The prosody of Focus and Givenness in Hindi and Indian English

Caroline Féry, Pramod Pandey, and Gerrit Kentner
Goethe University Frankfurt / Jawaharlal Nehru University, Delhi /
Goethe University Frankfurt

This paper reports the results of two identical experiments, one in Hindi and one in Indian English, that elicited semi-spontaneous sentences containing a focused agent or a focused patient. The primary aim of the experiments was to investigate the prosodic correlates of information structure in the two languages and to explain these correlates with a phonological model. The resulting phonological model proposes that focus is realized with enhanced correlates of phrasing and not with prominence, at least not of the same kind as languages using pitch accents. Secondary aims were to verify the ecological validity of similar data elicited with scripted speech (Patil et al. 2008) and to reflect on the place of Hindi and Indian English in a typology of intonation.

Keywords:

1. Introduction

In this article, the prosodic realization of information structure in Hindi and Indian English is investigated. The data underlying the study were obtained by means of a simple production experiment used in both languages, which elicited focus on an agent or a patient while the remainder of the sentence was given information. The obtained material can be called 'semi-spontaneous.' Utterances were elicited by picture prompts to which the speakers answered without script. The data elicited in Hindi fulfils the role of verifying the results that Patil et al. (2008) obtained by means of scripted speech. It thus assesses the ecological validity of generalizations obtained in the lab by scripted speech. It will be shown that semi-spontaneous data lead to results that are comparable to those obtained by scripted speech. The Indian English data serves the purpose of comparing the intonation structure of this language with those obtained for Hindi.

Studies in Language 40:2 (2016), **302–339**. DOI 10.1075/sl.40.2.02fer
ISSN 0378–4177 / E-ISSN 1569–9978

From a typological point of view, Hindi is particularly interesting, since it does not fit easily in a clear intonational class. In traditional typologies of intonation systems, languages are classified according to the presence or absence of lexical stress, lexically specified pitch accents, and tones. In Ladd's (1996/2008) review of the possible types of languages, Bangla and Hindi are defined with negative features [-lexical stress, -lexical accent, -tone], a proposal also adopted by Jun (2005) for Korean. In Gussenhoven's (2004) division of intonation in lexical and post-lexical intonational features, languages are primarily defined by their lexical features, like tone and stress. Languages can also have postlexical boundary tones, that is tones assigned to boundaries of tonal domains. Which of these lexical features and post-lexical boundary tones are present in Hindi and Indian English is an empirical question, but it also depends on the definition of post-lexical tones.

Both Hindi and Indian English have a rich system of tones mostly assigned at the level of the prosodic phrase or at the level of the prosodic word. Interestingly, and as has been noticed by a number of researchers (see next section), focus does not add much phonetic prominence on the focused constituent. Instead when a constituent is focused, the tones associated with the prosodic domain called 'prosodic phrase' (see below for definition) may be slightly lower and slightly higher than when they are without focus. A striking fact is that the prosodic phrase associated with the focus is set off from the remainder of the sentence in a clearer way. Among other factors, this is manifested by segmental cues.

The questions investigated in this paper are the following. First, the phonetic, prosodic and segmental correlates of a prosodic phrase when it is focused and when it is given are compared: F0, intensity and duration are potential phonetic correlates of focus and givenness; we will see that other correlates of phrasing must be considered as well. Second, the validity of scripted speech and its verification by semi-spontaneous speech are assessed: Are the results obtained by scripted speech in Patil et al. (2008) and those elicited by letting speakers choose their own wordings comparable, or do they differ? Third, a phonological interpretation of the data is proposed. It is assumed that the concept of prominence for focus is expressed primarily by correlates of phrasing, and not so much by pitch accents. As a side question we also raise the issue of the place of Hindi and Indian English in a typological system of intonation. The results presented below are limited in scope, and the conclusions drawn from them can only be preliminary and exploratory.

The elicited utterances discussed below are divided into a focused and a given portion. Focus is the part of a sentence eliciting a set of alternatives relevant for the interpretation of discourse. When presenting our data, a three-way distinction used for eliciting the data is introduced: information, selection and correction. In presenting the phonetic correlates, however, we treat focus as one category and givenness as another and do not make any distinction between the kinds of focus

used for elicitation. In that, we follow Rooth (1985, 1992) and Krifka (2008: 248) who do not distinguish between different kinds of focus.

Givenness characterizes the constituents that are present in the Common Ground of the interlocutors (Schwarzschild 1999, Krifka 2008: 262). In our experiments, given constituents have always been mentioned in the question, and they were often repeated in the answer. This minimal definition is in line with the anaphoric nature of givenness assumed by these authors.

The next section of this article provides a background to the work reported in this article. It passes review of some literature on the prosody of Hindi and Indian English and it also contains a brief survey of possible phonological analyses of Hindi and Indian English. Section 3 introduces the methodology used to elicit the data exposed in the subsequent sections. In Section 4, the results are presented for Hindi and in Section 5 for Indian English. Section 6 contains a discussion on the phonological issues and concludes the paper.

2. Background

In investigating the intonation of Hindi and Indian English, we do not enter an unexplored field. Indeed, there are quite a large number of studies dealing with these languages, and we can thus firmly ground our work on previous insights.

In Hindi the status of lexical stress, that is whether there is always a designated syllable for stress, is unclear. In the following, we assume that Hindi has lexical stress, and that as acknowledged by many authors, syllable weight plays a major role in its location. Pandey (1989) emphasized the fact that Hindi is spoken by a very large population, of whom a large part are native speakers of other Indian languages, Indo-Aryan or others, which may constitute an explanation for the difficult state of lexical stress in Hindi. Hayes (1995: 163), following Pandey, used the term "non-phonemic lexical accent", see also Pandey (to appear) for a formal proposal for word stress in Hindi and Puri (2013) for a thorough review of the literature on lexical stress in Hindi.

Lexical stress is realized in isolation, when the words are pronounced out of context, but it is often only weakly perceivable in spontaneous speech. The reason for this latter property is the absence of a systematic pitch accent corresponding to lexical stress. When accounting for the absence of the Nuclear Stress Rule in Hindi, Bansal (1976: 27) writes the following: "*The main differences [between Hindi and English] were found to be in tonicity. Except where the location of the tonic syllable is determined by contextual emphasis or contrast, English places the 'tonic' on the last important word in the tone-group, but no such rule operates in Hindi.*"

Intonational markers in the form of phrasal tones are often more prominently realized than lexical stress.

Moore (1965) showed that every non-final prosodic phrase in Hindi, which he called 'foot', starts with a low tone. The prosodic phrase is defined as "one to several syllables in length, which normally is uttered with a continuously rising pitch from beginning to end." Harnsberger (1994, 1999) and Harnsberger & Judge (1996) analysed the low part of the rising contour as a low pitch accent, annotated as L* in an autosegmental-metrical (AM) notation system (Pierrehumbert 1980, Ladd 2008), and the high part of the rising contour as a boundary tone H%, or H-. In the AM notation, a starred tone is associated with a metrically prominent syllable at the lexical or post-lexical level.

As for focus (or 'emphasis' as it is sometimes called in the literature), Moore (1965), Harnsberger (1994, 1998) and also Dyrud (2001) found that it has the effect of inserting a boundary tone at the left edge of the focused phrase. They find that the rising pitch pattern may also show a higher excursion, a greater intensity and longer duration. Post-focally the pitch range may be compressed or even be completely flat and deaccented, although rising pitch accents are still realized in compressed pitch range (Harnsberger & Judge 1996). Similarly, Puri (2013) finds a larger F0 excursion in narrow focus than in wide focus and post-focal compression.

Jyothi, Cole, Hasegawa-Johnson & Puri (2014) asked ten non-expert speakers of Hindi and one expert (the paper's last author) to identify prominent words and prosodic boundaries in 10 excerpts from spoken short narratives. They found that the non-experts do not agree with each other and with the expert as far as prominent words are concerned ($\kappa = 0.15$). The agreement is better with prosodic boundaries ($\kappa = 0.41$). Moreover, using automatic detection of prominence and boundaries (AuToBI) also led to different results from those of non-expert speakers. When asked to detect prosodic boundaries, all three groups agreed better. The expert and AuToBI had a higher agreement with each other than with the non-expert group. Altogether these results seem to confirm that prominence is a difficult notion in Hindi, and that prosodic boundaries are perceived better than prominence. Hindi speakers differ greatly from English speakers. For English speakers, Yoon, Chavarría, Cole & Hasegawa-Johnson (2004) found agreement rates at 89% for boundaries and 86% for pitch accent. Mo, Cole, and Lee (2008) found a mean kappa (κ) of 0.582 for non-expert speakers for prominence and boundaries.

Patil et al. (2008) conducted a simple production experiment to investigate the effects of word order and focus in controlled lab data using scripted speech. They let 20 native female speakers of Hindi, students at Delhi University, produce sentences as answers to questions eliciting different kinds of focus. There were three conditions for focus (wide focus vs. subject focus vs. object focus), and two conditions for word order (SOV vs. OSV), thus six conditions in all (3x2).

Figure 1 shows the results for the SOV word order conditions in the three focus conditions, in the form of pitch contours averaged over all speakers and all sentences. The results for this word order can be summed up as follows: First, the subject showed an earlier and slightly higher pitch rise when in focus, second, the pitch contour was compressed in the post-focal region, and third, there was no focus marking on the object. The results of the OSV order were very similar to those obtained in the SOV word order, except that there was less increase of F0 on a focus initial object than in the initial focused subject. All in all, the most robust effect was indirect: focus elicited compression of the postfocal material. A direct and slightly significant effect was only found on the focused constituent itself when it was the sentence-initial subject. In all other conditions, there was no effect on the height of the high tone of the focused constituent.

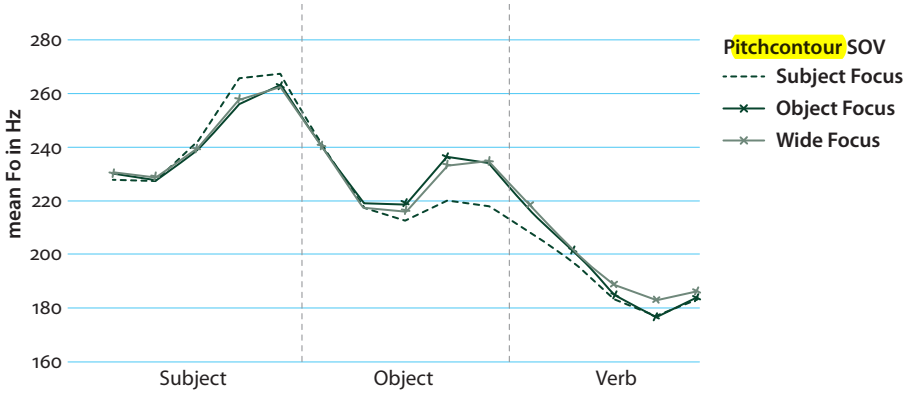


Figure 1. Averaged data for SOV word order in 3 focus conditions (10 sentences, 20 speakers).

