

THE PLACEMENT AND ACOUSTIC REALISATION OF PRIMARY AND SECONDARY STRESS IN INDIAN ENGLISH

Robert Fuchs and Olga Maxwell

University of Münster and University of Melbourne
robert.fuchs@uni-muenster.de, omaxwell@unimelb.edu.au

ABSTRACT

This study examined the acoustic correlates of primary and secondary stress in Indian English. Together with the patterns of lexical stress placement, the parameters of syllable duration, pitch slope, intensity and spectral balance were examined in six noun-verb pairs. Two L1 backgrounds (Hindi and Malayalam) were examined. Results showed that lexical stress placement varied substantially across the speakers, but was in the majority of cases on the same syllable as in American or British English. Second, speakers relied on (in order of importance) differences in intensity, spectral balance, duration, and pitch slope to distinguish primary from secondary stress. The results also showed that Indian English differs from other varieties in the phonetic realisation of the primary-secondary stress distinction.

Keywords: Indian English, prominence, acoustic correlates, primary stress, secondary stress

1. INTRODUCTION

Lexical prominence (stress) is realised in English with a range of acoustic cues. Previous research on American (AmE) and British English (BrE) suggests duration, intensity, vowel quality and spectral balance to be the most salient features [2, 12, 13, 23, 27]. Some also consider f_0 to be an acoustic correlate of lexical stress, but others of primarily accentual or post-lexical prominence [27, 28]. In order to disentangle the different levels of lexical prominence, syllables with primary stress should not only be contrasted with unstressed syllables, but also with syllables with secondary stress [23].

Some postcolonial varieties, such as Singapore English, have been shown to differ from AmE and BrE in the placement and realisation of lexical stress [15], and a similar debate surrounds Indian English (IndE). Around 4% of the population of India (around 50 million) speak English fluently [5].

Lexical stress in Indian English (IndE) has received a lot of attention in the literature but mostly in relation to the placement of stress, which was found

to be different from AmE and BrE as well as variable across IndE speakers. A handful of studies have looked at the acoustic characteristics of lexical stress in IndE [22, 29, 30]. Earlier research reported that lexical stress is cued by a distinct increase in syllable duration [29]. By contrast, [22] reported lower f_0 values in stressed syllables compared to unstressed syllables (both in accented words), but no differences in intensity. On this basis, the authors posited that IndE is similar to Japanese in that it uses f_0 but not amplitude as a cue to lexical prominence. According to this view, IndE would be a pitch accent or non-stress accent language [3], where stressed syllables are marked by low pitch.

[30] questioned these findings, suggesting they could have been due to the misinterpretation of the position of lexical prominence in IndE words, which was assumed to be the same as in AmE. In contrast, [30] reported that speakers of IndE and AmE used the same set of phonetic correlates to distinguish stressed and unstressed syllables, but the increase in intensity, duration and f_0 on stressed syllables was significantly smaller than in AmE (making the perceptual identification of stress in IndE more challenging for speakers of AmE and BrE). No differences were found on any of the acoustic parameters based on the speakers' L1s.

Previous research has thus shed light on the acoustic differences between syllables with primary stress and unstressed syllables, but not on differences between primary and secondary stress. In order to address the gap in present research, we examine the contribution of f_0 , intensity, syllable duration, and spectral balance to the difference between primary and secondary stress. This is particularly important as lexical stress appears to be variable in IndE, and it is unclear how consistent speakers are in its placement and acoustic realisation. This approach will help better understand the use of acoustic cues to signal stress in IndE.

2. METHODS

2.1. Speakers and stimuli

Eight speakers of Educated IndE (four L1 Hindi, four L1 Malayalam; five male, three female; all university students, 20-28 years old) participated in the study. They were recorded in 2012 in Hyderabad, India, in a quiet room with high-quality handheld audio recorders and head-mounted microphones. They had exclusively attended English-medium schools and universities, and had not resided outside of South Asia.

The stimuli consisted of six verb-noun pairs [23], where the verb (in BrE) has primary stress on the antepenultimate and secondary stress on the final syllable, and the noun has primary stress on the penultimate and secondary stress on the pre-antepenultimate syllable (e.g. *'conju, gate - , conju'gation*).¹ The keywords were embedded in a carrier phrase and presented to participants in written form with the keywords in capitals ('She said WORD again'), and interspersed with fillers. Three repetitions of each stimulus were recorded, leaving us with 288 tokens in total.

