

Tonal Alignment of Focal Pitch Accents in Two Varieties of Indian English

Olga Maxwell and Janet Fletcher

University of Melbourne

omaxwell@unimelb.edu.au, janetf@unimelb.edu.au

Abstract

This study examines the tonal alignment of focal rising pitch movements in two varieties of Indian English: Bengali and Kannada English. The results revealed that all speakers use a bitonal rising pitch accent to realise focal prominence. An examination of the alignment patterns of Low and High tone targets shows that a late rising L*+H accent on focally prominent words is characteristic of Bengali English speakers whereas Kannada English speakers use an L+H* accent. Consistent with the literature, the phonetic location of the tone targets is also strongly correlated with accented syllable duration and a range of other phonetic factors.

Index Terms: Indian English, intonational phonology, tonal alignment, pitch accents

1. Introduction

'Indian English' (IndE) is generally applied as a cover term to refer to the English language used by the speakers in India and the Indian diaspora around the world. IndE is considered a complex notion in the literature. On the one hand, it has developed distinct grammatical, phonological, lexical and discourse features, mutually intelligible between the speakers of IndE from various regions and L1 backgrounds, which makes this variety of English "identifiable as Indian" [1]. On the other hand, there is a substantial amount of variation, similar to other broad English varieties (i.e. American or British English). The definition of this variation is one of the ongoing debates in IndE research and could be related to the speakers' L1, language family, geographical location, socioeconomic or educational background [2,3,4].

In comparison to a vast body of research on IndE segmental phonology and, more recently, rhythm [2,4], intonation in IndE has received less attention. Earlier studies were often descriptive or based on auditory analyses, and the number of experimental studies within the Autosegmental-Metrical (AM) framework is rather small [5,6,7]. [5] and [6] reported rising or falling contours on accented words that could be described in the AM framework as pitch accents, indicating that further research was needed to determine the actual tonal categories. The findings suggest that there may be potential differences in the use of pitch accents across IndE varieties or in their phonetic realisation. [5], for example, reported a more frequent use of rises on accented words (an L*+H or an L+H* accent) in English spoken by Gujarati L1 speakers compared to the speakers whose L1 was Tamil. [7] reported an L*+H rising accent in Hindi English (HE), with a higher frequency of its use by late Hindi-English bilinguals. [6] found that the rises on focally accented words in HE showed greater lowering of f_0 at the beginning of the accented vowel compared to the rises produced by Telugu English speakers as well as the speakers of American English (AmE).

The present study adds to the ongoing research on IndE and examines the rises on focally accented words in IndE

spoken by L1 Bengali (Kolkata dialect, Indo-Aryan) and L1 Kannada (Dravidian) speakers with a view to determining whether these pitch movements are best described as either L+H* or L*+H pitch accents. One of the salient features shared across Indian languages is the use of rising pitch on accented words ([8] on Tamil; [9] on Hindi; [10] on Bengali). Focal pitch movements in these languages are modelled as an L* pitch accent followed by a high phrase tone, with pragmatic focus having little reflection on the differences in the use of tonal categories. This is with the exception of Bangladeshi Bengali [11], where the word in focus is signalled by an L*+H pitch accent. We predict therefore that there may be an L1 influence with regard to the way in which focal rises are realised in the two IndE varieties examined in this study. A second aim of this study is to test whether a prevailing hypothesis in the studies of tonal alignment, namely, the segmental anchoring hypothesis (outlined in the next section), can account for the phonetic realisation of focal rises in the two varieties.

2. Tonal alignment

Tonal alignment refers to the temporal relationship of tonal targets (L and H) with the segmental string or prosodic units [12,13]. It plays an important part in current theories of intonational phonology due to its relationship with meaning distinctions and the typology of intonational categories; and remains a complex notion in the AM framework with a number of unresolved issues [12,14]. The segmental anchoring hypothesis [15,12] assumes that the beginning and end points of the accentual f_0 movement are anchored to specific locations and that they independently align with specifiable points in the segmental string. Since its introduction, a number of studies have examined the segmental anchoring of tonal targets across a range of languages, with some of the studies even questioning the validity of the hypothesis in its strictest form [16]. Recently, the segmental anchoring hypothesis has been modified and is now accepted in a more loose form [13]. Moreover, recent experimental work on tonal alignment in typologically diverse languages has shown that the alignment of the tones, especially the timing of f_0 peaks, can be affected by a number of factors including tonal crowding, speech rate, proximity to prosodic boundaries, and syllable structure [e.g. 17,16,18,19].