Patil et al (2008) elicited their data with the use of scripted speech, thus in a totally controlled way. The data consisted in short sentences produced as answers to contexts eliciting information focus in form of wh-questions. The spoken material was read from a screen, leaving no place for word order changes or other individual variations from the participants.

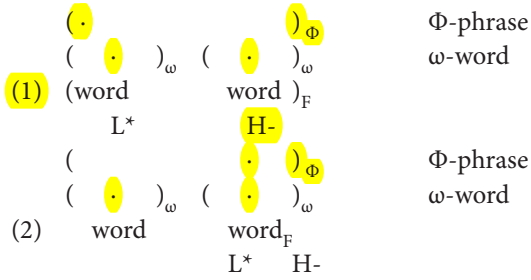
In the present article, the contexts used for elicitation of sentences are similar to those used by Patil et al. They also elicited focus on the agent and focus on the patient, and complementary givenness. In other words, both datasets were recorded under laboratory conditions under artificial circumstances. However, no scripted speech was used, which means that more natural data were elicited since. The participants were free to use their own wordings to answer the questions asked to them. Moreover, additional contexts were used, and more phonetic correlates were measured. This procedure allowed us to compare the phonetic correlates of

scripted speech with those of more natural data, and to discover a great overlap between the two kinds of data.

Since the data were only elicited in the context of narrow focus and givenness, we rely on the data by Patil et al. (2008) for a comparison with a wide focus context.

As for the phonological interpretation of the data, two competing models have been proposed for Hindi. The first one assumes that Hindi is similar to English in assigning prominent tones (L*) on lexically stressed syllables, as proposed first by Harnsberger & Judge (1996), see (1). The other one insists on the formation of prosodic phrases on focused material and does not necessarily associate the low and high tones with lexical prominence, as proposed by Patil et al (2008) and Féry (2010).

The new semi-spontaneous data helps to investigate these theoretical claims, in particular the status of the low and high tones at the beginning and end of prosodic phrases. Even if we do not provide a definite answer to this question due to the restricted amount of material elicited, some theoretical issues can nevertheless be clarified. In particular, the metrical and the phrasing hypothesis can be compared in the light of the results. The metrical interpretation of tones is illustrated in (1) and (2), with an idealized prosodic phrase (Φ-phrase) consisting of two prosodic words (ω-words). In the default case, illustrated in (1), the L* is assigned on the lexical stress of the first word, and the boundary tone H- is associated with the final syllable of the Φ-phrase. The metrical hypothesis predicts that the low tone is a pitch accent assigned to the most prominent syllable of a metrical grid. If the second word of the Φ-phrase is focused, as in (2), L* should be shifted to the metrically most prominent syllable of this new word. If prominence is culminative, the first word looses its prominence.



The phrasing analysis is illustrated in (3) and (4). It claims that both the low tone and the high tone are phrase tones, written L_Φ and H_Φ respectively, and that they are associated with the beginning and the end of a Φ-phrase. In case of a wide focus, the difference between the two approaches is minimal, because the distribution of the tones is identical. In (3), the entire Φ-phrase is focused and the two words are phrased in a single Φ-phrase with an initial low tone and a final high

tone. In case of narrow focus on the second word, as in (4), a new phrase is created on this word and additional tones are added on the first and second words, a high boundary tone on the first one, and a low one at the beginning of the focused word. These tones are indicative of a new Φ-phrase on the focused word. The first word is now phrased in its own Φ-phrase.

	( )$_{\Phi}$		Φ-phrase
	( )$_{\omega}$	( )$_{\omega}$	ω-word
(3)	(word	word)$_{F}$	
	L_{Φ}	H_{Φ}	
	( )$_{\Phi}$	( )$_{\Phi}$	Φ-phrase
	( )$_{\omega}$	( )$_{\omega}$	ω-word
(4)	word	word$_{F}$	
	L_{Φ} H_{Φ}	L_{Φ} H_{Φ}	

Let us first introduce the data and we will come back to the phonological analysis in Section 6.

3. Experimental method

The data discussed in this article were elicited by means of a task of the QUIS questionnaire (Skopeteas et al. 2006), called 'Anima'. The goal of the Anima task is to compare the outputs of the speakers for different kinds of focus and for givenness as far as word order and prosodic properties are concerned, and in this way, to be able to draw conclusions on the use of correlates of focus and givenness in different languages. The task uses two factors of variation. First, the focused constituent, which is the agent or the patient, and second, the focus type: information focus (answer to a *wh*-question), selection focus (choice between two possible answers) and correction focus (the suggested answer is wrong and needs to be corrected). These three types of focus are organized in a strength hierarchy. The weakest one is the information focus, selection focus has an intermediary strength, and correction focus is the strongest.

For example, one picture shows a man pushing a car in front of a well.[1] The questions used for this picture are listed in (1).

1. All agents were human, but half the patients were human and half inanimate. Since the distinction did not correlate with any difference in the results, human and inanimate patients were analysed together.

(1) Questions used in the Anima task
Information focus on the subject (S): 'In front of the well, who is pushing the car?'
Information focus on the object (O):'In front of the well, what is the man pushing?'
Selection focus on S: 'In front of the well, is a man or a woman pushing the car?'
Selection focus on O: 'In front of the well, is the man pushing a car or a bicycle?'
Correction focus on S: 'In front of the well, is a woman pushing a car?'
Correction focus on O: 'In front of the well, is the man pushing a bicycle?'

For the present study, we devised a presentation with a total of twelve pictures depicting simple events (involving an agent and a patient). The pictures were distributed over three slides with each slide containing four images. For each depicted event, six questions were formulated, inducing different focus-background configurations as in (1).

The participants were allocated to four groups and each group received one question per picture, such that focus conditions were roughly counterbalanced across groups.[2]

When presented with the slides, informants were instructed to observe the stimuli attentively and to memorize the details of each depicted event. When the participants were ready, the stimuli were taken away and the informants were asked four questions, one for each picture. They were instructed to give full answers.

Sixteen native speakers of Hindi and twenty native speakers of Indian English (mostly students) were recorded at Jawaharlal Nehru University (JNU), Delhi (in March 2011 for Hindi, and between March 2011 and January 2012 for Indian English) with the help of graduate students. Five bilingual speakers participated in both experiments.

Recordings were made in a quiet room with a Shure head-mounted microphone using an Edirol digital recorder. The stimuli were presented as a Powerpoint slide show. The Anima task used for this article was one of five and it was presented in a block, thus the twelve sentences of this experiment were produced in a row.

2. Not all the possible questions (12 x 6 = 72) were actually asked, but only 12 x 3 = 36.

4. Results for Hindi

4.1 Data

For Hindi 16 speakers (9 female and 7 male) were recorded. 139 sentences (from 192) were retained for analysis. The sentences that were not considered for analysis were eliminated for one of the following reasons.

In one of the pictures, a patient is visually ambiguous. It can be interpreted as a boy or a girl. In the case of information or selection, this was not a problem; all sentences, regardless whether they contained *boy* or *girl* as an answer, were considered. In the case of correction, however, this ambiguity led to plainly wrong answers: some speakers did not correct the object but confirmed the question instead. These sentences were eliminated.

Some further questions eliciting a correction were not corrected, and instead only a negation was produced, or a confirmation was produced. These sentences were eliminated.

The second author, a native speaker of Hindi, decided to eliminate six wrongly intonated sentences. For these sentences, the speaker answered for a patient focus instead of an agent focus or vice-versa.

On the other hand, since we were interested in the form of the sentences rather than in their contents, some erroneous answers were not eliminated, especially those eliciting information asked with a wh-question. Recall that the experiment was presented in the form of a memory test. The speakers had to remember four pictures, and at the time they answered the questions, the pictures were not visible anymore. This procedure lead to some additional errors.

The remaining sentences were transcribed in Hindi and English by a student in Delhi. There were 75 sentences with a focus on the subject and 64 sentences with a focus on the object, classified in kinds of focus as shown in Table 1. Due to the elimination of flawed sentences, the data set is considerably reduced and unbalanced. Correspondingly, a valid comparison of all conditions is impossible. With the data at hand, the analysis of the kinds of focus elicited by the question (information, selection and correction) did not reveal systematic differences in the prosodic form of the sentences. Therefore, in the following, we concentrate on the distinction between focused and given information but disregard the focus type.

Table 1. Analysed Hindi sentences (139)

	Subject	**Object**
Information	28	26
Selection	26	24
Correction	21	14
Sum	75	64

The utterances obtained were cut and labelled using the acoustic speech analysis software Praat (Boersma & Weenink 1993–2013) with the help of undergraduate students in Frankfurt, the result of which was verified and corrected by the authors.

The recorded sentences were digitized at a sampling rate of 16 kHz and were analysed in Praat. The recordings were manually divided into labelled substrings with the help of spectrograms and acoustic inputs. The F0 contours were smoothed using the Praat smoothing algorithm (frequency band 10 Hz) to minimize micro-prosodic perturbations. All frequency measurements were semi-automatically done using a script that detects the mean and the maximum F0 value (in Hz) and intensity value (in dB) within a given domain and that records the duration of the interval. The domains for measurements were the subjects and objects of each sentence, which were adjacent in most cases. The analysis was done in a number of steps. First, a Praat script located F0 maxima and minima plus 5 points in each domain. Second, the results of the Praat script were hand-edited to correct spurious labelling. The authors individually inspected the tone labels against the F0-track, the substring divisions, an auditory impression, and the spectrogram. Where the Praat script had assigned a tone label (L or H) that was not in the correct position (because of obvious errors due to the algorithm), the label was manually moved. In a third step, another Praat script recovered the F0-values at the positions of the tone labels as well as the tone labels themselves, and collected them in a table.

Two versions of the sentences were prepared. In the 'long' version (Tables 1–2), all subjects and all objects were labelled, including pre- and post-nominal articles, adjectives and case markers if any. In this way, variations in the position of the tonal correlates of the analysed constituents, as well as variations of intensity, could be measured, even if these correlates were not located exactly on the noun, but rather on an adjacent syllable. In the 'short' version (Table 3), only the disyllabic nouns were considered and separated by boundaries, without articles and modifiers. The short versions were used for measurement of duration, as we needed to measure entities that were comparable in size.

Most answers were uttered in the SOV word order; we will see examples of such sentences below. However, 22 out of 75 answers to questions eliciting a subject focus had an OSV word order. This happened only in the contexts eliciting an information focus on the subject. We strongly suspect an effect of the question, which was in the OSV order only in this case.

In Table 2, the results for word order are summed up: in addition to the sentences with SOV and OSV orders, there was one cleft construction, three relative clauses and six sentences containing a negation. Examples for these kinds of sentences are given from (2) to (4).

Table 2. Word order in Hindi

	SOV	OSV	Cleft	RelCl	SOV + Neg
Information S (n = 28)	6	22			
Information O (n = 26)	26	–			
Selection S (n = 26)	22	–	1	3	
Selection O (n = 24)	24	–			
Correction S (n = 21)	18	–			3
Correction O (n = 14)	11	–			3
Total 139	107	22	1	3	6

(2) Cleft construction in a subject selection sentence (Speaker 5, 2.1.1)
jəhã: ni:la: a:ka:ʃ hæ: ʋəhã: e:k əorət a:dmi: ko: ma:r rəhi: hæ:
where blue sky BE-3PS there one woman man DAT hit AUX-F BE-3PS
'Where the sky is blue, there a woman hitting a man.'