2.2. Auditory analysis

All stimuli were annotated by both authors independently for word stress, identifying each syllable as 'stressed' or 'unstressed'. The ratings were identical for 230/288 stimuli (Cohen's $\kappa = 0.72$, $p < 0.0001$), and agreement was reached for another 45 stimuli after discussion, bringing total agreement up to 275/288 (Cohen's $\kappa = 0.95$, $p < 0.0001$). The remaining stimuli, for which no agreement could be reached, were excluded from the analysis.

The analysis showed that all tokens can be divided into a left- and a right-prominent set. We refer to tokens with primary stress on the final or penultimate syllable (*calcu'late*, *calcu'lation*) as right-prominent, and tokens stressed on the antepenultimate or pre-antepenultimate syllable (*'calculate*, *'calculation*) as left-prominent. We also refer to final and penultimate syllables as right syllables, and antepenultimate and pre-antepenultimate syllables as left syllables.

2.3. Acoustic measurements

Duration, pitch slope, intensity and spectral balance of the ante-penultimate and final syllables of verbs (*CONjuGATE*), and the pre-antepenultimate and penultimate syllables of nouns (*CONjuGATION*) were analysed with a Praat script. We followed the

approach used by [23] as closely as possible.²

f_0 was measured with Praat's autocorrelation algorithm with parameters suitable for male (75-300 Hz) and female speakers (100-500 Hz), and minimum f_0 over the whole keyword and mean f_0 in the syllable were extracted and transformed to semitones (ST). Pitch slope (in ST/s) was then derived as $S = \frac{f_{max} - f_{min}}{t_{max} - t_{min}}$, where f_{min} and f_{max} are the minimum and maximum f_0 in the syllables, and t_{min} and t_{max} the times at which they are observed.

2.4. Statistical analysis

We used mixed effects regression models (R package lme4 [1]) to determine which contextual factors influence the acoustic correlates of stress in the data. The acoustic correlates were entered into separate regression models as dependent variables (following [23]). In addition, the following were used as independent variables: STRESS (syllable rated as having primary stress or not), PROMINENCE (whole token is left- or right-prominent), POS (part of speech: noun or verb), GENDER, L1 (Hindi or Malayalam). The first three were always entered into the analysis in order to determine (with post-hoc Tukey tests, p -level adjusted) whether syllables with primary and secondary stress in left- and right-prominent nouns and verbs differed in the size of the acoustic correlates of stress. The latter were only included when significant.

ITEM (i.e. the individual words) was used as a random factor. A second random factor, SPEAKER, could not be applied because this would have resulted in cells with sparse data and caused the regression algorithm not to converge. The data for each model was trimmed to remove outliers that unduly influenced the model (2.5 standard deviations above and below the residuals mean). This procedure never removed more than 5% of the data.

3. RESULTS

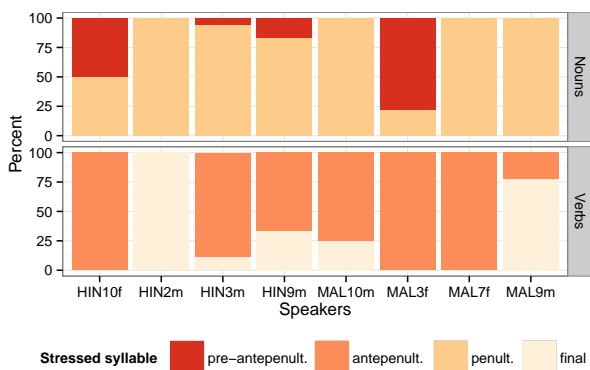
3.1. Auditory analysis

81% of the verbs (110 out of 135) were stressed on the antepenultimate syllable (e.g. *'calculate*), while 19% were stressed on the final syllable (e.g. *calcu'late*). Of the nouns, 69% (96 out of 140) were stressed on the penultimate syllable (e.g. *calcu'lation*) and 31% on the pre-antepenultimate (e.g. *'calculation*).

In the production of the nouns, five of the eight speakers always stressed the penultimate syllable (e.g. *calcu'lation*), while the remaining three used variable stress patterns (see Fig. 1). In the pro-

duction of the verbs, three speakers always stressed the antepenultimate (e.g. 'calculate), one always the final syllable (e.g. calcu'late), with the remaining four speakers relying on variable stress patterns. However, only one speaker was consistent in the production of both stimulus sets, and always stressed nouns on the penultimate and verbs on the antepenultimate. Stress patterns were also variable when comparing the different stimuli. Only *emulation* was always stressed on the penultimate syllable. Overall, this suggests that most speakers of IndE are variable in lexical stress placement, but that primary stress can only be shifted to syllables with secondary stress.

Figure 1: Stress patterns for individual speakers

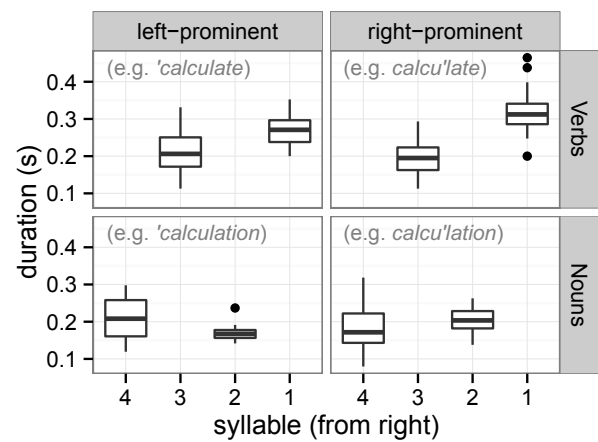


3.2. Acoustic analysis

3.2.1. Duration

As Fig. 2 shows, in left-prominent nouns, left (stressed) syllables were longer than right (unstressed) syllables ($p < 0.001$), whereas in all other cases, right syllables were longer than left syllables ($p < 0.001$). In addition, right (stressed) syllables of right-prominent verbs were longer than right (unstressed) syllables of left-prominent verbs and nouns ($p < 0.001$). The right (stressed) syllables of right-prominent verbs were also longer than those of right-prominent nouns ($p < 0.001$). This suggests that differences in duration provide cues for listeners to distinguish (1) right-prominent verbs and nouns from left-prominent verbs and nouns, respectively, but not (2) to distinguish left-prominent verbs from right-prominent nouns.

Figure 2: Duration



3.2.2. Pitch slope

PITCH SLOPE was steeper in the left (stressed) syllables of left-prominent words than in their right (unstressed) syllables, whereas in right-prominent words the right (stressed) syllable had a steeper pitch slope than the left (unstressed) syllable. However, this difference was only significant for right-prominent nouns. This suggests that differences in PITCH SLOPE provide auditory cues for listeners to distinguish (1) right-prominent nouns from left-prominent nouns and verbs, but that (2) right- and left-prominent verbs may not be distinguishable based on PITCH SLOPE alone.

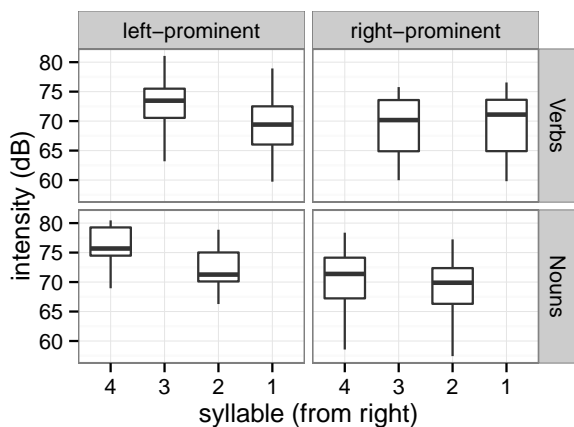
3.2.3. Intensity

As Fig. 3 shows, intensity was higher in left (stressed) than in right (unstressed) syllables of left-prominent words (verbs $p < 0.001$, nouns $p = 0.055$). In right-prominent words, the right (stressed) syllable had slightly and insignificantly higher (verbs) or lower (nouns) intensity than the left (unstressed) syllable. This suggests that differences in INTENSITY provide auditory cues for listeners to distinguish left-prominent from right-prominent pronunciations, including left-prominent verbs from right-prominent nouns.