Several studies investigating the tonal alignment of rising accents have demonstrated that the same phonological categories may show differences in their phonetic alignment across languages and language varieties (e.g. [15,19]). For example, [15] found that, unlike in English, the L and the H targets align beyond the stressed syllable in L*+H accents in Greek. In Spanish and Neapolitan Italian rising accents [19], both tonal targets (L and H) are located in the stressed syllable, where the L tone aligns with the onset of the stressed consonant. Differences have also been reported in the realisation of the same pitch accent categories in varieties of AmE [20] and varieties of British English [18]. The source of

phonetic variation is a debatable issue in the literature and could be explained in terms of a phonetic continuum in alignment across languages and language varieties [12].

This phonetic variation could also explain the alignment patterns of pitch accent categories in L2 intonation. [21]'s study on the intonational features in Greek L1 speakers of Dutch and in Dutch L1 speakers examined the phonetic realisation of the phonologically identical rise, common in Dutch and Greek and showed that most L2 speakers in the study were unable to produce the typical L1 f_0 peak alignment pattern in L2 intonation, suggesting that it is difficult to realise the fine phonetic detail of intonational pitch accents even when phonological categories may be relatively similar. Similarly, L2 English speakers of German showed later alignment patterns compared to English L1 speakers, possibly as a result of L1 influence on the phonetic alignment [22].

The study addresses the following related questions:

1. What are the alignment characteristics of the rising gesture as a whole and the tonal targets in relation to the prominent syllable in Bengali English (henceforth BE) and Kannada English (KE)?
2. What is the tonal category or categories used on focal words in BE and KE?
3. Are there any phonological or phonetic differences in the alignment of nuclear focal rises between BE and KE?

3. Method

3.1. Speakers

Four male speakers of L1 Bengali from West Bengal and four male speakers of L1 Kannada from Karnataka, India, were recorded in Melbourne, Australia. All had been living in Australia for 7 years on average. The speakers had completed their University degrees in India. Six speakers started learning English at the age of 3-5, and two speakers - at the age of 13.

3.2. Materials

The materials consisted of question-answer pairs designed to elicit narrow focus. Two target words were inserted in a carrier sentence with a declarative structure of various syntactic lengths. The answer with the target words in bold is shown below.

*Lulu/Maryellen memorised Manuela's
main/luminous/yellow-coloured manual.*

The target words consisted of sonorants in order to avoid any visible perturbation or errors in the f_0 tracking. It was also expected that using sonorants would minimise the differences in the alignment of the peak, reported for other consonants - for example, later alignment for fricatives in comparison to sonorants [15]. Each speaker was asked to read the sentences five times in a natural manner at normal speed. The data collection resulted in a total of 329 tokens. All speakers placed lexical stress on the first syllable in *Maryellen*.

3.3. Annotation and Analysis

The recordings were digitised in Praat [23] and labelled in the EMU Speech Database System [24]. The EMU annotation included the three tiers of word, phonetic, and target. The acoustic onset and offset of each word were marked on the word tier. The phonetic tier included segmentation of the words into corresponding sounds. The target tier was used for

the measurement points corresponding to segmental and word boundaries as well as tonal movements.

The digitised sound files were processed using EMU Speech and Signal Processing tool to create the additional .sfb and .sf0 files required for building hierarchies to allow the researchers to query the database and extract the necessary acoustic data for the relevant type of events and further manipulations of the data within the R statistical package [25] and the EMU-R interface [24].

The following acoustic landmarks were identified for each target word: L and H, syllable onset (C1, C2), vowel onset (V1), and word offset. The low and the high turning points were labelled as L and H accordingly. The f_0 landmarks were identified by eye. In general, the H target was relatively easy to locate. It was usually identified as a peak located towards the end of the accented syllable or in the post-accented syllable. The L target sometimes showed perturbation or a break in the f_0 due to creaky voice, usually at the syllable onset. In these cases, the L measuring point was labelled at the beginning of visible f_0 contour. For the measuring points corresponding to tonal targets and segmental/word boundaries, durational values in milliseconds were extracted.

3.4. Measurements

Following previous research on tonal alignment [15,18,20], several temporal distances were obtained in addition to segmental and syllable durations. These measurements are illustrated in Figure 1 below.

