Marking of Focus in Indian English of L1 Bengali Speakers

Olga Maxwell\(^1\)

\(^1\) School of Languages and Linguistics, University of Melbourne, Australia

\(\text{o.maxwell@pgrad.unimelb.edu.au}\)

\section*{Abstract}

A production experiment was designed to examine the effect of narrow focus structure on the intonational patterns of \(f_0\) and duration in English spoken by L1 Bengali speakers of English. The results show significantly higher pitch excursion on the accented words and a small increase in the duration of the accented syllable in narrow versus broad focus structures. In addition, the shape of the low rising \(f_0\) pattern used on narrow focused words is similar to the pattern observed in Bengali and could potentially be an additional cue to focus marking for English speakers of this L1 background.

\textbf{Index Terms:} intonational phonology, Indian English, focus, rising pitch

\section*{1. Introduction \& background}

Focus is known to have a prosodic reflection in many languages \([1]\). It is traditionally linked to relative prominence and prosodic phrasing in an utterance \([2]\); however, different languages rely on different prosodic means in the expression of focus and may vary in the way the focus domain is marked. Some of these means include pitch accent type \([1, 3, 4]\), pitch excursion \([5]\), post-focal pitch compression \([6, 7]\), change in phrasing \([8]\) and lengthening of the focus domain \([9, 19]\).

To-date, a substantial amount of research has been conducted on the realisation of focus in English and its varieties such as American English (AmE), British English (BE) and Australian English (AusE), where the intonational realisation of narrow focus structures is marked by the choice of relative prominence placement, lengthening of the focus domain \([9]\), an increase in the height of \(f_0\) peak and de-accenting \([9, 5, 6]\). Little attention has been given to the marking of focus in new and emerging varieties of English, while recent studies on other prosodic aspects reported an influence of first language on the prosodic structure of these varieties (i.e. Hong Kong English \([10]\) and Nigerian English \([11]\)). The current investigation attempts to fill this gap and examines the tonal and durational marking of focus in Indian English (IE) of L1 Bengali speakers, namely the pitch shape associated with the focal constituent, the effect of focus on the \(f_0\) peak scaling and pitch span, and segmental duration.

Despite an ongoing interest in the phonology of IE \([12]\), there has been no systematic research into its prosodic structure and the prosody of focus in particular. A preliminary account of intonation patterns in IE speakers of Tamil and Gujarati L1 backgrounds \([13]\) found that the utterances produced by the speakers had a greater number of accented words in a phrase as compared to AmE or BE, suggesting potential difference in the use of phrasal prominence and focus projection. The results also revealed that both L1 groups used rising pitch, analysed as a rising pitch accent (L\(^*_0\)+H). Tamil L1 speakers, in addition to a rising pitch pattern, applied high pitch and falling pitch to the accented words. The preference for using rising pitch on the accented words resembles the pattern recorded in the intonational phonology of a number of languages spoken in the subcontinent (e.g. Tamil \([14]\), Hindi \([7, 15]\)) including Bengali \([16, 4, 17]\), where the \(f_0\) movement on an accented word consists of a low \(f_0\) valley followed by a rise.

The phonological representation of the rise is treated differently in the two main models of Bengali. Using an autosegmental-metrical framework for intonation in Standard Kolkata Bengali (Western Bengali) \([16, 18]\) posit that a focused constituent forms a single phonological phrase, where rising pitch is treated as a low pitch accent (L\(^*_0\)) corresponding to the focus nucleus and a high tone (H) linked to the right edge of the phrase. A similar rising pattern is used in broad focus structures, but under narrow focus, the scaling of tones is more exaggerated with lower L and higher H targets \([16]\). For Standard Bangladeshi Bengali (Eastern Bengali) \([4, 17]\) propose that a low pitch accent (L\(^*_0\)) followed by a high edge tone functions as a default for broad focus structures, while a rising pitch accent (L\(^*_0\)+H) with no right edge marking, followed by post-focal deaccenting, are used as means for focus marking. In addition, the \(f_0\) peak of the bitonal pitch accent is realised higher than the H peak of the phrase tone. The low rising or “scooped” pitch accent (L\(^*_0\)+H) is also included in the inventory of pitch accents for English but is usually associated with a particular contextual meaning \([2]\). Another important difference is that in English and German, a range of pitch accent shapes can be used to indicate narrow as well as broad focus \([9, 6, 19]\) whereas in Bengali, tonal realisation of focus is restricted to a low rising pitch pattern on the focused constituent. However, narrow focus in all three languages expands overall \(f_0\) height or pitch span within the focal constituent.

\subsection*{1.1 Aims of the study}

This investigation is part of a larger study examining the realisation of focus in Indian English of L1 Bengali speakers. The following paper will concentrate on tonal and durational marking of focus and look at any evidence of possible L1 prosodic transfer. The aspects under investigation are the shape and height of \(f_0\) on the accented words in the two focus structures: broad focus and narrow focus.

Research questions:
1. Will there be a greater pitch excursion on a narrow focus constituent?
2. Will there be an increase in the duration of segmental domain for narrow focus constituents?
3. Will the speakers rely on a particular pitch shape to mark narrow focus?

