0 between unstressed and stressed vowels show no significant differences between Indian English and American English speakers. Our results also show that increases in max-F0 and min-F0 do not significantly differ for focus between all the varieties of English. However, Wiltshire and Moon (2000) shows no significant difference for mid-F0 while the current analysis shows that mid-F0 differs significantly for focus cues between AE speakers and the all Telugu L1 speakers.

In summary, while max-F0 does not appear to differ dialectally for both stress or focus, mid-F0 does. This suggests a cross-dialectal difference in F0 contour for focus which does not show up for changes in stress.

Duration

Wiltshire and Moon (2000) found that duration does not differ significantly between AE speakers and IE speakers for word-level stress. In contrast, our results show that duration differs significantly for focus between AE speakers and Hindi L1 speakers and for AE speakers and Telugu L1 speakers. While duration changes associated with stress remain constant between AE and IE speakers, duration changes associated with focus are significantly different between AE speakers and the two IE varieties in this study.

Summary and discussion

Of the five stress cues differences examined in WM (2000), two agree with our focus cue findings (max-F0, min-F0) and three disagree (mid-F0, duration and RMS amplitude/max-intensity). At best, we can offer a descriptive analysis of this comparison. The max-F0 and min-F0 similarities between word-level and sentence-level prominence, cross dialectally, may indicate that pitch-accents (which could correspond roughly to min-F0 and max-F0) associate themselves at the word- and sentence level in the same way, cross-dialectally. The difference between the word-level and sentence-level analysis of mid-F0 may be accounted for by differences in pitch contour associated with stress vs. focus.

The differences in duration and RMS amplitude/max-intensity and suggests that loudness and duration are associated with word-level stress in different way than with sentence-level focus. Since loudness and duration increases are often assumed to apply at the lexical-level, it is possible that lexical processes of loudness and duration augmentation are less universal than post-lexical processes of F0 augmentation.

In conclusion, max-F0 min-F0 changes at the word-level and at the sentence level in English behave the same cross-dialectally. In contrast, duration increases and loudness

increases—which have been shown to manifest themselves at both the word-level and the sentence-level—differ cross-dialectally at the word- and sentence-level. A model which explains the manifestation of loudness and duration at different prosodic levels has yet to be developed, as one has for F0. One clear area of future research is a comprehensive analysis of the behavior of acoustic correlates of speech at different prosodic levels.

Implications for the Pitch Accent Types of Focus in IE and AE

The following discussion assumes, as in Gussenhoven (1983), that focused words are associated with some type of pitch-accent, according to Autosegmental Metrical theory. Figures 4-1 to 4-7 in the Results section presented pitch contours for four two word combinations (*say bait*, *say bet*, *say boat*, *say bit*) based on the data from the inter-sentential analysis. The following discussion will attempt to develop a partial tone type for focused words in American English, Hindi L1 English, and Telugu L1 English, using the tone taxonomy of the AM model. The tone type will only be partial because this study did not develop a pitch contour model for *now*, the word which succeeds the target. Since pitch accents often extend beyond the boundaries of the syllable to which the central tone is attached (on both sides), this discussion can only make conjectures about those pitch-accent types with leading tones that precede the central tone (L+H*, H+L*) or those pitch accent composed of only a central tone (H*, L*).

Hindi L1 Pitch Accent Type

The F0 contour for Hindi L1 focused words undergoes a sharp decrease in F0 from the final-F0 of *say* to the onset-F0 of the target. The low onset-F0 of the target is then followed by a rapid rise. This contour suggests an L+H* tone type, in which an L tone is followed by a rapid rise to an H. According to the AM model, a pitch contour with a local peak that rises from a relatively low level (relative to the overall pitch contour) is an

L+H* tone type. Figure 4-6 seems to fit this description, but note that an H* tone type also seems possible. However, the L + H* tone type seems to best capture differences in onset-F0 and mid-F0 between Hindi L1 and Telugu L1 and also Hindi L1 and AE that are reported by the cross-dialectal results. Only further analysis (using longer pitch contours) will be able to determine the correct tone type.