3.2.4. Spectral balance

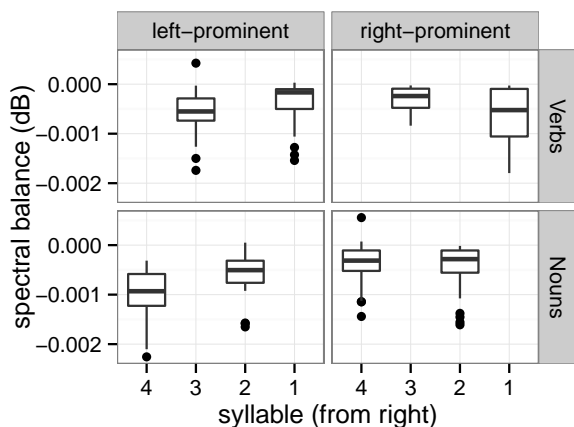
In left-prominent words, the left (stressed) syllable had a significantly lower (more skewed) spectral balance than the right (unstressed) syllable (verbs $p = 0.01$, nouns $p = 0.001$, see Fig. 4). In

Figure 3: Intensity



right-prominent nouns, right (stressed) and left (unstressed) syllable did not differ ($p>0.999$), whereas in right-prominent verbs the right (stressed) syllable had a lower (more skewed) spectral balance than the left (unstressed) syllable ($p<0.001$). This suggests that differences in SPECTRAL BALANCE provide auditory cues for listeners to distinguish left-prominent from right-prominent pronunciations, including left-prominent verbs from right-prominent nouns.

Figure 4: Spectral balance



4. DISCUSSION

This study set out to determine the acoustic realisation of the contrast between primary and secondary stress in Educated IndE, taking into account variable stress patterns in verb-noun pairs of the type *'conju,gate - ,conju'gation*. Results showed that

most speakers are variable in where they place lexical stress (confirming [11, 29]), regardless of L1 (contrary to [11]). However, this variability is restricted in that (1) primary stress falls in the majority of cases on the same syllable as in BrE and AmE and (2) primary stress can only be shifted to syllables with (underlying) secondary stress.

While similar cases of stress shift are uncommon in BrE and AmE, the primary-secondary stress distinction has a relatively low functional load in these varieties (compared to, e.g., Spanish and Dutch). Listeners rely less on the acoustic correlates of primary vs. secondary stress in word recognition, but mainly on the difference between syllables with primary stress and unstressed syllables [4]. Like other postcolonial varieties of English, IndE was formed in a process of dialect levelling and contact with local languages [25], which might have further reduced the functional importance of the primary-secondary stress distinction.

Regarding the acoustic realisation of the difference between primary and secondary stress, the results suggest that speakers of IndE rely on differences in (in order of importance) intensity, spectral balance, duration, and pitch slope. In certain cases, our results are supported by previous analyses of IndE. [7–9] showed that prominence-lending increases in intensity and duration, and variability in f_0 and duration, do not co-occur in IndE as often as in BrE and offset each other occasionally, contributing to a more syllable-timed rhythm in IndE (compared to BrE). Furthermore, [17] showed that, in IndE, more syllables receive pitch accents than in other varieties such as BrE. This might explain why pitch slope and duration are less reliable cues to the primary-secondary stress distinction in IndE than intensity and spectral balance. In addition, the role of f_0 may be more relevant in cueing accentual/post-lexical prominence in IndE.

Educated IndE appears to differ from other varieties in the realisation of the primary-secondary stress distinction. Our results are similar to previous research on AmE (based on a similar methodology [23]) in that intensity is a primary cue to the primary-secondary stress distinction in both varieties. However, some cues, which play a less important (spectral balance) or no role (duration, pitch slope) in AmE were found to be more important in IndE, giving further proof that IndE has developed its own phonological features distinct from the long-standing varieties of English.