(3) Relative clause in a subject selection sentence (Speaker 9, 3.1.4)
kuẽ: ke: sa:mne: ʤo: ka:r ko: dʱəkka: de: rəha:
well-SG-OBL GEN-OBL front-OBL who car DAT push give AUX-PROGR-M
hæ: ʋo: a:dmi: hæ:
BE-3PS that man BE-3PS
'In front of the well, the one who is pushing a car is a man.'

(4) SOV Neg in an object selection sentence (Speaker 10, 3.2.1)
əorət e:k pe:ɽ ko: ma:r rəhi: hæ: nə ki e:k phu:l ko:
woman one tree DAT hit AUX-PROGR-F BE-3PS NO COMP one flower DAT
'A woman is hitting a tree, not a flower.'

4.2 Results of the acoustic analysis

In order to get information not only about average phonetic values over whole constituents but also about the contour, the data were analysed as follows: each labeled constituent was automatically divided into five equal-sized intervals, and the mean pitch in Hz and the mean intensity in dB was extracted for each interval (one interval corresponding to less than a syllable). Because of the non-normal distribution (data came from male and female informants), the F0 data were normalized for plotting. The inverse of the overall mean F0 of each item was used as normalizing factor for the raw values, multiplied by the average F0 over the whole data set.

Figure 2 shows normalized pitch tracks broken down by focus and word order (note that OSV word order only occurs in subject focus conditions). Inspection of the SOV word orders reveals only little focus-induced difference in pitch on the

first argument. As for the second argument, there was a small but significant difference in pitch: its curve is altogether lower when the constituent is given (and the preceding argument is focused), most likely an effect of post-focal compression. This observation confirms the results of Patil et al. (2008) who used scripted speech.

The effect of word order is more difficult to interpret as it is confounded with focus and there are only 22 utterances with OSV word order: in any case, the unfocused object in OSV word order appears to be lower in pitch compared to the two SOV conditions, while the focused second argument (the subject) is higher than unfocused objects in SOV, but lower than focused objects.

These data were evaluated by means of linear mixed effects models (using R statistical software) with mean F0 as the dependent variable. (Using normalized values does not change the results qualitatively. We chose to report results from the raw values so that the coefficient estimates reflect the differences in Hz brought about by the fixed factors. The individual differences in overall F0 were controlled for by including speaker and item as random effects).

In order to ascertain the effects of focus and word order on pitch deflection on the two arguments, we specified models with the following factors included as fixed effects:

1. Effect of word order (SOV vs. OSV)
2. Effect of focus position (comparing focus on first vs. focus on second argument)
3. Effect of constituent position (comparing first vs. second argument)
4. Interaction of 2. and 3. (analysing the effect of focus position by constituent)

Speaker, item, and interval were included as random factors.

Note that word order–focus interactions could not be checked in addition to the interaction between focus position and constituent position because OVS word order only occurred in sentences with subject focus.

Analysis of F0: The full model yields a significant main effect for word order (Est.: 7.371, Std.Err: 3.123, t-value = 2.360). The main effect for constituent (Estimate: −24.594, Std.Err: 9.077, t = −2.710) is significant, too, reflecting the fact that the second argument is on average lower in pitch than the first. The factor focus position also has a significant effect on F0 (Est.: −9.111, Std.Err: 4.184, t = −2.178); that is, overall F0 is about 9 Hz lower when the first argument is in focus. A significant interaction of focus position and constituent reveals that this difference is mainly due to the second argument (Est.: 10.501, Std.Err: 2.1, t = 5.0), most likely reflecting an effect of moderate post-focal compression. The importance of each of these factors is confirmed by model comparison using ANOVAs: removing any of the fixed factors leads to significantly worse model fit:

Full model compared to
Model without word order: X2 = 5.2681, p = 0.022,
Model without focus position: X2 = 29.85, p < 0.001
Model without constituent position: X2 = 25.814, p < 0.001
Model without interaction of focus position and constituent: X2 = 24.774, p < 0.001

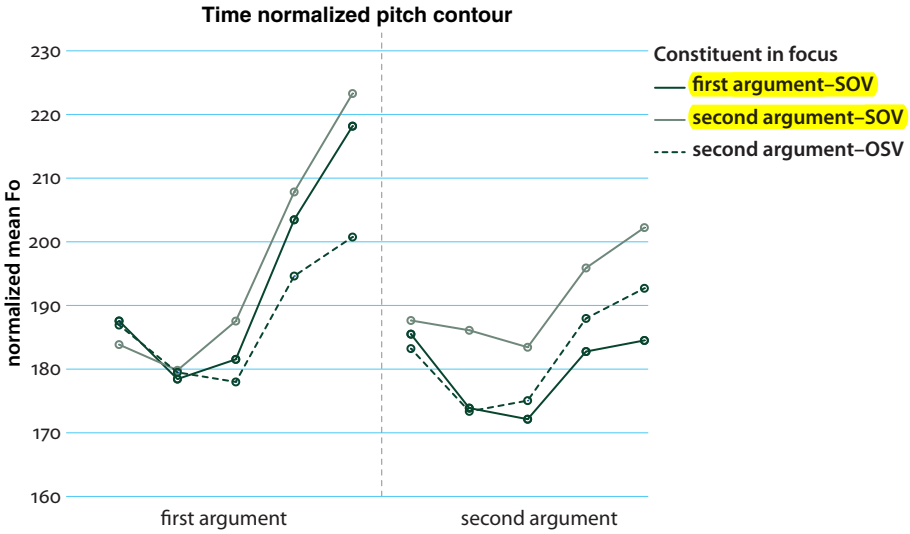


Figure 2. Hindi: Time normalized pitch contours for the two arguments (average across speakers and items)

4.2.2 *Intensity*

For the analysis of intensity, the mean amplitude in dB was extracted from the intervals. This data, too, was subjected to an analysis with linear mixed effects models, using the same fixed and random effects as in the analysis of F0. The model yields a significant effect of focus position on intensity (Est.: −2.9560, Std. Err: 1.0985, t-value = −2.691), reflecting overall higher amplitude when the second argument is focused. The effect of word order is non-significant (Est.: −0.697, Std.Err: 0.7478, t = −0.932). Furthermore, there is a marginal effect of constituent number (Est.: −2.998, Std.Err: 1.944, t = −1.542), reflecting decreasing intensity over the course of the utterance.

A marginal effect of the interaction of focus position and constituent number (Est.: 1.144, Std.Err: 0.602, t = 1.901) suggests higher intensity on focused arguments. A model comparison using ANOVAs shows that model fit does not suffer significantly from removing word order as fixed effect ($X^2 = 0.8053$, p = 0.37); however, model fit deteriorates from removing any of the other factors (all ps < 0.05), confirming the importance of even the marginally significant effects. The intensity data are summarized in Figure 3.

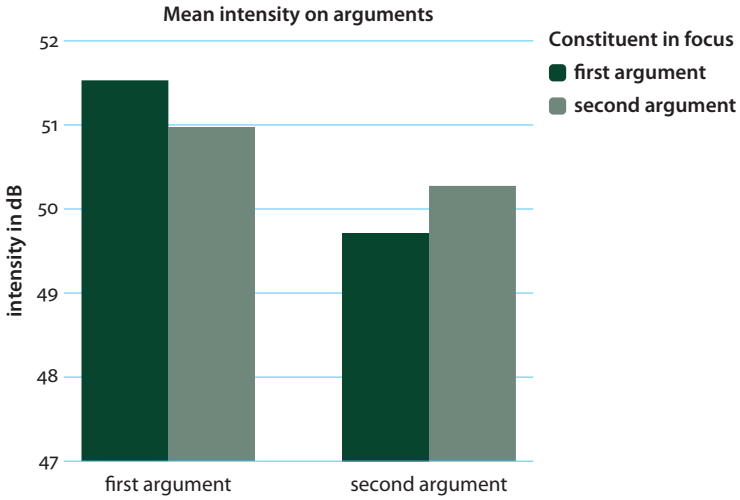


Figure 3. Hindi: Mean intensity on arguments (average across speakers and items)

4.2.3 *Duration*

The analysis of duration is complicated by the fact that monosyllabic, disyllabic and trisyllabic arguments with intrinsically different lengths were used in the experiment. The disyllabic words were by far the most frequent ones. To avoid a possible confound in the measurements, all and only disyllabic subjects and objects without article or adjective were compared. Altogether 121 sentences were retained for the analysis of duration, which contained 70 focused subjects and 51 focused objects, as shown in Table 3.

Table 3. Analysed Hindi sentences in the short version (121)

	Subject	**Object**
Information	23	22
Selection	26	17
Correction	21	12
Sum	70	51

Figure 4 illustrates the results for all disyllabic focused subjects and objects, as well as all disyllabic given subjects and objects in the two word orders. As can be gathered from the comparison, there is no significant effect of focus on duration: rather, irrespective of focus and word order, the first argument is longer than the second one. Overall, the arguments are shorter in OSV order compared to SOV. In SOV word order, both arguments are shorter when the first argument (i.e. the subject) is focused.

A linear mixed effects model with speaker and item as random effects reveals a marginally significant effect of constituent (Est.:−0.0651, Std.Err: 0.0336, t = −1.935), reflecting the longer duration of the first argument. There is a significant main effect of word order (Est.: 0.0606, Std.Err: 0.025, t = 2.409), confirming that durations of OSV sentences are shorter. The marginally significant effect of focus position (Est.: 0.0639, Std.Err: 0.0377, t = 1.695) suggests that both arguments are longer when the second one is in focus. The interaction of constituent and focus position (i.e. the effect of focus per se) is non-significant (Est.: 0.0015, Std.Err.: 0.0209, t = 0.073). In fact, discarding the interaction of constituent and focus position does not worsen model fit ($X^2 = 0.015$, p = 0.903). That is, the duration data remain inconclusive with respect to the question of whether arguments under focus are lengthened (as would be expected from the Patil et al. 2008 study on Hindi).