\section*{2. Method}

\subsection*{2.1. Speakers}

Four male speakers of Bengali from West Bengal, India, were recorded in Melbourne, Australia. Four speakers were from Kolkata and one speaker – from a smaller town in close proximity to Kolkata. The speakers were aged between 37-55 years of age at the time of recording, had completed their
degree in India, had English as the medium of instruction at school and higher educational institutions, and started learning English at the age of 6-7. All speakers were international medical graduates practising psychiatry in Australia and spoke English fluently.

2.2. Materials

A number of simple declarative utterances were designed to elicit broad and narrow focus structures. Sentence length was kept constant for each utterance and had the same syntactic structure [6, 19]. 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 focus for utterances in the narrow focus condition. The structure of the target words varied from monosyllabic to tri-syllabic. Each target utterance consisted of a subject, a proper noun, followed by a verb with the preceding pronoun ‘may’, followed by an object with the preceding pronoun ‘my’. Consistent with previous research [6, 1] target words contained sonorant consonants (m, n, l, r) in the onset position and with no preceding coda to avoid gaps in the visible \( f_0 \) contour due to voiceless segments.

In order to elicit a particular focus condition for each utterance, the stimuli where designed as question-answer pair sets. This method has been used successfully in previous studies [6, 9, 19]. Each set consisted of four prompt questions and four answers where the target word would vary depending on the focus structure and was marked in bold for the narrow focus reading. An example of question-answer set is given below.

<table>
<thead>
<tr>
<th>Prompt question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Who may move my mill?</td>
<td>1. Lee may move my mill. N</td>
</tr>
<tr>
<td>2. What may Lee do?</td>
<td>2. Lee may move my mill. N</td>
</tr>
</tbody>
</table>

The speakers were asked to speak at their normal speaking rate, as if they were talking to a friend. After recording, any sentences containing any type of reading error such as unexpected phrasing break, disfluencies, etc, were eliminated or reclassified. For the purpose of this paper only the utterances with broad focus and sentence-initial narrow focus have been analysed, in total 360 utterances, 15 tokens for each target word (Lee, Nina, Marina).

2.3. Annotation criteria and analysis

The digitised files were annotated and analysed using EMU speech database system [20]. Labels included the target word, syllables, onset and offset of the consonant and vowel in both accented and post-accented syllables, and tone targets. Statistical analyses were performed using the ‘R’ statistical package [21].

Summary of acoustic measurements:

To analyse pitch shape on the focused word, the following annotation criteria were used:

L - the \( f_0 \) valley defined as the local minimum (min \( f_0 \)), taken in the vicinity of the stressed syllable before the rise in the \( f_0 \) contour;

H - the \( f_0 \) peak defined as the local maximum (max \( f_0 \)), taken at the highest point reached in the vicinity of the accented syllable immediately before the fall in the \( f_0 \) contour;

Syllable duration - duration of the accented syllable in the accented word.

3. Preliminary Results

3.1. Pitch shape

Table 1 summarises the types of pitch shapes used by each speaker on the accented words in narrow and broad focus structures. As the middle column shows, the speakers are consistent in using a rising pitch contour on the focused word in narrow focus. The L target is aligned just before or on the accented syllable, while the alignment of the H is often moved to the post-accented syllable in bi- and tri-syllabic words. In broad focus structure (third column), the data show some variation across the speakers and repetitions. Speaker B4 apply a rise in pitch to the accented word, while speakers B1, B2 and B3 are less consistent and, in addition to rising pitch, apply a sustained low or high pitch to the accented words. For speakers B2 and B3 low pitch is used more frequently, especially in bi- and tri-syllabic target words.

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Narrow Focus</th>
<th>Broad Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>LH</td>
<td>LH, L</td>
</tr>
<tr>
<td>B2</td>
<td>LH</td>
<td>LH, L, (H)</td>
</tr>
<tr>
<td>B3</td>
<td>LH</td>
<td>LH, L, (H)</td>
</tr>
<tr>
<td>B4</td>
<td>LH</td>
<td>LH</td>
</tr>
</tbody>
</table>

Table 1. Pitch shape in the two focus structures.

Figure 1. \( f_0 \) track from the acoustic onset of the word Marina in the two focus structures (narrow – top line, broad – bottom line) for speaker B2, time normalised and averaged, averaged consonant onset timing of the accented and post-accented syllables in the two focus structures superimposed.

For those repetitions where rising pitch is also applied to the accented target words in broad focus, the data suggest possible differences in pitch shape depending on the focus structure. Figure 1 shows an example of differences in the phonetic detail in the rising \( f_0 \) contour on the target word Marina in two focus structures for Speaker B2. Broad focus is marked with a top line, narrow focus – with a bottom line. The observed pattern is representative of all speakers and shows that in narrow focus the H target aligns earlier in relation to the accented syllable and has a higher \( f_0 \). In addition, a number of narrow focused utterances show a phrase break or a pitch reset after the focused word, which may indicate the effect of focus on phrasing.