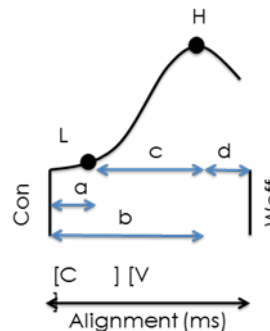


Figure 1. Schematic representation of main measurements, shown relative to the segmental landmarks

- (a) ContoL - temporal interval between the accented syllable onset (C1) and the beginning of the rise (L)
- (b) ContoH - temporal interval between the accented syllable onset (C1) and the f_0 peak (H), also referred to as *peak delay*
- (c) LtoH - the temporal interval between the L and the H targets corresponding to the duration of the rising gesture
- (d) HtoWoff - peak (H) relative to word offset (Woff)

A series of LMM analysis were performed on each interval with GROUP (BE or KE) and WORD as fixed factors, and SPEAKER as a random factor. The raw values for a, b and c were examined as a proportion of the accented syllable duration.

4. Results

4.1. L alignment

Figure 2 shows the pattern of L tone alignment for each speaker of BE (B1, B2, B3, B4) and KE (K1, K2 and so on).

It illustrates the distance from syllable onset to L as a proportion of the accented syllable duration. WORD had no effect on L alignment [$t=0.419$, $p>.05$]. Similarly, no differences in L alignment were reported based on L1 GROUP [$t=-0.536$, $p>.05$], with some variation across the speakers. As a result, the values for L alignment are presented by speaker (x-axis).

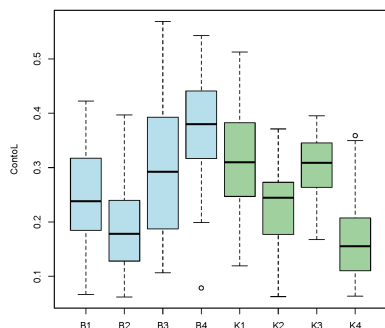


Figure 2. Box plot for the distribution of proportional mean ContoL distance, presented by speaker

Depending on the speaker, the start of the rise (L) was aligned at 0.18 to 0.38 relative to the duration of the stressed syllable, or at 20 to 60 ms interval from the syllable onset. The words produced by speaker B4 displayed the latest mean onset of the rise (L), and speaker K1 produced later rise onsets compared to K2, K3 and K4. This may have been due to the speaker-specific durational differences among the word-initial sonorants in the focally accented tokens examined here. The nasal and the lateral had the longest duration for speakers K1 and B4 who also produced later L alignment. A correlation test between ContoL timing interval and the duration of the stressed consonant confirmed a strong relationship [$R^2=0.73$; $p<.0001$]. This finding indicates that L is anchored to the initial sonorant consonant for both groups and explains the later L alignment for some of the speakers.

4.2. Peak delay

Figure 3 shows the alignment of H for each group. It illustrates the distance from syllable onset to the f_0 peak (ContoH) as a proportion of the accented syllable duration. There was no effect of WORD and a lack of interaction between GROUP and WORD. GROUP had a strong effect on the alignment of the peak [$t=3.115$, $p<.002$], showing significantly later alignment for the BE group. On average, the interval ContoH had a measured ratio of 1.38 relative to the duration of the stressed syllable [238 ms, $SD=46$] for the BE group, indicating that the H target was realised well into the post-stressed syllable rhyme. Conversely, mean peak delay for the KE group measured at 1.2 ratio of the stressed syllable [175 ms, $SD=34$ ms], and was realised in the post-stress consonant. Speaker K2 consistently produced the earliest peaks.

As shown in Figure 4, the f_0 peak (H) alignment strongly correlated with the length of the stressed syllable for both language groups [BE: $R^2=0.87$, $p<.0001$; KE: $R^2=0.87$, $p<.0001$], further supporting the looser version of the segmental anchoring hypothesis. In addition, there is ample evidence that the peak is part of a pitch accent gesture and does not demarcate a minor phrase. The distance from the

peak to word offset was affected by WORD [$t=53.42$, $p<.000$], indicating a greater distance of the peak or H tone target to word offset for *Maryellen*, with no interaction between GROUP and WORD. The number of syllables in *Maryellen* did not lead to the H tone target moving towards the end of the word edge as the word lengthened in terms of the number of syllables. L1 GROUP showed a significant effect [$t=3.057$, $p<.002$] on H alignment relative to word offset further suggesting that the two groups were realising pitch accent alignment differently in focal nuclear contexts.