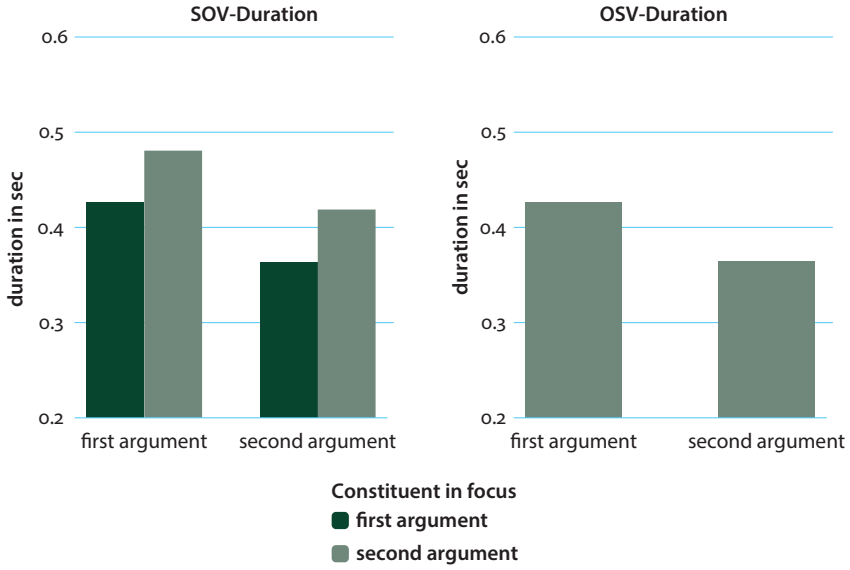


Figure 4. Hindi: Duration on arguments (average across speakers and items)

Durational effects were measured here in a rather coarse way, but more study is needed in order to understand better whether focus really affects duration in Hindi. Note that Patil et al. (2008) and Puri (2013) report durational effects of focus (longer duration on focused as opposed to given constituents). It is quite possible that such an effect is blurred here by the exclusion of monosyllabic and trisyllabic constituents, by the imbalanced data set, or by the generally more spontaneous elicitation procedure. In any case, the onset of a word is often perceptively much longer in the focused version. This needs further investigation, which we cannot pursue here.

4.2.4 *Summary*

To sum up this section, the results of Patil et al (2008) were confirmed. Focus in Hindi is not accompanied by as clear acoustic correlates typical of pitch accents that are commonly observed in languages like English and German. In these latter languages, a narrow focus is accompanied by a clear increase of pitch height, and by a tonal post-focal compression. In Hindi, only a small increase of pitch height could be measured and a moderate post-focal compression. Analysing the contour as a 'hammock' consisting of a high, a low and a high tone, the change due to focus is reduced to the following effect: the last pitch value of the focused constituent is higher than when the same constituent is given; this is equivalent to observing that the last rise of the prosodic phrase is higher when the prosodic phrase contains a focused word. The changes in the pitch values of given constituents are also weaker than in Germanic languages: the post-focal compression in Hindi is real and measurable, but it is not as important as in these languages.

There is a consistent increase in intensity in the focused constituent. The intensity of the first constituent is always larger than the second constituent, but the difference between intensity of the first and intensity of the second constituent is significantly smaller when the second constituent is focused.

No effect of duration was found in our data.

At an impressionistic level, we often find hammock and amplitude together, and occasionally we find only one of these correlates. Another observation is related to the possibly complementary use of pitch increase and pitch hammock. Pitch is used more often for emotional expressions such as agitation, excitement etc. than for strictly linguistic prominence placement. Thus in non-emotional contexts, the pitch hammock could be expected to replace pitch accent in Hindi, along with amplitude. Obviously, the small-scale experiment reported here does not allow us to formulate strong claims in this respect. Additional studies are needed to verify this hypothesis.

4.3 Correlates of phrasing of the focused constituent

The analysis of acoustic correlates associated with a focused constituent deliver only weak results in terms of the association of focus with a pitch accent, which contrast with the larger effects found in Germanic languages. However, Hindi speakers do have the feeling that prosody is operational in highlighting a focused constituent in a sentence.

And indeed, there are additional effects to be found in the Hindi data, which relate to prosodic phrasing of the focused constituent. In this section, we review the properties found in the data examined, which tend to show that phrasing of a

constituent is enhanced when the constituent is focused, and that phrasing is less clear or blurred when the constituent is given.

4.3.1 *Occurrence of the indefinite article e:k*

In some sentences, the numeral *e:k* 'a, one' which serves as indefinite article accompanies the subject or the object. This article is optional and, as a result, is not always used in our data. We find a correlation of its presence with focus, as illustrated in Table 4 and 5. Table 4 gives the raw quantities of *e:k* and Table 5 the probability of its occurrence. It is visible from these tables that *e:k* is used much more frequently when a noun is focused than when it is given: the proportion is 113 to 31. It is also more frequent on a subject than on an object (88 to 56). The highest occurrence of *e:k* is thus on a focused subject (69 instances) and the lowest on a given object (12 instances).

Table 4. Number of articles *e:k* in the corpus

	Focused	**Given**	**Sum**
Subject	69	19	88
Object	44	12	56
Sum	113	31	144

Table 5. Probability of *e:k*

	Focused	**Given**	**Average**
Subject	.92	.30	.61
Object	.69	.16	.43
Average	.81	.23	.52

4.3.2 *Encliticisation of e:k 'one/a'*

In many cases, *e:k* was prosodically encliticised to the preceding noun even if it syntactically belonged to the following noun. When it happened, *e:k* was part of the contour of the preceding constituent. Encliticisation happened 27 times on the subject and 16 times on the object, as shown in Table 6. Note that in the case of the subject, the article could generally only encliticise to the locational adverb or clause preceding the answer. This locational expression was always part of the question, and the participants just repeated it. When the locational expression was lacking, the expression started with the subject, in which case there was no possibility for the article to encliticise to preceding material.

Table 6. Number of encliticization of *e:k* in the corpus

	Encliticization of *e:k*
Focused S	27
Focused O	16
Sum	43

Encliticisation of *e:k* is illustrated with the sentence (5) in Figure 5, with word order OSV. The article *e:k* is cliticised to the preceding object *ləʈki:ko:* 'girl.DAT' and is separated from its following nominal head *a:dmi:* 'man' by a glottal stop. Both the object and the subject are realized with a rising contour, but the article *e:k* is realized as a prolongation of the rising contour of *ləʈki:ko:*. In this sentence, the subject *a:dmi:* is downstepped relative to the object, and it starts low. The perceptive impression of emphasis provided by separating the noun from the preceding *e:k* in this way is quite effective. A hammock is formed by the high pitch on *e:k*, the following low tone at the beginning of *a:dmi:* and the high tone at the end of this word.

(5) Information subject {Who is watching a girl?}
ʋəha: jəhã: a:ka:ʃ mẽ: ba:dəl hæ: ləʈki: ko: e:k ʔa:dmi: de:kh
there where sky LOC cloud-PL be-PL girl DAT one man see
rəha: hæ:
AUX-PROGR-M be-PS
'Where there are clouds in the sky, a man is watching a girl.'

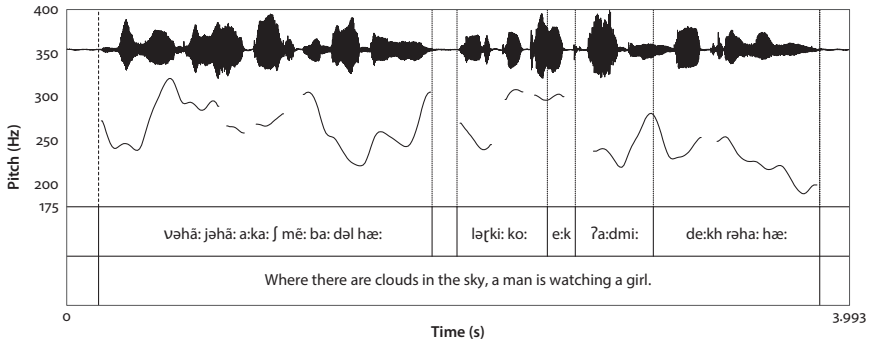


Figure 5. Encliticisation of *e:k* in an information subject sentence (1.2.9. OSV word order)

The same strategy can be illustrated with an example of object focus (6). In this case, the subject preceding the object served as a host for encliticisation. As illustrated in the corresponding Figure 6, it is clearly visible that the subject *əorət* 'woman' has a rising contour including the article *e:k*, which is part of the following object. This word reaches the highest peak of the phrase. Between *e:k* and the object *pe:ʈ-ko:* 'tree.DAT', there is a short phrase break. The object forms a coherent

contour with the verbal complex *ma:r rəhi hai* 'is hitting', which is low and flat. The oscillogram reveals that the object has relatively high intensity.

(6) Selection object {Is the woman hitting a tree or a flower?}
vəhã: ʤəhã: ni:la: a:ka:ʃ hæ: ɘorət e:k pe:ɽ ko: ma:r rəhi hæ:
there where blue sky be-3s woman one tree DAT hit AUX-PROG-F be-3s
'There where the sky is blue, the woman is hitting a tree.'

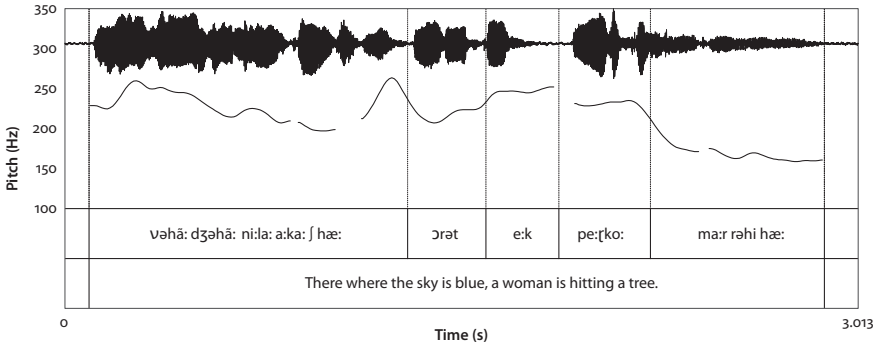


Figure 6. Encliticisation of *e:k* in a selection object sentence (3.9.5)

Encliticisation of the article and the following break enhance the separation (and thus: highlighting) of the focused constituent. It is optional, and, as shown above, it was used 43 times in the data, thus quite frequently. It rarely appeared on the given constituent.

4.3.3 *Glottal stop*

A third strategy for highlighting a focused constituent is insertion of a glottal stop. This strategy was usually found before a vowel-initial focused element, and it was also occasionally found in given constituents. The proportion of occurrences in the focused constituents is actually quite large, as can be seen from Table 7. It was inserted 77 times before a focused noun and 39 times before the article of a focused noun. Notice that it sometimes appears twice in the nominal phrase as in *ʔe:k ʔa:dmi* 'a man'.

Table 7. Number of insertions of a glottal stop in the corpus

	Insertion of glottal stop before the noun	**Insertion of glottal stop before the article**
Focused subject	55	34
Focused object	22	5
Sum	77	39

Glottal stop epenthesis sometimes correlates with encliticisation, in which case the article is encliticised to the left and the noun is separated by a glottal stop. But in many cases, it is just a different way to enhance a focus word, and to phrase it individually. The next illustration shows a sentence with focus on the subject (7). In this case, the indefinite article accompanying the subject is not cliticised to the word preceding it, which is a locational adverb, but it is nevertheless prosodically separated from *ɔ̀rət* 'woman', the word it syntactically attaches to. This strategy provides the impression that *e:k* is phrased individually.

(7) Selection subject {Is a man or a woman hitting the man?}
ʤəhã: ni:la: a:ka:ʃ hæ: ʋəhã: ʔe:k ʔəorət a:dmi: ko: ma:r
where blue sky be-3s there one woman man DAT hit
rəhi: hæ:
AUX-PROG-F be-3S
'Where the sky is blue, there, a woman is hitting a man.'

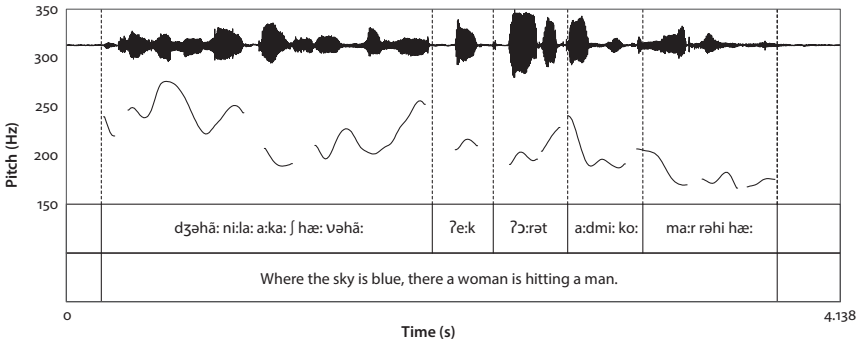


Figure 7. Glottal stop insertion in a subject selection sentence (2.5.1)

In the next example (8) illustrated in Figure 8, *e:k a:dmi*: 'man' is used twice, once as a focused subject and once as a given object, and as a result, it appears as a minimal pair. In the first occurrence, a glottal stop is inserted between the indefinite article and the noun. Moreover there is a break between *e:k* and the following noun. The beginning of *ʔa:dmi:* sounds prominent and loud. When the same sequence appears as the given object of the sentence, it is realized with the same correlates (glottal stop, small break, intensity on the first vowel) but in a reduced way: *e:k* is reduced to a very short vowel plus *k*. It is also enclitised to the preceding noun.

(8) Selection subject {Is it a man or a woman who is killing the man?}
[e:k ʔa:dmi:]$_F$ [e:k ʔa:dmi: ko: ma:r rəha: hæ:] (subject selection)
oneman one man DAT hitting/killing is
'{In the stony house,} a man is killing a man.'

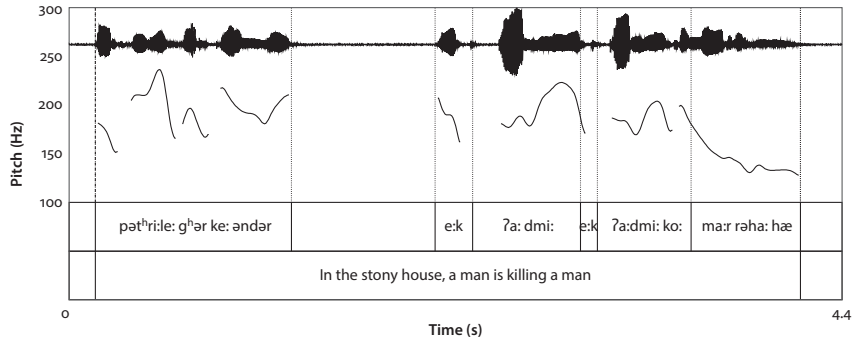


Figure 8. Glottal stop insertion in a subject selection sentence (4.13.16)

4.4 Summary for Hindi

To sum up the results of the acoustic correlates of focus in the Hindi experiment, we can confirm the results of Patil et al. (2008). We found a small increase of F0 and intensity as a result of focus, as well as a small amount of post-focal compression in pitch and intensity accompanying givenness. These correlates are not as large as in Germanic languages. We deduced that Hindi does not use pitch accents as a correlate of focus in the same way as Germanic languages do. One possible interpretation of the results is that lexical stresses are only weakly implemented in intonation and are overwritten by tonal properties at the level of the prosodic phrase. This suggestion has to be tested experimentally with appropriate data.

Correlates of phrasing at the level of the focused constituent are indicative of an effect of phrasing for focus: a boundary at the left edge of the focused constituent is realized by different means which are not all prosodic: i. occurrences of the article *eːk* 'a,' ii. cliticization of this article to the preceding constituent although, morphologically, the article belongs to the following constituent, and iii. glottal stop insertion before the noun and before the article. These three strategies have the same result: they enhance the separation of the focused constituent from the rest of the sentence and, in this way, increase its perceptive prominence.

5. Indian English

5.1 Data

The same experiment 'Anima' as described in Section 4 was conducted with twenty speakers of Indian English (11 female and 9 male), all of whom are also speakers of Hindi. The participants were recorded at Jawaharlal Nehru University (JNU),

Delhi in the same conditions and with the same recording devices. Most of the speakers were students. They all came from Delhi and spoke the same dialect of Indian English.

All answers were uttered in the SOV order. As was done for the Hindi data, subjects and objects of all utterances were cut and labelled with the help of undergraduate students in Frankfurt. Both Subject and Object were present in all sentences; in other words, there was not a single case of ellipsis in the elicited speech.[3] Apart from subject and object, the constituents of interest included pre-nominal articles, and in some rare case an adjective, mostly *another.*

From the 240 elicited sentences, 211 sentences were considered for analysis, 114 with focus on the subject, and 97 with focus on the object, see Table 8. 29 sentences were not considered for analysis.

Table 8. Analysed Indian English sentences in the long versions (211)

	Subject	**Object**
Information	42	39
Selection	36	41
Correction	36	17
Sum	114	97

The same pictures were used as for Hindi, and as explained above, in one of the pictures, a patient was ambiguous. It could be a boy or a girl. In the case of information or selection, this was not a problem, as all sentences, regardless of whether they contained *boy* or *girl* as an answer, were considered, except for one in which the speaker hesitated too long before answering. But in the case of correction, some speakers did not correct the object, but instead confirmed the question. Two such sentences were eliminated.

Some other questions eliciting a correction were not corrected, and instead, a confirmation was produced. These sentences were eliminated as well. In one case, the roles of the agent and the patient were inversed. This sentence, produced as a correction, was also eliminated because it is not clear which of the roles was corrected.

Further, a question eliciting a correction has been answered only negatively (*No, the man is not pulling a chair*), and no correction was provided. This sentence was also eliminated.

Some wrong answers were not eliminated, especially those eliciting an information focus, asked with a wh-question. As already commented above for Hindi,

3. We suspect that this homogeneity was the result of the slightly directive instructions given by the graduate student who conducted the recordings.

we were interested in the form of the answer and not in their correctness. But some wrong answers were eliminated, in which the answer was considered too deviant. For instance, as an answer to *Is the man pulling a table or a chair?*, one informant answered *'The man is pulling a woman.'* This sentence was eliminated.

The remaining eliminated sentences contained disfluencies.

As for the syntactic variation, there were 11 sentences containing a cleft construction, all of them with focus on the subject. An example of cleft sentence in a selection subject appears in (9).

(9) In front of the well, it is the man who is pushing the car.

In two of these cleft sentences, the subject was introduced and the action was described with the help of a nonfinite participle construction, as illustrated in (10), a selection subject.

(10) It is a man cutting the watermelon

Most elicited answers were syntactically very simple, and contained only the subject, the verb and the object. An untypically complex sentence is reproduced in (11), with a relative clause and a negation, in the context of a subject correction.

(11) Inside the stone house, there is a man who is killing another man, not a woman.

5.2 Results of the acoustic analysis

In this section, the results of the acoustic analysis are discussed, starting with the results for F0.

5.2.1 *Pitch contour (F0)*

As can be observed in Figure 9, in case of subject focus, there is no clear reflection of focus on the subject itself by means of F0. By contrast, there is an F0 difference on the object: the object had a significantly lower pitch contour when it is given than when it is focused. As can be seen in Figure 9, the entire domain from the end of the subject through the verb to the end of the object is lowered when focus is on the subject as compared to when it is on the object. This strongly suggests (moderate) post-focal compression in the case of subject focus.

This interpretation is backed up by a linear mixed model: as in the case of the Hindi data, we used focus position and constituent and their interaction as fixed effects (word order was SVO throughout, so this effect was not included). Speaker, item, and interval position were used as random effects (to make the analysis com-

parable to the Hindi data, we only used measurements from the subject and object intervals, data points from the verb were disregarded in the models).

This model yields three significant main effects: First, overall mean F0 is higher in sentences with object focus (Est.: −5.319, Std.Err.: 2.189, t-value = −2.43), which is likely due to post-focal compression on the object in this condition. Secondly, there is a strong effect of constituent (Est.: 25.954, Std.Err: 5.202, t = 4.99), reflecting that F0 on the sentence-initial subject is higher than on the sentence-final object. Importantly, the interaction is significant, too (Est.: 5.231, Std.Err: 1.411, t = 3.71), confirming that the F0 on the sentence final object is more strongly affected by focus (again, due to post-focal compression) than the subject that does not appear to show clear effects of focus. Model comparison justifies the inclusion of each of the factors, as their removal would deteriorate model fit considerably (all p-values < 0.001).

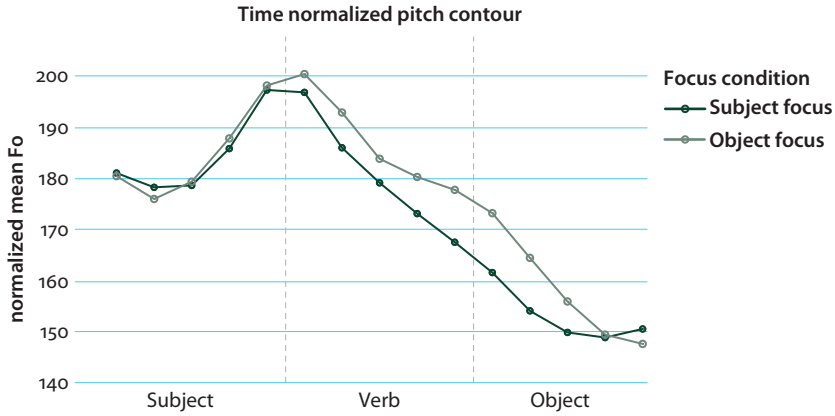


Figure 9. Indian English: Time normalized pitch contour (average across speakers and items)

Comparing Figure 9 with Figure 2 for Hindi, the similarity between the pitch contours of the first argument is striking. In both cases, there is a moderate fall before the low tone, and the following rise reaches a higher level, even more so in Hindi than in Indian English. On the other hand, the difference in the form of the pitch contours on the second argument is also evident. The rising pattern identified in Hindi for both arguments, focused and given alike, is absent in the case of objects in Indian English, which show a falling contour. The difference is mainly due to the final position of the object in Indian English, as compared to the non-final position in Hindi. Since the object in IE is sentence-final it is bound to be accompanied with the final falling contour typical of a declarative sentence. However, this is not the whole story. In some cases, there is also a difference between the contour of the verb, which is displaying just a plain fall in the subject focus sentences, but not so

in the object focus sentences.[4] It will be shown below that a frequent pattern in this case is an additional high boundary tone on the verb, guaranteeing in this way that the object is phrased separately, see Figure 13 for illustration. This additional high boundary tone influences the contour of the verb and object observed in Figure 9.