3.2. Pitch range

Figure 2 shows scaling difference between the L and H targets in the target words in two focus structures for each speaker. Grey-coloured boxes correspond to the words in broad focus,
while white-coloured boxes correspond to the words in narrow focus structure. The words are presented in pairs, starting with a monosyllabic word Lee. A number of ANOVA analyses were performed on three word pairs in the two focus conditions for speakers B1, B3 and B4, and a paired t-test on one pair for Speaker B2. The ANOVA confirmed a significant difference in the scaling between L and H targets in the two focus structures across the speakers (speaker B1 – $F(2,20)=11.89$, $p<0.002$, speaker B3 – $F(2,14)=8.66$, $p<0.01$, speaker B4 – $F(2,14)=10.2$, $p<0.0001$).

Figure 2. Box plots showing the confidence interval in mean and standard deviation of the scaling difference between L and H targets (Hz) for each speaker in relation to the two focus structures.

For speaker B2, a significant difference was recorded in the scaling of the rise for the word Lee in two focus structures ($p<0.01$).

### 3.3. Duration

Figure 3 plots the mean duration of the accented syllable in each word in the two focus structures averaged across all speakers. Target words in narrow focus structure are marked with grey colour. The analysis shows differences in the mean duration values of the accented syllable between the narrow and broad focus structures ($F(3,114)=18.05$, $p<0.0004$). The monosyllabic target word Lee has a greater durational increase under narrow focus as compared to the other two target words. One speaker (B4) significantly ($F(2,24)=35.28$, $p<0.0001$) lengthens the accented syllable in focused words, with a mean duration of 288 ms (Lee), 271 ms (Nina), and 224 ms (Marina) in narrow focus.

It is important to take into consideration that the tokens compared are accented, which explains why the actual differences are small. Further comparison is needed to examine segmental duration in accented versus non-accented words.

Figure 3. Differences in mean and standard deviation of the accented syllable duration averaged across speakers and repetitions in the two focus structures (ms).
4. Discussion

The L1 Bengali speakers of English in this study realise focus using intonational patterns of \( f \) and duration similar to speakers of well-described varieties of English. Despite observed individual differences in the use of pitch range, all speakers manipulate \( f \) height in the two focus structures and produce higher peaks on the narrow focused words. Similar patterns are observed in other varieties of English [9, 5, 6] as well as Bengali [16, 4].

Duration is another correlate of narrow focus structure for all speakers, which has also been observed in other varieties of English [3]. Speaker B4, who uses similar patterns of rising pitch in narrow and broad structures, relies on lengthening accented syllables to realise the focus contrast, whereas for the remaining speakers, durational differences are relatively small. In line with previous research [9], the amount of durational increase on the accented syllable of the focused word correlates with the overall length of the word. Thus, longer words show a proportionally smaller increase in duration than the molten target word. Further durational movement of target words in sentence-medial and sentence-final position is required to determine the details of duration of the segmental domain, especially due to the possible final lengthening [9, 3].

While duration and pitch excursion are used similarly to patterns observed in other English varieties, this investigation reveals a difference in the expression of focus by L1 speakers of Bengali. Typically, English allows for more variation in the use of pitch accents and tonal configurations in both broad and narrow focus structures. The speakers in this study always use rising pitch on the accented word in narrow focus structures. The pattern observed is similar to the pitch pattern used in Bengali [4, 16, 18] and has previously been recorded for English spoken by Gujarati and Tamil L1 backgrounds [13].

In broad focus structure, on the contrary, the data show a degree of variation in the pitch shape across the target words. Perhaps the speakers differentiate the tonal shape between narrow focus and broad focus structures by applying a sustained high or low pitch to the accented words in broad focus structures. Two speakers (B2, B3) use low or, less frequently, high pitch in broad focus structure, while one of the speakers (B1) uses low pitch. In Bengali, the location of a high pitch accent (H*) is restricted to the final position in the prosodic phrase [16] or is used to mark sarcastic or unexpected information [4].

Preliminary analyses suggest that L1 Bengali speakers rely on a particular pitch shape (LH) as means of marking narrow focus, which in English could be interpreted as L+H* or L*+H. In addition, it is unclear at this stage whether there are one [16] or two [4, 17] phonological pitch accent categories for narrow and broad focus structures. Further analysis of tokens in different prosodic conditions and an investigation of the alignment of the tonal targets to the segmental material will help test the hypotheses and determine the tonal categories or categories of the rising pitch pattern on the accented words in different focus structures. In addition, the effect of focus on phrasing needs to be examined. The results show that in a number of repetitions, the speakers use either a pitch reset or a pause after the accented word indicating that the focused word may form a separate prosodic constituent.

5. Conclusion

To conclude, preliminary results present some interesting aspects of the intonational realisation of focus by L1 Bengali speakers of English. Typically for the expression of focus in English, the speakers use lengthening of the accented syllable and a greater pitch excursion in narrow focus structures. In addition to this, the speakers also seem to rely on a particular rising pitch shape to mark narrow focus structures.

5. References
