The prosodic expression of contrast in Hindi

Susanne Genzel [1], *Frank Kügler*[2]

[1] SFB 632 "Information structure", Potsdam University
[2] Department of Linguistics, Potsdam University
susonne7@gmail.com, kuegler@uni-potsdam.de

Abstract

This production study examines the prosodic means of encoding contrast in Hindi. Different target words were embedded in carrier sentences and were put in two information structural contexts, wide and contrastive focus. Contrary to what is expected from earlier findings Hindi uses prosodic means of expressing contrast on the focused word, namely an increase in pitch span and duration. These results may contribute to the understanding of intonational phonology and to the prosodic classification of Hindi.
Index Terms: Hindi, contrast, pitch span, duration

1 Introduction

Hindi is one of the world's most widely spoken languages with approximately 150 mill. native speakers in north and central India in the states Uttar Pradesh, Uttaranchal, Rajasthan, Bihar, Haryana, Madya Pradesh, Sikkim, Jharkhand Chattisgarh und Himachal Pradesh [1]. Hindi belongs to the Indo–Iranian branch of the Indo–European language family, and is genetically related to other European languages like German, English and Russian.

With respect to its prosodic properties most researchers agree that Hindi has lexical stress [2], [3], [4], [5], [6], [7]. The position of stress is driven by syllable weight, whereas the ultima is assumed to be extrametrical [8].

The few studies examining the prosody of Hindi agree that each content word except the last one of an intonation phrase is associated with a rise in pitch [2], [4], [9]. For instance, Moore [2] subdivides sentences into feet. The neutral distribution is one foot per content word, whereas polymorphematic words can be subdivided in two or more feet. According to Moore it is the foot that is prosodically realized with a F0 rise (LH). This indicates that Hindi demarcates the prosodic word prosodically. A study on word order and intonation in Hindi indicated that the constituents in a sentence appear to be in a strict downstep relation [9].

Apart from its demarcative function, intonation also functions as a means of highlighting information that in human communication a speaker transfers to a hearer (e.g. [10]). At the same time other information gets in the background. This function is known as prosodic focus [11], [12], [13]. A focus of a sentence represents a word or constituent that receives prominence either by means of morphology, syntax, prosody or combinations of these linguistic devices. Semantically, focus defines a set of alternatives from which one element is chosen by the speaker [14], [15]. If a focus highlights new information, parts of a sentence may contain already old information which is usually called given information, defined as previously mentioned in the discourse [16], [17]. The information structural notions important for the course of the paper are broad focus and contrastive focus. Following [11] we assume that "... if the focus constituent is the whole sentence, we get `normal stress'; if not, we get a narrow focus on the constituent identified by the placement of the accent". In the wide or broad focus, all information is new and the whole sentence is focused and serves as baseline condition, also sometimes referred to as out-of-the-blue utterance [19]. In the case of corrective or contrastive focus the speaker has chosen an element from a set of alternatives [14], [15]. The contrastive focus is narrow concerning size [13]. Based on the assumptions of Focus-Prominence-Theory (e.g. [20], [21]) the expression of abstract focus prominence is language specific. Prosodic reflexes due to information structure are well known for intonation languages like German, English and many more. Prominence due to focus is marked by a pitch-accent which shows longer duration and higher F0 whereas given information is deaccented [22], [23], [24]. The greatest amount of focal lengthening in German [25], English [26] and in Swedish [27] can be found in the stressed syllable.

Focus may also be expressed syntactically or morphologically, yet Hindi does not use morphological means to express focus. Hindi is a head-final language with SVO as base word order, and a focused constituent typically would occupy the immediately preverbal position [18]. Since this study is concerned with prosodic reflexes of focus we chose to study the prosody of focus on adjectives. An adjective belongs to a noun phrase, and cannot be split and moved into the preverbal focus position. This strategy ensures that any prosodic effect of focus appears with no other influences.

Previous studies on the effects of focus in Hindi have shown three effects: Greater pitch excursion on the rise, higher intensity and longer duration. In the post–focal region the rise is compressed [2], [4], [28]. In contrast, [9] revealed that a narrow focus in sentence medial position showed no difference in high tone scaling as compared to the all-new pattern. Given these contradictory findings this study will examine the prosodic expression of contrast in Hindi in more detail.

2 Method

2.1 Speech materials

Table 1 shows the different adjectives used as target words. The adjectives differed in number of syllables as well as in position of stress. These were put in carrier sentences with SVO structures modifying the object (1).