Telugu L1 Pitch Accent Type

The F0 contour for Telugu L1 focused words is, on average, a gentle slope up from *say* to the mid-F0 of the target. This contour suggests an H* pitch accent for Telugu L1 focus. The AM definition of an H* tone type is a local peak in a pitch track, which is what Figure 4-7 shows for Telugu L1. Note that for *boat* and *bit* the pitch contour approaches its maximum toward the end of the vowel. This suggests, perhaps that, at least for these two words, the pitch accent does not directly align itself with the center of the vowel.

American English Pitch Accent Type

The AE F0 contour also begins with a gentle slope up from *say*, which terminates at approximately the mid-F0 pitch position at the target and then begins to decline toward the final-F0 of the target. As with Telugu, this contour could suggest an H* pitch accent for AE. However, the decline in contour after mid-F0 could suggest an H*+L accent. Note that this low trailing accent may be associated with the declarative boundary tone in American English and therefore may independent of the preceding tone (Ladd, 1996).

Summary

Admittedly, these findings are quite tentative. Further research can determine the F0 contour which follows the focused item in order to develop a complete and reliable pitch accent model. Also, the presence and influence of edge tones on the pitch contour need

to be considered in future analyses. For example, an intermediate phrase may intervene between the target and the adverbial *now*, creating an additional phrase tone (the H or L tone attached to end of an intermediate phrase) at the end of the target. This additional phrase tone might alter the pitch contour over the target and confound the present - analysis of the pitch accent.

Lending support to this analysis of different contours among different English varieties is the observation that these varieties differ significantly most strongly in the measures of onsetF0 and mid-F0. Since these measures are most closely associated with pitch contour, we would expect the pitch accents associated with focus to also differ among varieties. This analysis of pitch accent is descriptively rather than statistically valid. However, the statistical differences in mid-F0 and onset-F0 among varieties does lend it some degree of statistical validity.

Summary

The Introduction listed three goals (or question) which this research sought to answer:

- 1) To describe the acoustic correlates of focus in Hindi L1 English, Telugu L1 English, and American English.
- 2) To describe the differences in focus among these dialects and to discuss the ways in which differences in focus are relevant to cross-dialectal differences in word-level stress.
- 3) To develop a partial model of pitch accent for focus, according to the principles of Autosegmental Metrical Theory (Pierrehumbet 1980).

With regard to these goals, this study found that (1) the acoustic correlates of focus of American English accord with other such studies, and it also described the acoustic cues of focus for Hindi L1 and Telugu L1 English. It found that (2) cross-dialectal differences in focus are not equivalent to cross-dialectal differences in word-level stress, which indicates that the acoustic correlates of speech at different prosodic levels are at variance.

Finally, this study found that (3) the pitch accents associated with focused items differ cross-linguistically.

_____	from _____	to _____
_____	from _____	to _____
_____	from _____	to _____
_____	from _____	to _____

What percentage do you speak *English* in these situations? *Your native language?*

ENGLISH

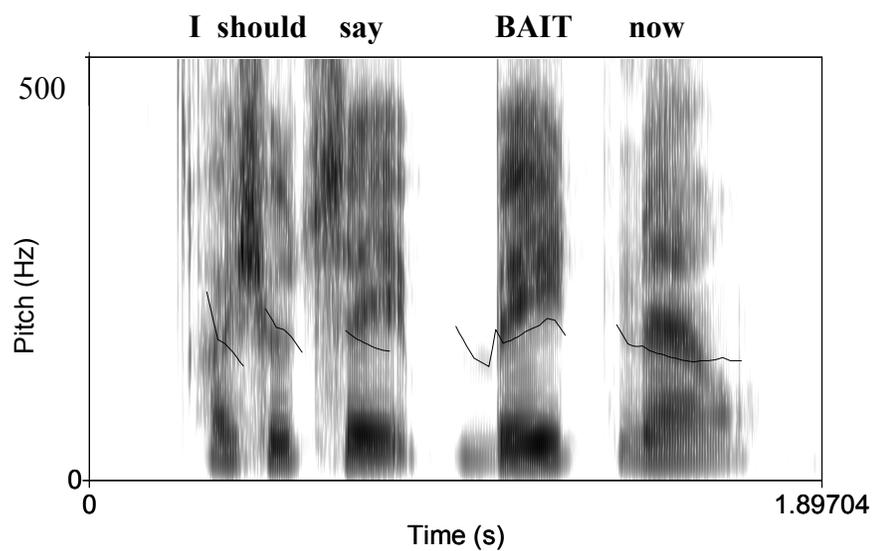
NATIVE LANGUAGES OTHER THAN ENGLISH—Please specify:

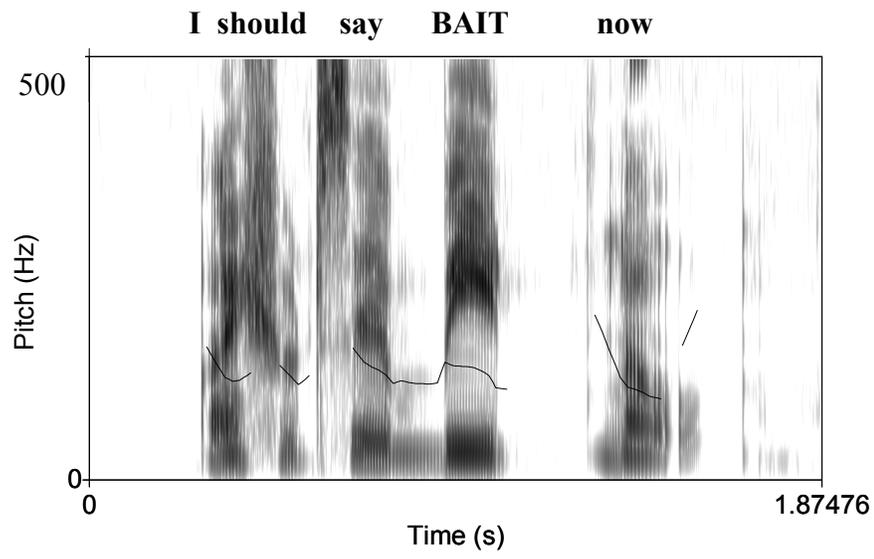
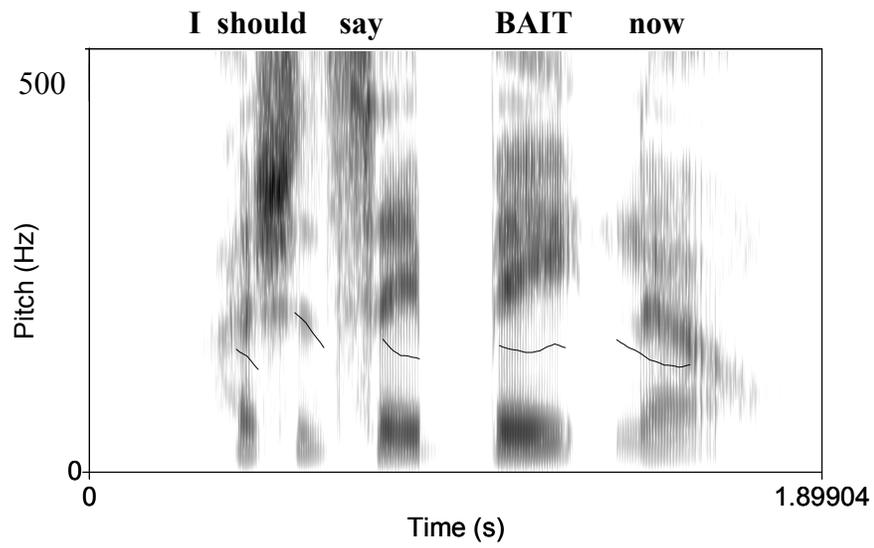
_____ %	at work/school?	_____ %
_____ %	at home?	_____ %
_____ %	with friends?	_____ %
_____ %	other?	_____ %

APPENDIX B
PITCH TRACK AND SPECTROGRAM EXAMPLES

*** The pitch track overlays the spectrogram in each diagram. Note that the units on the y-axis give the pitch range for F0 (the pitch track) but not the spectrogram.*

Hindi L1 English



Telugu L1 English**American English**

LIST OF REFERENCES

- Atkinson, J. (1973). Aspects of intonation in speech: implications from an experimental study of fundamental frequency. Unpublished PhD thesis, University of Connecticut.
- Beckman, M.E. (1986). Stress and non-stress accent. Foris Publications: Dordrecht.
- Beckman, M.E. and Pierrehumbert, J.B. (1986). Intonational structure in English and Japanese. *Phonology Yearbook* 3. 255-310.
- Berinstein, A. (1979). A cross-linguistic study on the perception and production of stress. *UCLA Working Papers of Phonetics* 47.
- Bolinger, D. (1958). A theory of pitch accent in English. *World* 14, 109-149.
- Chen, Y., Robb M. P., Gilbert, H.R., and Lerman, J.W. (2001). A study of sentence stress production in Mandarin speakers of American English. *JASA* 109 (4), 1681-1690.
- Cooper, W.E. and Sorensen, J. (1981). Fundamental frequency in sentence production. Springer-Verlag: New York.
- Cooper, W.E., Eady, S.J., and Mueller, P.R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *JASA* 77 (6), 2142-2156.
- Cooper, W.E., Soares, C., Ham, A., and Damon, K. (1983). The influence of inter- and intra-speaker tempo on fundamental frequency and palatalization. *JASA* 73, 1723-1730.
- Crystal, D. (1969). Prosodic systems and intonation in English. Cambridge University Press: Cambridge.
- Eady, S. J. and Cooper, W.E. (1986). Speech intonation and focus location in matched statements and questions. *JASA* 80 (2), 402-415.
- Eady, S.J., Cooper, W.E., Klouda, G.V., Mueller, P.R., and Lottis, D.W. (1986). Acoustical characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments. *Language and Speech* 29 (3), 233-251.
- Eefting, W. (1990). The effect of information value and accentuation on the duration of Dutch words, syllables, and segments. *JASA* 89, 412-424.

Ferreira, F. (1993). Creation of prosody during sentence production. *Psychol. Rev.* 100, 233-253.

Folkins, J.W., Miller, C.J., and Minifie, F.D. (1975). Rhythm and syllable timing in phrase level stress patterning. *J. Speech Hear. Res.* 18, 739-753.

Fry, D. (1958). Experiments in the perception of stress. *Language and Speech* 1, 126-152.

Gussenhoven, C. (1983a). Focus, mode, and the nucleus. *Journal of Linguistics* 19, 377-417.

Gussenhoven, C. (1983b). Testing the reality of focus domains. *Language and Speech* 26, 61-80.

Hayward, K. (2000). Experimental phonetics. Pearson Education Limited: Harlow, England.

Ladd, R.D. (1996). Intonational phonology. Cambridge University Press: Cambridge.

Ladefoged, P. (1996). Elements of acoustic phonetics. The University of Chicago Press: Chicago.

Lieberman, M. (1967). Intonation, perception and language. MIT Press: Cambridge.
O'Shaughnessy, D. (1979). Linguistic features in fundamental frequency patterns. *Journal of Phonetics* 7, 119-145.

Pell, M.D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *JASA* 109 (4), 1668-1680.

Pierrehumbert, Janet (1980). The phonology and phonetics of English intonation. PhD thesis, MIT.

Pike, K.L. (1945). The intonation of American English. University of Michigan Press: Ann Arbor.

Potisuk, S., Gandour, J., and Harper, M. (1996). Acoustic correlates of stress in Thai. *Phonetica* 53, 200-220.

't Hart, J., René, C., and Cohen, A. (1990). A perceptual study of intonation: an experimental-phonetic approach. Cambridge University Press: Cambridge.

Trager, G.L. and Smith, H.L. (1951). An outline of English structure. Battenburg Press: Norman, OK. Reprinted 1957 by American Council of Learned Societies, Washington.

Weismer, G., and Ingrisano, D. (1979). Phrase-level timing patterns in English: Effects of emphatic stress location and speaking rate. *J. Speech Hear. Res.* 22, 516-533.

Wells, R. (1945). The pitch phonemes of English. *Language* 21, 27-40.

Wiltshire, C.R., and Moon, R. (2000). Phonetic correlates of stress in Indian English. Paper presented at the International Conference on Stress and Rhythm. CIEFL: Hyderabad.

BIOGRAPHICAL SKETCH

Russell Moon was born in Texas and grew up in South Florida, but is still really a Texan. He attended college at the University of Florida, where he received a B.A. in English and an M.A. in linguistics. He plans to pursue a Ph.D. in linguistics at the University of Florida.

Russell is a happy soul who treasures his friends and family. He remains a staunch defender of linguistics.