5. REFERENCES

- [1] Bates, D., Maechler, M., Bolker, B. 2013. *lme4: Linear Mixed-Effects Models Using Eigen and S4*. R package version 0.999999-2.
- [2] Beckman, E. J., M. 1994. Articulatory evidence for differentiating stress categories. In: Keating, P. A., (ed), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press 7–33.
- [3] Beckman, M. E. 1986. *Stress and Non-Stress Accent*. Dordrecht/Riverton: Foris.
- [4] Cutler, A. 2012. *Native listening: Language experience and the recognition of spoken words*. MIT Press.
- [5] Desai, S. B., Dubey, A., Joshi, B. L., Sen, M., Shariff, A., Vanneman, R. 2010. *Human Development in India*. Oxford: Oxford University Press.
- [6] Fuchs, R. 2012. Focus marking and semantic transfer in indian english: the case of *also*. *English World-Wide* 33(1), 27–53.
- [7] Fuchs, R. 2013. *Speech Rhythm in Educated Indian English and British English*. PhD thesis University of Münster.
- [8] Fuchs, R. 2014. Integrating variability in loudness and duration in a multidimensional model of speech rhythm: Evidence from indian english and british english. Campbell, N., Gibbon, D., Hirst, D., (eds), *Proceedings of Speech Prosody 7, Dublin* 290–294.
- [9] Fuchs, R. 2014. Towards a perceptual model of speech rhythm: Integrating the influence of f0 on perceived duration. Li, H., Meng, H., Ma, B., Chng, E. S., Xie, L., (eds), *Proceedings of Interspeech 2014* Singapore. 1949–1953.
- [10] Fuchs, R. 2015. You're not from around here, are you? - a dialect discrimination experiment with speakers of british and indian english. In: Delais-Roussarie, E., Avanzi, M., Herment, S., (eds), *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations*. Berlin: Springer 121–146.
- [11] Gargesh, R. 2004. Indian english: Phonology. In: Schneider, E. W., Burrige, K., Kortmann, B., Mesthrie, R., Upton, C., (eds), *A Handbook of Varieties of English* volume 1. Berlin: Mouton de Gruyter 992–1002.
- [12] Gussenhoven, C. 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- [13] de Jong, K. 2004. Stress, lexical focus, and segmental focus in english: Patterns of variation in vowel duration. *Journal of Phonetics* 32, 493–516.
- [14] Lange, C. 2012. *The Syntax of Spoken Indian English*. Amsterdam: Benjamins.
- [15] Low, E. L., Grabe, E. 1999. A contrastive study of prosody and lexical stress placement in singapore english and british english. *Language and Speech* 42(1), 39–56.
- [16] Masica, C. P. 1972. *The Sound System of Indian English*. Hyderabad: Central Institute of English and Foreign Languages.
- [17] Maxwell, O. 2014. *The Intonational Phonology of Indian English. An Autosegmental-Metrical Analysis Based on Bengali and Kannada English*. PhD thesis University of Melbourne.
- [18] Maxwell, O., Fletcher, J. 2009. Acoustic and durational properties of indian english vowels. *World Englishes* 28(1), 52–69.
- [19] Maxwell, O., Fletcher, J. 2010. The acoustic characteristics of diphthongs in indian english. *World Englishes* 29(1), 27–44.
- [20] Maxwell, O., Fletcher, J. 2011. Phonetic cues to accentual prominence in bengali english. *Online Proceedings of the 17th International Congress of Phonetic Sciences* Hong Kong. 1350–1353.
- [21] Mukherjee, J. 2007. Steady states in the evolution of new englishes: Present-day indian english as an equilibrium. *Journal of English Linguistics* 35(2), 157–187.
- [22] Pickering, L., Wiltshire, C. 2000. Pitch accent in indian-english teaching discourse. *World Englishes* 19(2), 173–183.
- [23] Plag, I., Kunter, G., Schramm, M. 2011. Acoustic correlates of primary and secondary stress in north american english. *Journal of Phonetics* 39(3), 362–374.
- [24] Sailaja, P. 2012. Indian english: Features and sociolinguistic aspects. *Language and Linguistics Compass* 6(6), 359–370.
- [25] Schneider, E. W. 2007. *Postcolonial English: Varieties Around the World*. Cambridge: Cambridge University Press.
- [26] Sirsa, H., Redford, M. A. 2013. The effects of native language on indian english sounds and timing patterns. *Journal of Phonetics* 41(6), 393–406.
- [27] Sluijter, A., van Heuven, V. 1996. Acoustic correlates of linguistic stress and accent in dutch and american english. *Proceedings of the International Congress of Spoken Language Processing* Philadelphia. 630–633.
- [28] Sluijter, A. M., Shattuck-Hufnagel, S., Stevens, K. N., van Heuven, V. J. 1995. Supralaryngeal resonance and glottal pulse shape as correlates of stress and accent in english. Elenius, P. B., K., (ed), *Proceedings of ICPHS 13* Stockholm. 630–633.
- [29] Vijaykrishnan, K. G. 1978. Stress in tamilian english: A study with the framework of generative phonology. Master's thesis Central Institute of English and Foreign Languages Hyderabad.
- [30] Wiltshire, C., Moon, R. 2003. Phonetic stress in indian english vs. american english. *World Englishes* 22(3), 291–303.

¹ The following verb-noun pairs were used: *conjugate - conjugation, delegate - delegation, emulate - emulation, hibernate - hibernation, terminate - termination, illuminate - illumination*.

² The only difference being that [23] relied on the duration of the sonorous part of each syllable, whereas we used the total duration of the syllable.