Wide focus was elicited without a context, i.e. out of the blue [19]. The contrastive focus elicitation consisted of a question-answer pair: a question and a response to the question (1). The focus domain is indicated by squared brackets and the F mark.

(1)

क्या मोहिंदर ने मोटी बिल्ली को मारा?

Kyaa Mohindar ne moti billii ko mara?
wh Mohindar erg fat cat acc hit.past
'Did Mohindar hit a fat cat?'

नहीं! मोहिंदर ने एक पतली बिल्ली को मारा।

Nahi! Mohindar ne ek $[patli]_F$ billii ko mara.
no Mohindar erg det slim cat acc hit.past
'No! Mohindar hit a slim cat.'

Table 1: *Target words with different numbers of syllables and stress position*

initial stress	penutlima/ antepenultima	penutlima
nai – 'new' lal – 'red' gol – 'round'		
asli – 'natural' kali – 'empty' lambe – 'big' nakli – 'artificial' sundar – 'nice' kale – 'black' namkin – 'salty' bure – 'bad' gande – 'dirty' halke – 'light' patli – 'thin'		
bartiye – 'indian' kurderi – 'rough' hoshiyaar –'intelligent' bahari – 'foreign' sabdahik – 'weekly'	jaberdust– 'pretty' mulayam – 'soft' imandaar – 'honest' abmanit– 'offended' samanit–'honorable' gulabi – 'pink' narangi – 'orange' andheri – 'dark' sugheri – 'golden' samajik – 'social'	
parvetiye – 'hilly'	parishrami – 'busy' sherareti – 'naughty' akashije – 'high'	etihasik – 'historic' parivarik –'internal' djegeralu–'polemic'

2.2 Procedure

Eleven native speakers of Hindi (4 female, 7 male) participated in the experiment. All participants were Ph.D.-students in Berlin or Potsdam. The age average was 29 years. Each speaker was paid a small fee for participation. One male speaker had to be dropped from the F_0 analysis due to creaky voice.

The experiment was carried out using presentation software. The whole presentation was carried out in Devenagari script. Participants were digitally recorded on a laptop (Levono R61) using Audacity (Version 1.2.6) in a quiet room at their houses or working places using a headset (Logitech Internet Chat Headset). The headphones were binaural with a frequency spectrum from 20-20000 Hz and an acoustic impedance of 32 Ohm with an integrated volume control, so that every participant could adjust the volume. The microphone was an electret condenser type with a sensitivity of -39 dbV/Pascal.

The participants were familiarized with the task through written and verbal instructions, followed by four practice trials. Trials for the contrastive focus consisted of a visual presentation of the question and its answer on the computer screen. For elicitation of the wide focus condition only the target sentence was presented. Participants heard the pre-recorded question over headphones, spoken by a middle aged male. At the same time the target sentence was presented on the screen. The instructions forced the participants to listen to the question first and thereafter to read the answer quietly. Then the question was presented again, and afterwards the participants had to speak out the answer displayed on the screen as a response to the question. If the answer started with *nahi* the participant was asked to put a phrase break after it. The presentation was pseudo randomized and fillers were interspersed.

2.3 Acoustic analysis

The recordings were digitized at a sampling frequency of 44.1 kHz and 32 bit resolution. The target word in all 814 sentences (37 target words × 2 information structural conditions × 11 speakers) was labeled by hand at the level of the word, the syllable, and the segments. Labeling in Praat [29] was based on a see-listen-label method, visually evaluating the spectrogram of the sound files and listening. Standard cues for segmental labeling were employed [30], and boundaries were set automatically at zero crossings via a Praat script. For each target word, the duration of the word, the syllables and the segments was extracted in ms using a Praat script.

In the target word the rise was labelled. The H tone was set at the local pitch maximum, the L tone at the local low if present, and if not, at the beginning of the actual pitch rise. To normalize for pitch range and sex, two reference points were set, one for high pitch (R1) before the target word and one for low pitch (R2) after it. Obvious errors in the pitch track were manually corrected, and F_0 was smoothed at 10Hz. The pitch values were extracted via a Praat script. The actual Hz (x) values were normalized applying the formula in (3) following [31].

$$(3)\quad y = \frac{x - R2}{R1 - R2}$$

2.4 Statistic analysis

The statistical analysis relied on the dependent variables 'word duration' and 'F_0' calculated for L, H and span. A paired samples T–test was carried out in *R*. A *p*-value smaller than 0.05 indicates statistical significance.