5.2.2 *Intensity*

Intensity, illustrated in Figure 10, is significantly higher on the first argument than on the second one, reflecting an overall decrease in intensity over the course of the utterance (main effect for constituent: Est: −9.0484, Std.Err: 0.3978, t = −22.75). Overall, the effect of focus on intensity is non-significant, i.e. there is no difference in overall loudness between sentences with subject focus and those with object focus (main effect of focus: Est.: −0.398, Std.Err: 0.706, t = −0.56). However, a significant interaction clearly reveals that focused constituents are produced with higher intensity than given ones (Est.: 2.012, Std.Err: 0.589, t = 3.42). Note, however, that the effect is distinctly larger in the case of object focus compared to subject focus. A possible interpretation of this difference is a ceiling effect on the subject in initial position: it is much louder in all cases, and there is thus not much margin to increase its intensity while speakers have more freedom to vary intensity on the object.

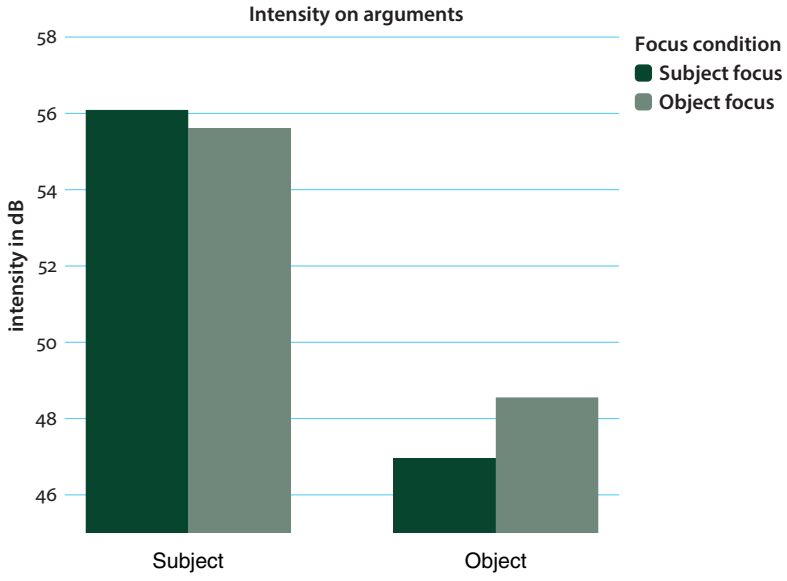


Figure 10. Indian English: intensity on arguments (average across speakers and items)

4. This difference is obscured in Figure 9, which depicts averaged data.

5.2.3 *Duration*

Turning next to the results for duration, a corpus of constituents with equal length was prepared for Indian English, as was shown for Hindi. In our corpus of Indian English, monosyllabic nouns were by far the most frequent ones. To make the duration analysis comparable, only the monosyllabic nouns were measured, to the exclusion of articles and adjectives. There were 38 sentences containing a monosyllabic subject and 59 sentences containing a monosyllabic object. Altogether 97 nouns entered the analysis.

Table 9. Analysed Indian English sentences in the short versions (97)

	Subject	Object
Information	10	26
Selection	21	18
Correction	7	15
Sum	38	59

The most conspicuous result, shown in Figure 11, is that the utterance-final object is always longer than the subject, probably as a result of final lengthening (Estimate: 0.1245, Std.Err: 0.0202, t = 6.156). This result differs from Hindi, where the first constituent was longer than the second. A significant interaction between focus and constituent confirms that duration increases on focused constituents compared to given ones (Est.: −0.079, Std.Err: 0.027, t = −2.875). There is no overall effect of focus position on duration (Est.: 0.041, Std.Err.: 0.027, t = 1.510).

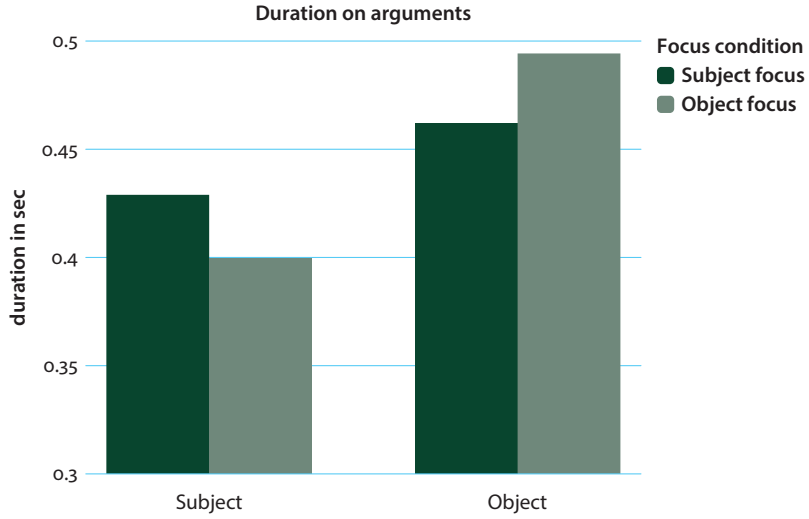


Figure 11. Indian English: duration on monosyllabic arguments (averaged across speakers and items).

5.2.4 *Summary*

Summing up this section, the results showed altogether a similar picture to what we found for Hindi, despite the difference in F0 and in duration. The results for the first argument (the subject in IE) were very similar in the two languages, while the results for the second argument differed substantially. However, we assume that the differences in the case of the second argument are due to word order. The second argument precedes the verb in Hindi, and is thus non-final in the sentence, whereas the object is final in Indian English. Summing up the results of the acoustic measurements, we showed the following. In the case of subject, there was no increase in F0 as a result of focus, and in the case of object, only a minor increase could be observed. The results also revealed a small but significant post-focal compression (10–15 Hz in average), and no real deaccenting. As for intensity, the subject was always much louder than the object. Moreover, the object was louder when focused than in the given condition, whereas the effect of focus on intensity was non-significant for the subject. Finally, two effects could be identified for duration. First, the object was longer than the subject probably because of final lengthening. And second a focused constituent was longer than a corresponding given one.

5.3 Other correlates of the focused constituent

5.3.1 *Phrasing*

As was the case for Hindi, the perceptive impression of a separate phrasing on the focused constituent is paramount in Indian English. We investigate further correlates of phrasing in this section.

In some sentences with subject focus, as for instance in (12), the auxiliary is cliticized to the preceding constituent, the subject in our data. In Figure 12, there is a rising contour on the subject *woman* going on well into the vowel of the following auxiliary *is*. It will be shown in Section 6 that in the sentences with subject focus, the verb and the object are phrased in a single prosodic phrase with a falling contour. There are two adjacent high tones, one at the end of the subject, the other one at the beginning of the VP, which are in a downstep relation to each other. Moreover, some amount of post-focal compression may be present, as indicated by the phonetic data in the preceding section.

In British English the focused constituent, i.e. the subject in this sentence, generally correlates with a falling contour (Pierrehumbert 1980, Liberman & Pierrehumbert 1984, Steedman 2000 and many others), and deaccenting of the entire post-focal domain is the consequence. As can be gathered from Figure 12, this is not the case in Indian English. The auditory impression coming from the long rising contour on the focused subject plus the auxiliary is very different. The result is that, at least to the ear of someone familiar with Germanic intonation,

the following fall might be perceived as accented, especially since the fall is rather smooth and interpolating between the high tone of the first phrase and the final low tone (see next section).

(12) Correction subject {Is a man hitting the man?}
No, a woman is hitting the man

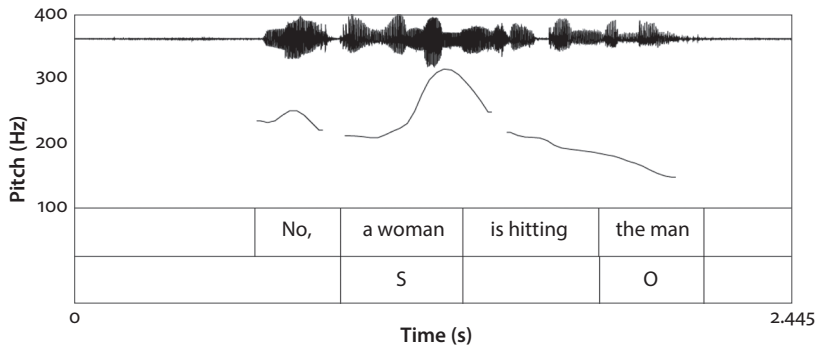


Figure 12. 'A woman is hitting the man.'

In some sentences with object focus, as for instance in sentence (13), illustrated in Figure 13, the auxiliary is cliticised to the left, as shown above for Hindi. The verb (*carrying*) has a low tone on its first syllable and a high tone at the boundary between *carrying* and the object delimiting a phrase with a final high tone. Notice that the focused object constituent *a girl* does not carry a special contour signalling a pitch accent. On the contrary, it has the same falling contour indicating its finality as the object in Figure 12. The correlate of object focus appears on the boundary tone of the verb, separating it from the following focused constituent.

(13) Selection object{Is the man carrying a boy or a girl?}
The man is carrying a girl

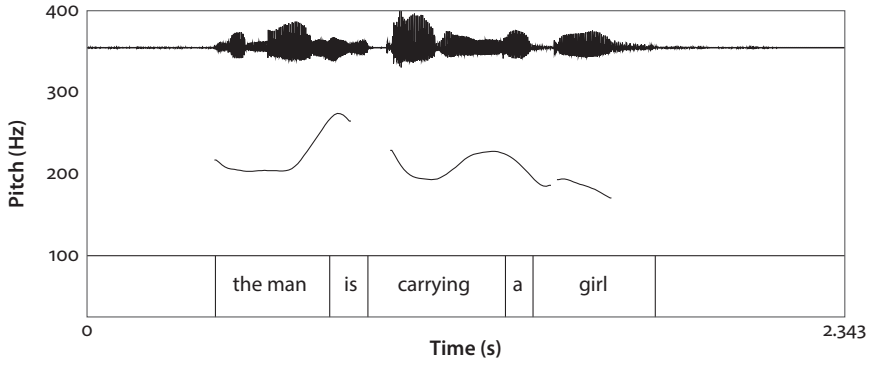


Figure 13. Indian English: 'The man is carrying a girl.'

5.3.2 *Indefinite article and glottal stop insertion*

Compare next the distribution of definite and indefinite articles as a function of the focused or given status of the subject and object. When a constituent was focused, it was generally introduced by the indefinite article *a/an* and less so by the definite article *the*, see Table 10 for comparison. The opposite distribution of the articles appeared on given constituents.

Table 10. Distribution of the definite and indefinite article

	Indefinite Article (*a/an*)	**Definite Article (*the*)**
Focused Subject	90	24
Focused Object	87	10
Given Subject	6	91
Given Object	15	99

As we saw for Hindi, the presence of a glottal stop enhances a phrase boundary. This strategy is speaker-dependent: some speakers use glottal stops more regularly than others, and some speakers did not use it at all. A focused subject with the indefinite article was in most cases preceded by a glottal stop on the indefinite article. In a non-negligible number of VPs, the auxiliary *is* was preceded by a glottal stop: 19 in sentences with focused subject and 18 in sentences with focused object, see Table 11. And a focused object was also often preceded by a glottal stop, although only half so often as a subject. This difference is expected if the subject is always phrase-initial and the object is only exceptionally phrase-initial, as will be shown in the next section.

Table 11. Distribution of glottal stop

	Glottal stop before the Subject	**Glottal stop before the V (VP)**	**Glottal stop before the Object**
Focused Subject	62	19	5
Focused Object	4	18	30

5.4 Summary for Indian English

As far as the acoustic correlates of focus are concerned, it could be observed that in the sentence-initial position, constituents have the same rising contour as the one found in Hindi. As in Hindi and British English, a declarative sentence ends with a falling tone. But a major difference with British English is that the last falling tone can be located on the post-focal element, as illustrated in Figure 12, denying the hypothesis that the focus of a sentence carries the sentence nuclear pitch accent, as is the case in British English. We also observed that intensity was increased in a

focused constituent as compared to a given one. These results resemble those for Hindi. Additionally, duration was increased as well. This differs from Hindi, and this difference may be due to the fact that the position of the object is not the same in the two languages. It is prefinal in Hindi and final in Indian English.

Correlates of phrasing like encliticisation of the article, glottal stop insertion, and additional tonal contours speak for insertion of an occasional phrase boundary to the left of the focused constituent. Again, this strategy is similar to focus-triggered phrasing in Hindi.

6. Discussion, phonological analysis, and conclusion

In this section, the questions raised in Section 1 are answered. First, the phonetic, prosodic, and segmental correlates of focus in both Hindi and in Indian English mainly suggest a difference in phrasing but only marginally a difference in prominence (cf Sections 4.4 and 5.4). The survey of phonetic correlates in Hindi answers the question about the validity of lab speech. The tasks were similar to those employed by Patil et al. (2008) with the difference that Patil et al. used scripted speech, while the present paper used unscripted, semi-spontaneous utterances. In both cases, the measurements of F0 and intensity delivered very similar results. This implies that the behaviour of speakers when performing scripted speech is similar to their behaviour during spontaneous speech production, a welcome result confirming the ecological validity of scripted speech. Moreover, the unscripted data revealed that cliticization and segmental correlates should be considered as well when looking for correlates of focus. The large amount of variation obtained in this kind of speech can be a source for discovering new factors. Extracting prosodic generalizations from spontaneous data remains a realistic task as long as the data obtained are controlled and compared with lab data.

In the present experiment, data with narrow focus on one of the constituents were compared to the same data in which the constituent was given, and we used Patil et al. for a comparison with the wide focus context. It was shown there that narrow focus does not change the pitch excursion on the object, but that it does change the excursion on the subject. Givenness had a post-focal effect, but no pre-focal effect. The intention of the present experiment was thus not to reproduce the effects of Patil et al. but to find additional correlates of information structure that are only to be found in (semi-)spontaneous data.

Turning to the phonological model for the data, the third question, a clear finding of this paper is that in both languages the low and high tones are functioning as phrase tones or boundary tones, see the phonological model illustrated. Focus is expressed primarily by correlates of phrasing, as predicted by the model

illustrated in (3) and (4), and not so much by pitch accents, as predicted by the prominence model. An open question that we do not try to answer here because of lack of relevant data is whether these tones coincide with lexical stresses or whether the two categories are largely independent from each other. It could be that Hindi and Indian English differ in this respect.

Starting with Hindi, every non-final low tone is analysed as aligned with the beginning of a non-final prosodic phrase (a prosodic phrase is called Φ-phrase, following recent literature) and every high tone as aligned with the end of a non-final prosodic phrase (Gussenhoven 2004: 147ff for tone alignment and tone association). Tones which are aligned with prosodic domains are not predicted to be associated with particular syllables, and a mild amount of variation in the ultimate association of tones with syllables is predicted. We suggest that, in default phrasing of Hindi sentences, the subject forms its own Φ-phrase, and the verb plus the object are phrased together in a second Φ-phrase. Additionally, the object is also phrased individually and is forming a recursive Φ-phrase with the verb, as illustrated in (14) and (15).[5] This phrasing is motivated by the scaling of the tonal excursions. The high tones of an intonation phrase (that we call ι-phrase) are in a downstep relation: a high tone is scaled lower than a preceding high tone in the same ι-phrase, as visible in Figures 2, 5, 7, and 8.

(14) L_Φ H_Φ L_Φ H_Φ L_ι
[(e:k ʔa:dmi:)$_\Phi$ ((e:k a:dmi: ko:)$_\Phi$ ma:r rəha: hæ:)$_\Phi$]$_\iota$
one man one man DAT hitting/killing is
'A man is killing a man.'

It was shown that the article *e:k* is in some cases enclitized to the preceding constituent, and that enclitisation is more likely in the case of a following focused noun. Encliticisation can be considered as rephrasing of a function word, the result of which is a phrase boundary just before the focus. This is illustrated in (15).

(15) L_Φ H_Φ L_Φ H_Φ L_ι
[(əorət e:k)$_\Phi$ ((pe:ɽ ko:)$_\Phi$ ma:r rəhi hæ:)$_\Phi$]$_\iota$
woman one tree DAT hit AUX be-3S
'The woman is hitting a tree.'

Turning to Indian English, it is proposed that the phonological phrasing is similar to Hindi. However, there is an optional difference between the sentences with focus on the subject and those with focus on the object. In the former cases, the phrasing

5. It has sometimes been proposed that every content word is accompanied by a low and a high tone. However these tones can be interpreted as delimiting embedded Φ-phrases. This analysis is in line with the observation that not all content words are delimited by tones.

and tonal structure shown in (16) for sentence (12), illustrated in Figure 12, is sufficient to explain the phonetic correlates identified in the preceding section.

(16) Correction focus {Is a man hitting the man?}
L_{Φ} H_{Φ} L_{ι}
$[(\text{a woman})_{\Phi}\ (\text{is hitting}\ (\text{the man})_{\Phi})_{\Phi}]_{\iota}$

In the case of focus on the object, there may be an additional phrase on the verb. As a result, the focused word *girl* in (17), illustrated in Figure 13, is separated from the verb by a boundary tone. In this case, the verb ends with a high boundary tone and the sentence consists of three Φ-phrases. The high tone at the end of the verb *carrying* can only be a boundary tone delimiting the verb from its following focused object. There is no possible analysis in which this high tone is a pitch accent. In this case, it is visible that focus is changing the prosodic phrasing of the sentence.

(17) Selection object {Is the man carrying a boy or a girl?}
L_{Φ} H_{Φ} L_{Φ} H_{Φ} L_{ι}
$[(\text{The man is})_{\Phi}\ (\text{carrying})_{\Phi}\ (\text{a girl})_{\Phi}]_{\iota}$

One speaker followed the same strategy in Indian English that was observed above for Hindi to enhance the phrasing of the focused constituent. She phrased the article with the verb, realizing a very high boundary tone on the article and making a short break before continuing. An illustration appeared in Figure 14. The proposed phrasing is shown in (18) which is similar to (17) with the exception that the definite article of the final object is part of the phrase of the verb, allowing in this way that the noun *plant* is phrased individually. The same strategies of encliticization and of tonal marking as before are used. Only the register and the duration relationships are changed. The high tone on *the* is much higher than the comparable high tones in (13), illustrated in Figure 13.

(18) Information object{What is the woman hitting?}
L_{Φ} H_{Φ} L_{Φ} H_{Φ} L_{ι}
$[(\text{The woman is})_{\Phi}\ (\text{hitting the})_{\Phi}\ (\text{plant})_{\Phi}]_{\iota}$

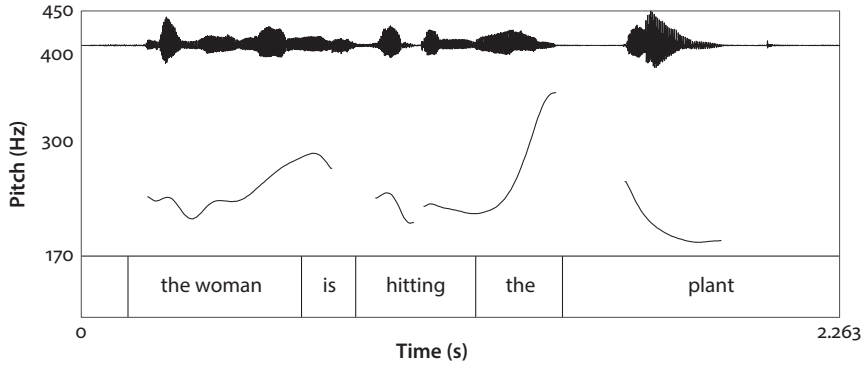


Figure 14. 'The woman is hitting the plant' Inf O 1.2.5

Even if the pattern illustrated in Figure 14 was exceptional in our data for Indian English, it is quite common in everyday conversation. The next example, illustrated in (19) and in Figure 15, is taken from a (recorded) conversation between the first author and a student at JNU in March 2011.[6] The word *to* belongs syntactically to the embedded infinite clause, but prosodically it is encliticized to the main clause. The conversation is only a few minutes long, but several occurrences of encliticizations were produced by this speaker. In (19), the phrasing and the tones are indicated.

(19) L_Φ H_Φ L_Φ H_Φ L_Φ H_Φ
[(that he is able to)$_\Phi$(speak both (languages)$_\Phi$(fluently)$_\Phi$]$_\iota$

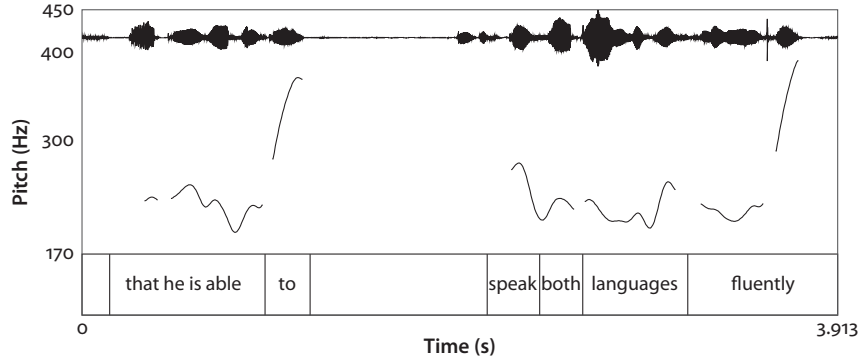


Figure 15. From a conversation: cliticisation of 'to' to the left

6. The longer sentence from which the example is taken is as follows: 'Between his L1 and L2 he is just trying to adjust the constraints in such a way that he is able to speak both languages fluently, and there is a stage when the constraint hierarchy is a little awkward which can be called the interlanguage.'

Consider now the predictions of the prominence analysis, illustrated in (1) and (2). This approach implies that focus is reflected by added prominence on the lexical stress of the focused word. However, this is not confirmed by the data: as demonstrated above the focused word does not get additional prominence. What we see in (6) and in (18), for instance, is a different strategy: the focus is clearly separated from what precedes by a boundary tone indicative of a phrase delimitation on the preceding material.

Postfocal compression is present in both languages. However, its amount is much less in Indian English and Hindi than in British English.

As has been mentioned throughout this paper, the analysis proposed here can only be tentative. Quite a number of important insights are still missing before a satisfying analysis of the intonation pattern of Hindi and Indian English can be completed. During the past decades, intonation research has been intensively conducted with particular kinds of languages, namely those using pitch accents associated with lexically stressed syllables for focus, like English and German. These languages fulfil the axiom of prominence established by Jackendoff (1972) and other researchers after him, who showed that a focus is prosodically realized by a syntactically or semantically determined pitch accent, alternatively high or low in pitch, depending on the context, and by deaccenting of the post-focal material.

The languages investigated in this article have been shown to have different properties, which leads us to the question raised in Section 1 as to the place of Hindi and Indian English in a typological system of intonation. Several researchers have emphasized the similarities between the suprasegmental properties of different South Asian languages and their difference from British or American English. Kachru (1983) and Mahboob & Ahmar (2004: 1014) emphasise that '*it is the suprasegmental features of South Asian English (SAsE) such as stress and rhythm, rather than segmental features, that mark its uniqueness. Kachru argues that while the segmental features of SAsE are heavily influenced by mother tongues and may therefore be different between various speakers, non-segmental features are shared*.' Even if on the surface, intonation alternates high and low tones, some of which are associated with prominent syllables and some others with boundaries, there is some evidence that the grammatical rules underlying the assignment of these tones differ greatly from language to language.

Languages choose between different possible word orders and optimize in this way the communication of information structural needs. They may also have different phonetic correlates to accompany either the edges of constituents or the focused exponents. Syntactic constituents are often mapped with prosodic constituents (see Cinque 1993, Zubizarreta 1994, Selkirk 1995, Reinhart 2006 among others) and it is this syntactic structure which is marked under focus. Patil et al. (2008) and Féry (2010) suggest that Hindi may be called a phrase language,

because some of the phonological strategies to mark focus are different from all classes of languages proposed by Hyman (1999), Ladd (1996/2008), Gussenhoven (2004) and others, intonation languages, tone languages and pitch accent languages, see also Féry (2016) for a typology of intonation. This article suggests that not only Hindi but also Indian English might be phrase languages.

Acknowledgements

We would like to thank Mahesh M., Meiraba and Shekhar Sudhanshu, graduate students at JNU, who were involved in the recording of the material used in this paper, and also Divvya Rathore, who provided glosses, transcriptions and translations for all sentences. Many thanks also to Katrin Müller, Anna Roth and Viviana Haase, the undergraduate students in Frankfurt who prepared the material for analysis. Thanks to the DFG which financed the cooperation between Pramod Pandey (JNU Delhi) and Caroline Féry (Goethe University, Frankfurt). The paper profited greatly from comments from two anonymous reviewers and from Hubert Truckenbrodt.

References

Astésano, C., E. Bard & A. Turk. 2007. Structural influences on initial accent placement in French. *Language and Speech* 50. 423–446. doi: 10.1177/00238309070500030501

Bansal, Ram Krishna. 1976. *The Intelligibility of Indian English*. 2nd edn. Monograph 4. Hyderabad: CIEFL.

Boersma, Paul & David Weenink. 2005. Praat: Doing phonetics by computer (Version 4.5).

Botinis, Antonis (ed.). 2000. *Intonation: Analysis, modeling and technology*. Dordrecht: Kluwer.

Cole, J., Y. Mo & S. Baek. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes* 25. 1141–1177. doi: 10.1080/01690960903525507

Cooper, William E., Stephen J. Eady & Pamela R. Mueller. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America* 77. 2142–2156. doi: 10.1121/1.392372

Cruttenden, Alan. 2006. The deaccenting of given information: A cognitive universal? In G. Bernini & M. L. Schwartz (eds.), *Pragmatic organization of discourse in the languages of Europe*, 311–356. The Hague: Mouton.

Dyrud, Lars O. 2001. *Hindi-Urdu: Stress accent or non-stress accent?* Grand Forks, ND: University of North Dakota PhD thesis.

Féry, Caroline. 2010. The intonation of Indian languages: An areal phenomenon. In Imtiaz Hasnain & Shreesh Chaudhury (eds.), *Problematizing language studies. Festschrift for Ramakant Agnihotri*, 288–312. New Delhi: Aakar Books.

Féry, Caroline. 2016. *Intonation and prosody*. Cambridge: Cambridge University Press.

Féry, Caroline & Frank Kügler. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36. 680–703. doi: 10.1016/j.wocn.2008.05.001

Gambhir, Vijay. 1981. *Syntactic restrictions and discourse functions of word order in standard Hindi*. Philadelphia: University of Pennsylvania PhD thesis.

Gargesh, Ravinder. 2004. Indian English: Phonology. In W. Edgar, Bernd Burridge, Rajend Mesthrie Kortmann & Clive Upton (eds.), *A handbook of varieties of English, Vol 1: Phonology*, 993–1002. Berlin: Mouton de Gruyter.

Gussenhoven, Carlos. 2008. Notions and subnotions of information structure. *Acta Linguistica Hungarica* 55: 231ff. doi: 10.1556/ALing.55.2008.3-4.11

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511616983

Harnsberger, James D. 1994. *Towards an intonational phonology of Hindi*. Ms., University of Florida.

Harnsberger, James D. 1999. The role of metrical structure in Hindi intonation. In South Asian Language. Analysis Roundtable, University of Illinois, Urbana-Champaign.

Harnsberger, James D. & Jasmeet Judge. 1996. Pitch range and focus in Hindi. *The Journal of the Acoustical Society of America* 99. 2493–2500. doi: 10.1121/1.415628

Hayes, Bruce. 1995. *Metrical stress theory. Principles and case studies*. Chicago: University of Chicago Press.

Hayes, Bruce & Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9. 47–96. doi: 10.1007/BF00133326

Hussain, Sarmad. 1997. *Phonetic correlates of lexical stress in Urdu*. Evanston, IL: Northwestern University PhD thesis.

Ito, Junko & Armin Mester. 2007. Categories and projections in prosodic structure. Paper presented at *Old World Conference in Phonology-4*, Rhodes, Greece.

Jyothi, Preethi, Jennifer Cole, Mark Hasegawa-Johnson & Vandana Puri. 2014. An investigation of prosody in Hindi narrative speech, Proceedings of Speech Prosody. http://www.ifp.illinois.edu/~pjyothi/files/SP2014.pdf. (Accessed 12 Jan 2016).

Kachru, Braj B. 1983. *The other tongue: English across cultures*. Urbana: University of Illinois Press.

Kachru, Yamuna. 1987. Hindi-Urdu. In B. Comrie (ed.), *The world's major languages*, 470–489. London: Routledge.

Keane, Elinor. 2007a. Distribution and alignment of F0 contours in Tamil. In *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbruecken, Germany.

Keane, Elinor. 2007b. Investigating the intonational phonology of Tamil. Poster presented at *the satellite meeting of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbruecken, Germany, August 5, 2007.

Khan, Sameer. 2007. Phrasing and focus in Bengali. Poster presented at *the satellite meeting of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbruecken, Germany, August 5, 2007.

Kidwai, Ayesha. 2000. *XP-adjunction in universal grammar: Scrambling and binding in Hindi-Urdu*. Oxford: Oxford University Press.

Kuno, Susumo. 1981. The syntax of comparative clauses. *Proceedings of the Chicago Linguistic Society* 17. 136–155.

Ladd, D. Robert. 1996/2008. *Intonational phonology*. Cambridge: Cambridge University Press.

Liberman, M. & J. Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In M. Aronoff & R.T. Oehrle (eds.), *Language sound structure*, 157–233. Cambridge: MIT Press.

Mahboob, Ahmar & Nadra Huma Ahmar. 2004. Pakistani English: Phonology. In W. Edgar, Bernd Burridge, Rajend Mesthrie Kortmann & Clive Upton (eds.), *A handbook of varieties of English, Vol 1: Phonology*, 1003–16. Berlin: Mouton de Gruyter.

Mo, Y., J. Cole & E-K. Lee. 2008. Naïve listeners' prominence and boundary perception. *Fourth Speech Prosody*. 735–736. http://www.isca-speech.org/archive/sp2008. (Accessed 5 Jan 2016).

Moore, Robert R. 1965. *A study of Hindi intonation*. Ann Arbor, MI: University of Michigan PhD thesis.

Nair, Rami. 1999. *Syllables and word edges*. Evanston, IL: Northwestern University PhD thesis.

Nair, Rami. 2001. Acoustic correlates of lexical stress in Hindi. In A. Abbi, R. Gupta & A. Kidwai (eds.), *Linguistic structure and language dynamics in South Asia — papers from the proceedings of SALA*. XVIII roundtable, 123–143.

Ohala, Manjari. 1986. A search for the phonetic correlates of Hindi stress. In B. Krishnamurti (ed.), *South Asian languages: Structure, convergence, and diglossia*, 81–92. Delhi: Motilal Banarsidass.

Pandey, Pramod. 1989. Word accentuation in Hindi. *Lingua* 77. 37–73. doi: 10.1016/0024-3841(89)90038-7

Pandey, Pramod K. 1994. On a description of the phonology of Indian English. In R. K. Agnihotri & A. L. Khanna (eds.), *Second language acquisition. Socio-cultural and linguistic aspects of English in India. Research in applied linguistics*, 198–207. Delhi: Sage.

Patil, Umesh, Gerrit Kentner, Anja Gollrad, Frank Kügler, Caroline Féry & Shravan Vasishth. 2008. Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics* 1. 53–70.

Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. Bloomington, IN: MIT PhD thesis.

Post, Brechtje. 2000. *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.

Puri, Vandana. 2013. *Intonation in Indian English and Hindi late and simul-taneous bilinguals*. Urbana-Champaign: University of Illinois at Urbana-Champaign PhD dissertation.

Reinhart, Tanja. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27. 53–94.

Rooth, Mats. 1985. *Association with focus*. Amherst: University of Massachusetts PhD thesis.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116. doi: 10.1007/BF02342617

Selkirk, Elisabeth. 2007. Contrastive focus, givenness and the unmarked status of discourse-new. In C. Féry, G. Fanselow & M. Krifka (eds.), *Interdisciplinary studies on information Structure (ISIS 6)*. 125–145.

Steedman, Mark 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 31. 649–689. doi: 10.1162/002438900554505

Truckenbrodt, Hubert. 1995. *Phonological phrases: Their relation to syntax, focus and prominence*. Cambridge, MA: MIT PhD thesis.

Yoon, T-J., S. Chavarría, J. Cole & M. Hasegawa-Johnson. 2004. Intertranscriber reliability of prosodic labeling on telephone conversation using ToBI. *Interspeech* 2004. 2729–2732.

Corresponding author's address

caroline.fery@gmail.com
Institut für Linguistik
Norbert-Wollheim-Platz 1
60629 Frankfurt
Germany