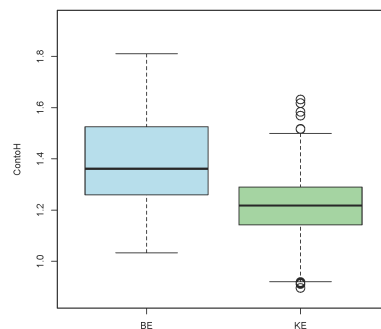


Figure 3. Box plot for the distribution of proportional mean peak delay (BE – blue, KE – green)

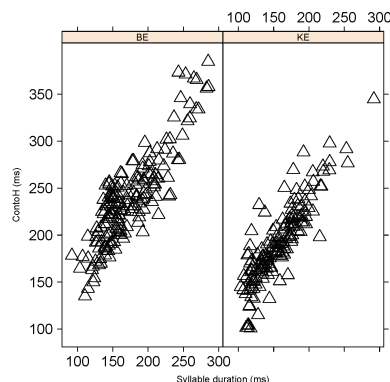


Figure 4. Scatter plot of mean peak delay and accented syllable duration, both in milliseconds (BE – left panel, KE – right panel)

5. Discussion and Conclusions

Consistent with the tonal alignment research on a variety of languages [15,16,17,18,19,22] and segmental anchoring hypothesis [15,17,12], the tonal targets aligned with certain segmental landmarks but the alignment of the peak was affected by the duration of the accented syllable. In both IndE varieties, the valleys (L tones) were sensitive to the length of the accented consonant but nonetheless showed consistent anchoring to that consonant. A stable alignment of the first tonal target in bitonal accents has been reported in cross-linguistic research [15,22,16,19]. The results on f_0 peak alignment revealed that the H tone was anchored to the post-stress vowel in BE and to the post-stress consonant in KE.

The results clearly demonstrated that the f_0 peak (H) did not align with the right edge of the focally accented word, thus excluding the possibility of an L* accent followed by a high

edge tone, reported for a number of languages spoken in India [8,9,10]. Instead, the H moved further away from the word offset as word length increased in terms of the number of post-stress syllables. All speakers produced rising pitch accents similar to the speakers of Gujarati, Tamil, Hindi and Telugu English [5,6,7], indicating that IndE intonation phonology bears a certain resemblance to the intonation of well-established varieties of English and includes rising accents in the phonological inventory.

Based on the alignment patterns of L and H tone targets, the following pitch accent categories have been posited for each variety: L+H* in KE and L*+H in BE. Despite the peak being realised beyond the boundary of the accented syllable, the L+H* pitch accent is a more suitable analysis for a focal pitch accent in KE given the consistently earlier realisation of the *f0* peak in comparison to BE. The structure of the experimental tokens may also have been a factor. Several studies have shown that the peak moves to the post-stress syllable in syllables with an open syllable structure [18,19,12]. In BE, a significantly later peak, anchored to the rhyme of the accented syllable, is interpreted as an L*+H accent similar to Hindi English [7] and Bangladeshi Bengali [11].

The L*+H accent has a marked status in many varieties of English [20] including Australian, Southern British and American English. Moreover, combined with the falling boundary tone, this tune can be associated with a particular pragmatic meaning and is often used to express greater incredulity or uncertainty [26]. Therefore, BE shows significant difference with other Englishes based on the function of the L*+H accent if this is the typical focal accent shape that is used regardless of the pragmatic context. It remains to be seen whether this could potentially have an impact on communication and influence the interpretation of an utterance in English-speaking situations outside of India.

The findings also reveal that the phonetic realisation of the two rising accents in KE and BE exhibit differences compared to other languages [15,19] and English varieties [20,26]. In the canonical interpretation of L*+H in English, the L usually aligns with the accented vowel and has a stretch of low *f0*. To the contrary, BE speakers in this study timed the L target with the onset of the accented syllable, a pattern more consistent with L alignment in L+H*.

To conclude, the examination of focal accents in these two varieties of IndE adds to the research on tonal alignment of rising accents and demonstrates important phonetic and phonological variation of these categories within IndE.