3 Results

3.1 Information structural effects on syllable duration

The results from the statistical analysis reveal that contrastive focus significantly affects the duration of the target words in Hindi, $t(10)=2.2281$, $p = 0.01$. To gain a better understanding of focal lengthening in the words Table 2 presents the mean syllable duration for both information structural conditions, with wide focus (wFoc) as baseline condition compared to contrastive focus (cFoc). From Table 2 it can be seen that in contrastive focus each syllable of a target word is lengthened compared to wide focus renditions.

Table 2: *Mean syllable duration in ms for wide focus (wFoc) vs. contrastive focus (cFoc); grey shading indicates the stressed syllable, n see right column.*

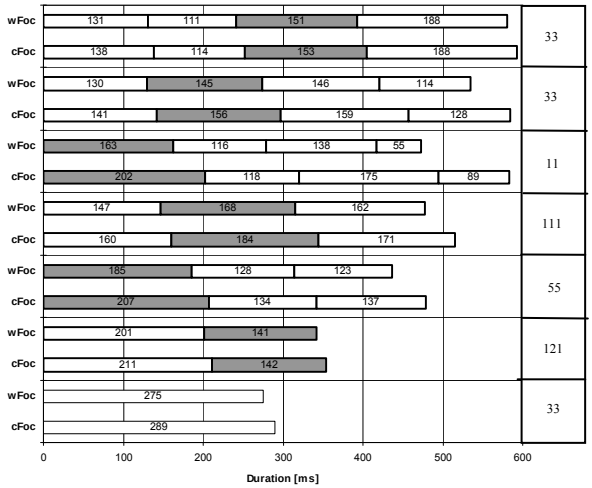

Table 3: *Number of syllables (nos), stress pattern (0=unstressed, 1=stressed), amount of durational change in ms (ΔC), and as percentage (%L) for word and syllable level, the syllable-to-word-lengthening-ratio (%Δσ /ΔW) for contrastive focus compared to wide focus.*

nos	stress	ΔCW	%L	ΔCσ	L%	%Δσ/ ΔW
1	1	15	6	15	6	100
2	1	11	3	10	5	90
2	0	11	3	1	1	10
3	1	43	10	22	12	51
3	0	43	10	6	5	14
3	0	43	10	15	12	35
3	0	37	8	12	8	32
3	1	37	8	16	10	43
3	0	37	8	9	6	24
4	1	110	23	39	24	35
4	0	110	23	2	2	2
4	0	110	23	36	26	33
4	0	110	23	22	40	20
4	0	51	10	12	9	24
4	1	51	10	12	8	24
4	0	51	10	13	9	25
4	0	51	10	14	12	27
4	0	13	2	8	6	62
4	0	13	2	3	3	23
4	1	13	2	2	1	15
4	0	13	2	0	0	0

For target words with four syllables all syllables except one appear to contribute equally to word lengthening. For the initially stressed four syllable target word, which contains only one item, the whole word gets lengthened by 23%. Equally to the two and three syllable target word the stressed syllable contributes the biggest amount 35% of the durational difference. The smallest contribution is made by the second unstressed syllable 2%, the third syllable is lengthened by 33% and the final one 20%. For the two remaining 4 syllable target bunches the picture is different. For the antepenultima stressed items all syllables contribute nearly equally (24-27%) to the lengthening of the words which amounts 51%. In the penultima stressed four syllable target words which are lengthened by 13% the stressed syllable contributes only 15% to the total amount of lengthening. The first syllable contributes 62%, the second one 23% and the final syllable is not involved in the durational difference.

3.2 Information structural effects on F0

Figure 1 shows the mean normalized F_0 values for the L, the H tone and pitch span, which is calculated as the difference between the two tones [13]. The L tone gets significantly lower under contrastive focus ($t(9) = 2.301$, $p < .05$), whereas the H tone gets significantly higher ($t(9) = 4.130$, $p < .01$). As a consequence also the pitch span is significantly enhanced under contrastive focus ($t(9) = 4.613$, $p = .001$).

To get an insight into speaker specific strategies Table 4 shows the mean pitch change for L and H tones between contrastive and wide focus for all speakers. Concerning the L tone Table 4 reveals that only half of the speakers lower of the L tone. This is true for both females and males. All speakers, however, increase the height of the H tone.

Figure 1: *Mean F0 in normHz for L, H, and pitch span (H-L) for contrastive compared to wide focus; n=10.*

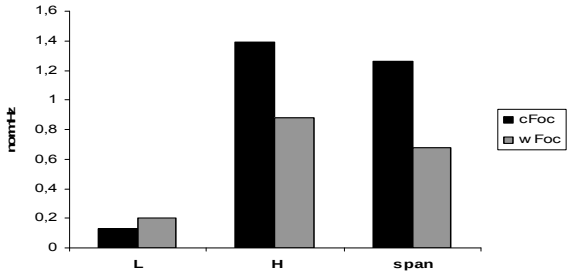

Table 4.: *Mean amount of pitch change in normHz (ΔC) for L and H for contrastive focus compared to wide focus, for all subjects, sex.*

subject	sex	ΔCL	ΔCH
1	♂	-0,32	0,27
2	♂	-0,30	1,16
3	♂	0,06	0,88
4	♀	-0,45	0,08
6	♂	-0,04	0,32
7	♀	0,00	0,42
8	♀	-0,14	0,56
9	♀	0,05	0,17
10	♂	-0,03	0,56
11	♂	-0,32	0,27

Any change in tonal realization results in a larger pitch span, and speakers differ as to which tone mostly contributes in the change of pitch span. Most speakers lower L tones and raise H tones at the same time. Speaker 4 only lowers the low tone with almost no increase of the H tone, and speakers 3, 7 and 9 only raise the H tone to achieve a larger pitch span.

4 Discussion

The results of our experiment suggest that contrastive focus affects the pitch rise associated with a prosodic word in Hindi declaratives in two ways. First, duration is affected. The word is lengthened under contrastive focus. The involvement of each syllable was studied in detail because stress in Hindi is not an uncontroversial issue. Most researchers agree that there

is stress ([2], [3], [4], [5], [6], [7]), the phonetic cues and listeners agreement on the placement of stress are yet still unclear. Ohala [3] suggests that stress is far weaker than in English, but for many words there is only one syllable that can be made prominent. Regarding the involvement of the stressed syllable in the focal lengthening the results show that for di- and trisyllabic target words the stressed syllable contributes the biggest amount to the word lengthening effect. The same effect was reported for the four syllable target word with initial stress, though not for the other four syllable words. Putting these words aside there seems to be a hint that quantitatively strong syllables are more involved in the lengthening than the unstressed syllables like in German [25] and Swedish [27]; a result which may contribute to the understanding of Hindi stress. In non-stress-accent languages like Japanese only F_0 should be used as a phonetic correlate of accentual prominence [32]. It would be worth studying the lengthening in more detail looking for segmental and syllable structure effects as well as for domain effects.

Second, the scaling of the pitch contour is affected. Contrary to [9] a higher scaling for the H tone in contrastive focus was observed in comparison to the wide focus baseline. And the L tone was also affected showing a significant lowering due to contrastive focus which together with a change of the H tone results in an increased pitch span. Concerning speaker variation we showed that all speakers modulated pitch span but to a different extend to express contrast prosodically. The tonal configuration is not changed under contrastive focus. Speaker specific strategies expressing focus are also reported for German [33]. A perception experiment testing the relevance of the phonetic cues found in this study would shed light on the perceptual utilization and maybe also on functional ranking of the cues.

Apart from the well established H raising effect in contrastive focus reported for intonation languages [17], [22], [23], [24] this study shows a reverse effect for the L tones which results in a difference in pitch span. The tonal distinctions are made sharper, an effect also reported for Mandarin Chinese [34]. Thus, it is the strategy of pitch span change (lowering L tones, raising H tones, or both) that makes Hindi an interesting case in terms of prosodic typology regarding the expression of focus.

5 Acknowledgements

Thanks to Anjana Singh for preparation of the target words, to Shravan Vasishth for taping of the questions, Umesh Patil for preparation of the material in Devenagri script and all participants attending. This research has been funded by a DFG research grant (SFB 632:D5) at Potsdam University.

6 References

[1] Shapiro, M.C. (1986). A Primer of Modern Standard Hindi. Delhi: Motilal Banarsidass.

[2] Moore, P.R. (1965). A study of Hindi intonation. University of Michigan. Unpublished dissertation.

[3] Ohala, M. (1986). A search for the phonetic correlates of Hindi stress. South Asian languages: structure, convergence, and diglossia, ed. by Bh. Krishnamurti, C. Masica, and A. Sinha, 81-92. Delhi: Motilal Banarsidass.

[4] Harnsberger J. (1994). Towards an intonational phonology of Hindi. Unpublished manuscript. (http://www personal.umich.edu/~jharns/hindi.html)

[5] Hussain, S. (1997). Phonetic Correlates of Lexical Stress in Urdu. Northwestern University. Unpublished dissertation.

[6] Nair R. (2001). Acoustic Correlates of Lexical Stress in Hindi. In: Abbi A., Gupta R.S., Kidwai A. (Ed.) (2001) Linguistic Structure and Language Dynamics in South Asia- papers from the proceedings of SALA XVIII roundtable.

[7] Dyrud L. O. (2001). Hindi-Urdu: Stress accent or non-stress accent? M.A. Thesis. University of North Dakota. (http://www.und.nodak.edu/dept/linguistics/theses/2001Dyrud)

[8] Hayes B. (1995). Metrical stress theory: principles and case studies. Chicago. University of Chicago Press.

[9] Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., Vasishth, S. (2008). Focus, word order, and intonation in Hindi. Journal of South Asian Linguistics 1, 53-70.

[10] Chafe, W. L., (1974). Language and consciousness. Language 50 (1), 111–133.

[11] Ladd, D. R., (1980). The structure of intonational meaning: Evidence from English. Bloomington: Indiana University Press.

[12] Gussenhoven, C., (1984). On the Grammar and Semantics of Sentence Accents. Dordrecht: Foris.

[13] Ladd, D. R., 1996. Intonational Phonology. Cambridge: CUP.

[14] Rooth, Mats E. (1985). Association with Focus. PhD dissertation. University of Massachusetts.

[15] Rooth, M. (1992). A Theory of Focus Interpretation. NLS, 1, 75-116.

[16] Allerton, D. J. (1978). The notion of 'givenness' and its relations to presupposition and to theme. Lingua 44 (2-3), 133–168.

[17] Baumann, S., (2006). The Intonation of Givenness. Evidence from German. Tübingen: Niemeyer.

[18] Kidwai, A. (2000). XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu. Oxford: OUP.

[19] Cruttenden, A. (1997). Intonation. Cambridge: CUP.

[20] Truckenbrodt, H. (1995). Phonological phrases: their relation to syntax, focus, and prominence. PhD thesis, MIT.

[21] Kratzer, A., Selkirk, E. (2007). Phase theory and prosodic spellout: The case of verbs. The Linguistic Review 24:93-135.

[22] Féry, C., Kügler, F., 2008. Pitch accent scaling on given, new and focused constituents in German. Journal of Phonetics 36, 680–703.

[23] Eady, S. J., Cooper, W. E., Kloouda, G. V., Mueller, P. R., Lotts, D. W., (1986). Acoustical Characteristics of Sentential Focus: Narrow vs. Broad and Single vs. Dual Focus Environments. Language and Speech 29, 233–251.

[24] Cooper, W. E., Eady, S. J., Mueller, P. R., (1985). Acoustical Aspects of Contrastive Stress in Question-Answer Contexts. Journal of the Acoustical Society of America 77, 2142–2156.

[25] Kügler, F., Genzel, S. (submitted) Sentence length, position, and information structure effects on segmental duration in German. Submitted to Journal of Phonetics.

[26] Turk, A. E., White, L., (1999). Structural influences on accentual lengthening in English. Journal of Phonetics 27 (2), 171–206.

[27] Heldner, M., Strangert, E., (2001). Temporal effects of focus in Swedish. Journal of Phonetics 29 (3), 329–361.

[28] Harnsberger J. (1999). The role of metrical structure in Hindi intonation. South Asian Language Analysis Roundtable 1999. (http://www-personal.umich.edu/~jharns/hindi.html)

[29] Boersma, P., Weenink, D., (2008). Praat: doing phonetics by computer (Version 5.0.35) [Computer program], retrieved from http://www.praat.org/.

[30] Turk, A., Nakai, S., Sugahara, M., (2006). Acoustic segment durations in prosodic research: A practical guide. In: Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N., Schließer, J. (Eds.), Methods in empirical prosody research. Berlin: Mouton de Gruyter, 1–27.

[31] Truckenbrodt H. (2004). Final lowering in non-final position. