PHONETIC CUES TO ACCENTUAL PROMINENCE IN BENGALI ENGLISH

Olga Maxwell & Janet Fletcher

University of Melbourne, Australia
o.maxwell@pgrad.unimelb.edu.au; janetf@unimelb.edu.au

ABSTRACT

This paper examines the acoustic correlates of accentual prominence in English spoken by L1 Bengali speakers. The acoustic phonetic correlates of stressed unaccented vowels, and stressed vowels produced in positions of narrow focus were compared. As observed for many varieties of English, the main correlates of accentual prominence in narrow focused contexts were presence of/and extent of f_0 movement throughout the stressed vowel, followed by a concomitant increase in RMS-db. Vowels were also longer compared to stressed unaccented vowels although the differences were relatively small and highly variable.

Keywords: prominence, phrasal accentuation, narrow focus, Indian English intonation

1. INTRODUCTION

Previous work on Dutch and English [2, 3, 10, 11] has established that duration, overall intensity, vowel quality and spectral balance, and f_0 are important acoustic correlates of prosodic prominence. It is important however, to distinguish two phonological levels of prominence in English, namely lexical prominence (stress) and phrasal prominence (accent). For example, f_0 movement is considered to be a more significant cue to accentual prominence than to lexical stress e.g. [2]. In an investigation of the acoustic correlates of stress and accent in American English (AE), [11] found that duration and spectral balance correlate with lexical prominence, while overall intensity correlates more with accentual prominence.

Despite a relatively large body of research on prominence in English, limited attention has been given to this aspect in new or recently emerged World English varieties. Most studies have focused on the segmental phonology of Indian English (IE) varieties (e.g. [4]). However, [13] compared the phonetic cues to lexical prominence in IE and AE and reported that IE stress is cued by an increase in f_0, duration and amplitude although the affects were less strong than in AE. In a related vein, [12] found L1 Gujarati and L1 Tamil speakers accented words more frequently in a phrase compared to AE speakers, suggesting differences in phrasal accentuation patterns. All agree e.g. [12, 13] that it is important to consider the potential influence of L1 patterns on the realization of different levels of prominence of IE varieties. The L1 of speakers examined in the current study is Bengali, which like English, is generally classified as an intonation language. Unlike English for the most part, word stress placement in Bengali is restricted to the initial syllable of a word [6]. On the other hand pitch is likely to cue post-lexical prominence [9], as in AE. Moreover, accented words have a low rising pitch pattern where the f_0 movement consists of a low f_0 valley followed by a rise. In Standard Kolkata Bengali, the initial low pitch (L*) is associated with the focused syllable and the high pitch is associated with the right edge of the phrase [6]. Recent research on Standard Bangladeshi Bengali [8] proposes that a low pitch accent (L*) followed by a high edge tone functions as a default accent, while a rising pitch accent (L*+H) with no right edge marking is used on narrow accented words.

No research to-date has examined how lexical or accentual prominence is realized in Bengali English (BengE). The current investigation seeks to redress this by examining acoustic correlates of accentual prominence in BengE. In particular, we examine the contribution of f_0 (i.e. tonal patterns) and non-tonal cues such as duration and intensity to the realization of accentual prominence and lexical stress.

2. METHOD

2.1. Speakers

Four male speakers of Bengali from West Bengal, India, were recorded in Melbourne, Australia. The speakers had completed their University degrees in

India, and had started learning English at the age of 6-7. All were professional medical doctors and spoke English fluently.

2.2. Materials and procedure

A corpus of declarative utterances was designed to elicit different patterns of accentual prominence placement and/or realization that would normally indicate broad or narrow focus structures in many varieties of English e.g. [9]. Two variables were manipulated for each utterance: a) the length of the target words in each utterance for both narrow and broad conditions and b) the location of semantic focus for utterances in the narrow focus condition. The structure of the target words varied from monosyllabic to tri-syllabic.

In order to elicit a particular focus condition, each set of materials consisted of four prompt questions and four answers [1, 14]. An example of the question-answer set is given below. The first prompt question in bold font was designed to elicit narrow focus on the verb. The second question was designed to produce broad focus across the entire utterance.

Prompt questions
What may Lee do? / What did you say?
Answer
Lee may ***move/borrow/minimize*** my mill.

The speakers produced five repetitions of each utterance type. They were asked to speak at their normal speaking rate. For the purpose of this paper, target words *move, borrow* and *minimize* in the utterances with sentence-medial focus and broad focus conditions were analyzed, which came to a total of 422 tokens.

2.3. Annotation criteria and analysis

The digitized files were annotated and analyzed using the EMU speech database system [5]. Annotations included the target word, syllables, onset and offset of the consonant and vowel of all syllables of the target words. Syllables of the target words were also identified as stressed or unstressed on the bases of an auditory analysis performed by the first author. A further analysis of each utterance was performed to ensure that speakers produced the two different focus structure patterns. A narrow focus pattern was noted if the target word received a prominence-lending pitch movement. Tokens were analyzed as stressed and unaccented if they did not receive a prominence-lending pitch movement. The f_0 contour was used to perform an intonational annotation of the target words. Where there was an obvious f_0 rise in and around the target word, L tone targets were annotated at the local pitch minimum or elbow (min f_0), which was usually in the vicinity of the lexically stressed vowel. H tone targets were annotated at the local pitch maximum (max f_0) at the end of the rise. These syllables were analyzed as "accented". In the remaining utterances if there was no identifiable pitch movement on the target word, and the initial and/or final word in the utterances had prominence-lending pitch movements, the target words were assumed to be stressed and "unaccented".

2.4. Measurements

The following measurements were extracted using Emu/R [5, 7]: - vowel duration (ms) in syllables that were identified as stressed unaccented, and stressed accented in narrow focus utterances; dB-RMS values extracted at the stressed/accented vowel midpoint. The f_0 trajectory across all stressed syllables was extracted from the onset to offset of the main stressed vowel. For narrow focus tokens, this was measured by subtracting the min f_0 (i.e. L tone target) from max f_0 (i.e. H tone target). Subsequent statistical analyses were performed using the 'R' statistical package [7].

3. RESULTS

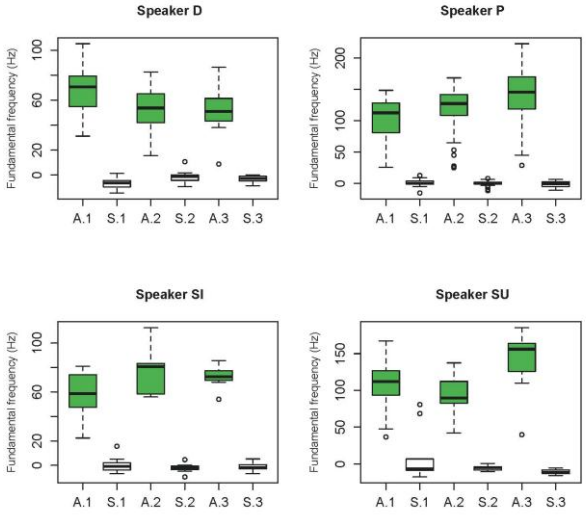
3.1. Pitch

Figure 1 shows box plots of mean f_0 (Hz) trajectories through the narrow focus accented vowels and non-focal stressed vowels produced by each speaker separated by word. All speakers produced significant differences in pitch trajectories through the narrow focus stressed vowel compared to the non-focal contexts (D:$F(1,70)=360.123$ $p<0.05$; P:$F(1,196)=732.64$, $p<0.05$; SI:$F(1,86)=1009$, $p<0.05$; SU:$F(1,66)=278.41$, $p<0.05$). In other words, the results confirmed our impressionistic observations that stressed unaccented vowels showed little pitch variation compared to accented vowels.

With the exception of speaker D, word identity also affected the extent of pitch trajectory (P:$F(2,196)=8.7$, $p<0.05$; SI:$F(2,86)=5.66$, $p<0.05$; SU:$F(2,66)=5.8$, $p<0.05$). The longest token ***mi****nimize* showed the greatest differences. This also relates to the strong significant interaction between accent and word, observed for all speakers (D:

F(2,70)=4.68, p<0.01; P:F(2,196)=7.6 p<0.05; SI:F(2,86)=5.65,p<0.05;SU:F(2,66)=5.82, p<0.05). For example, speakers P and SU produced greater pitch rises on ***mi**nimize* compared to *b**o**rrow* or *m**o**ve*. Speaker D, by contrast produced a wider accentual rise on *m**o**ve*, compared to the other tokens.

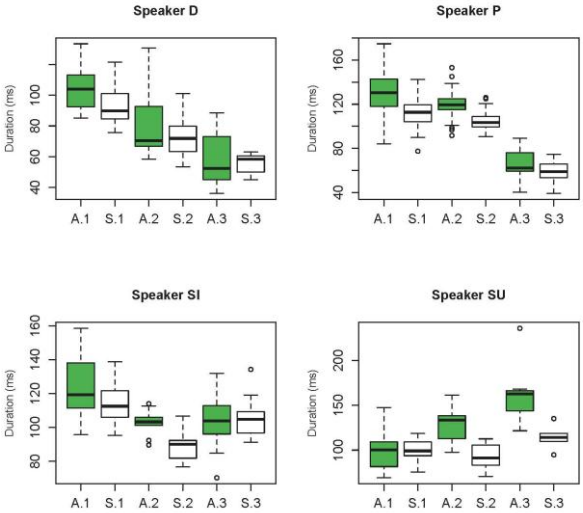
Figure 1: Mean f_0 (Hz) and confidence intervals of pitch trajectory in accent (A-green) and stress (S-white) conditions of the target words *move* (1), *borrow* (2) and *minimize* (3) for each speaker.



3.2. Duration

Figure 2 plots mean vowel duration for each speaker broken down by target word, accent/stress condition, and speaker. In spite of a high level of variation overall, accented stressed vowels were significantly longer than stressed unaccented vowels for three of the four speakers (P:F(1,195)=86.61, p<0.05; SI:F(2,96)=4.22, p<0.05; SU:F(2,90)=32.11, p<0.05). For example, the mean value of the accented initial vowel in *b**o**rrow* for speaker P was 119 ms (sd=12 ms), compared to 105 ms (sd=9 ms) for stressed unaccented tokens. While there was no main effect of accent on target vowel duration for speaker D (p>0.05), there was however a main effect of word which was also apparent for all speakers (D:F(2,71)=52.22, p<0.05; P:F(2,195)=349.97, p<0.05;SI:F(2,96)=25.52,p<0.05;SU:F(2,90)=46.76, p<0.05). Stressed and accented vowel durations were longer in the monosyllabic target word *m**o**ve* compared to trisyllabic *mi**n**imize* for three of the four speakers (D, SI & P). For example, mean accented duration *m**o**ve* was 130ms (sd=19ms) compared to 65m (sd=12ms) for accented /ɪ/ in *mi**n**imize.* Conversely, target vowel duration was longer in *mi**n**imize* compared to *m**o**ve* for speaker SU.

Figure 2: Mean vowel duration (ms) and confidence intervals in accented (A-green) and stressed (S-white) vowels of the target words *move* (1), *borrow* (2) and *minimize* (3) for each speaker.

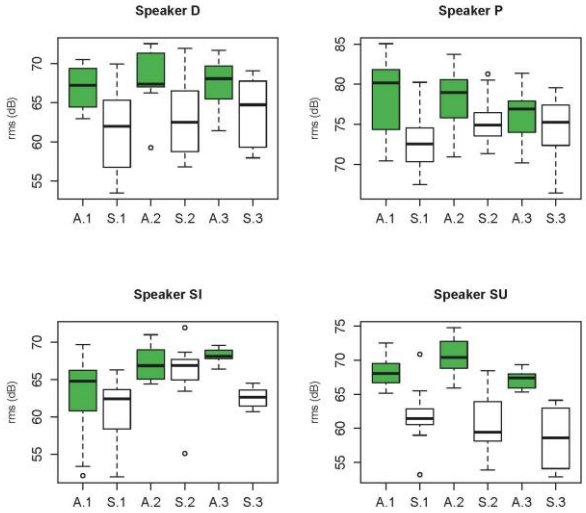


A significant interaction between accent and word was also observed for two speakers (P: F(2,195)=3.33, p<0.05; SU:F(2,90)=5.57, p<0.05). Accented vowels were longer than stressed vowels for *m**o**ve* and *b**o**rrow* for speaker P, but there was no significant difference for *mi**n**imize*. Conversely, for speaker SU, vowels in accented *b**o**rrow* and *mi**n**imize* were longer than in stressed tokens but there was no significant difference between accented and stressed *m**o**ve* tokens. For speaker D, there was only a significant effect of accent in *move* tokens.

3.3. Intensity

Figure 3 shows mean intensity (dB-RMS) of stressed and accented vowels broken down by target word. Accented vowels had higher dB-RMS values than stressed unaccented vowels (D: F(1,71)=33.96, p<0.05, P: F(1,195)=56.84, p<0.05, SI:F(1,85)=17.79, p<0.05, SU:F(1,66)=102.98, p<0.05), although there was a high level of variability across tokens and speakers. There was a strong effect of word for speaker SI (F(2,85)=16.12, p<0.05. For two speakers the effect was weak, but statistically significant (P:F(2,195)=5.99, SU: F(2,66)=3.84, p<0.05). The interaction between accent and word was only significant for two speakers (P:F(2,195)=5.99, p<0.05,SI:F(2,85)=3.85, p<0.05).

Figure 3: Mean intensity and confidence intervals (in dB-RMS) for accented (A-green) and stressed (S-white) vowels of the target words *move* (1), *borrow* (2) and *minimize* (3) for each speaker.



4. DISCUSSION AND CONCLUSION

In this study of BengE, the strongest and most consistent cue to accentual prominence in narrow focused words was presence of, and the extent of pitch movement through the accented token, with a concomitant increase in db-RMS. In addition, the analyses show a degree of variation across the speakers, e.g. the use of pitch range, overall intensity values, realization of target tokens.

The contrast between stressed accented and stressed unaccented vowels was also marked by differences in acoustic duration; in line with previous research on AE, Dutch [11], and German [1], durational differences were smaller and somewhat more variable in this corpus. Moreover, for one of the speakers (D), there was no significant increase in stressed accented vowel duration suggesting that this speaker relies on f_0 and overall intensity as primary cues for marking accentual prominence. Predictable durational variation due to word length was observed for at least three of the four speakers (P, SI and SU), showing longer duration of a particular constituent in smaller focus domain [1]. However, the effects of accent, word length, and the interaction between the two factors were significant only for two speakers (SI and SU). It is possible that vowel duration is less reliable in marking differences between lexical stress and accentual prominence in BengE compared to AE or Dutch [11]. Future research will examine the duration of unstressed vowels in both accented and unaccented words. Vowel quality is also one of the key acoustic correlates of lexical prominence in English [2, 3, 11], and this needs to be examined.

In terms of potential L1 influence on accent realization, accented words produced by these BengE speakers, consistently bore a rising pitch movement, similar to an accentual pattern of *f0* in Bengali [6, 8]. It is also similar to L*+H pitch accent observed for IE speakers of Tamil and Gujarati L1 backgrounds [12]. One issue not examined here is the complex interaction between accentuation and phrasing mentioned in the introduction. When placing narrow focus on the verb, speakers occasionally inserted a phrase break after the focused word in a number of repetitions, suggesting a potential role for phrasing in the realization of a focus domain.

5. REFERENCES

[1] Baumann, S., Becker, J. Grice, M., Mucke, D. 2007. Tonal and articulatory marking of focus in German. *Proc. 16th ICPhS* Saarbrucken, 1029-1032.

[2] Beckman, M. 1986. *Stress and Non-stress Accent*. Dordrecht: Foris.

[3] Campbell, N., Beckman, M. 1997. Stress, prominence and spectral balance. In Botinis, A. et al (eds.), *Intonation: Theory, Models and Applications*. ESCA and University of Athens, 67-70.

[4] Gargesh, R. 2004. Indian English: Phonology. In Edgar, W., Burridge, K., Kortmann, B., Mesthrie, R., Upton, C. (eds.), *A Handbook of Varieties of English, Vol 1: Phonology*. Berlin: Mouton de Gruyter, 993-1002.

[5] Harrington, J. 2010. *The Phonetic Analysis of Speech Corpora*. Blackwell Publishing.

[6] Hayes, B., Lahiri, A. 1991. Bengali intonational phonology. *NLLT* 9, 47-96.

[7] Ihaka, R., Gentleman, R. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5(3), 299–314.

[8] Khan, S.D. 2008. *Intonational Phonology and Focus Prosody in Bengali*. PhD dissertation, UCLA.

[9] Ladd, D.R. 2008. *Intonational Phonology*. CUP.

[10] Slujter, A.M.C., van Heuven, V. 1996a. Spectral balance as an acoustic correlate of linguistic stress. *JASA* 100(4), 2471-2485.

[11] Slujter, A.M.C., van Heuven, V. 1996b. Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of ICSLP* 96, 630-633.

[12] Wiltshire, C., Harnsberger, J. 2006. The influence of Gujarati and Tamil L1s on Indian English: A preliminary study. *World Englishes* 25(1), 91-104.

[13] Wiltshire, C., Moon, R. 2003. Phonetic stress in Indian versus American English. *World Englishes* 22, 291-303.

[14] Xu, Y., Xu, C. 2005. Phonetic realization of focus in English declarative intonation. *J. of Phon* 33,159-97.