6. References

- [1] Sailaja, P. *Indian English*. Edinburgh University Press, 2009.
- [2] Sirsa, H., & Redford, M. A. "The effects of native language on Indian English sounds and timing patterns". *Journal of Phonetics*, vol. 41(6), pp. 393-406, 2013.
- [3] Wiltshire, C. The "'Indian English' of Tibeto-Burman language speakers". *English World-Wide: A Journal of Varieties of English*, vol. 26(3), pp. 275-300, 2005.
- [4] Fuchs, R. *Speech rhythm in educated Indian English and British English*. Unpublished PhD dissertation, University of Münster, 2013.
- [5] Wiltshire, C., & Harnsberger, J. "The influence of Gujarati and Tamil L1s on Indian English: A preliminary study". *World Englishes*, vol. 25(1), pp. 91-104, 2006
- [6] Moon, R. *A comparison of the acoustic correlates of focus in Indian English and American English*. Unpublished Master's thesis, University of Florida, 2002.
- [7] Puri, V. *Intonation in Indian English and Hindi late and simultaneous bilinguals*. Unpublished PhD dissertation, University of Illinois, 2013.
- [8] Keane, E. "Tamil intonation". In S.-A. Jun [Ed], *Prosodic typology 2: The phonology of intonation and phrasing*. Oxford, UK: Oxford University Press, 2014, pp.118-153.
- [9] Patil, U., Kentner, G., Gollard, A., Kuhler, F., Féry, C., & Vasishth, S. "Focus, word order, and intonation in Hindi". *Journal of South Asian Linguistics*, vol. 1(1), pp. 1-21, 2008.
- [10] Hayes, B., & Lahiri, A. "Bengali intonational phonology". *Natural Language & Linguistic Theory*, vol. 9(1), pp. 47-96, 1991.
- [11] Khan, S. D. "The intonational phonology of Bangladeshi Standard Bengali". In S.-A. Jun [Ed], *Prosodic typology 2: The phonology of intonation and phrasing*. Oxford, UK: Oxford University Press, 2014, pp. 81-117.
- [12] Ladd, D. R. "Segmental anchoring of pitch movements: autosegmental association or gestural coordination?". *Italian Journal of Linguistics*, vol. 18(1), pp. 19-38, 2006.
- [13] D'Imperio, M. "Tonal alignment". In A. C. Cohn, C. Fougeron, & M. K. Huffman [Eds] *Oxford Handbook of Laboratory Phonology*. Oxford, UK: Oxford University Press, 2012, pp. 275-287.
- [14] Prieto, P. "Tonal alignment". In M. Van Oostendorp, C. Ewen, B. Hume, & K. Rice [Eds], *The Blackwell Companion to Phonology*. Oxford, UK: Wiley-Blackwell, 2011, pp. 1185-1203.
- [15] Arvaniti, A., Ladd, D.R., & Mennen, I. "Stability of tonal alignment: The case of Greek prenuclear accents". *Journal of Phonetics*, vol. 26, pp. 3-25, 1998.
- [16] Prieto, P., & Torreira, F. "The segmental anchoring hypothesis revisited. Syllable structure and speech rate effects on peak timing in Spanish". *Journal of Phonetics*, vol. 35, pp. 473-500, 2007.
- [17] Ladd, D. R., Mennen, I., & Schepman, A. "Phonological conditioning of peak alignment in rising pitch accents in Dutch". *Journal of the Acoustical Society of America*, vol. 107, pp. 2685-2696, 2000.
- [18] Ladd, D.R., Schepman, A., White, L., Quarmby L., & Stackhouse, R. "Structural and dialectal effects on pitch peak alignment in two varieties of British English". *Journal of Phonetics*, vol. 37, pp. 145-161, 2009.
- [19] Prieto, P., D'Imperio, M., & Gili-Fivela, B. "Pitch accent alignment in Romance: Primary and secondary associations with metrical structure". *Language and Speech*, vol. 48(4), pp. 359-396, 2006.
- [20] Arvaniti, A., & Garding, G. "Dialectal variation in the rising accents of American English". *Laboratory Phonology*, vol. 9, pp. 547-576, 2007.
- [21] Mennen, I. "Bi-directional interference in the intonation of Dutch speakers of Greek". *Journal of Phonetics*, vol. 32, pp. 543-563, 2004.
- [22] Atterer, M., & Ladd, D. R. "On the phonetics and phonology of 'segmental anchoring' of F0: Evidence from German". *Journal of Phonetics*, vol. 32, pp. 177-197, 2004.
- [23] Boersma, P., & Weenink, D. *Praat: Doing phonetics by computer* (Version 5.1.35), 1992-2010. Available from <http://www.praat.org/>
- [24] Harrington, J. *The phonetic analysis of speech corpora*. Oxford, UK: Wiley-Blackwell, 2010.
- [25] Ihaka, R., & Gentleman, R. "R: A language for data analysis and graphics". *Journal of Computational and Graphical Statistics*, vol. 5(3), pp. 299-314, 1996.
- [26] Pierrehumbert, J., & Hirschberg, J. "The meaning of intonational contours in the interpretation of discourse". In P. R. Cohen, J. Morgan, & M. Pollack [Eds], *Intentions in communication*. Cambridge, MA: MIT Press, 1990, pp. 271-311.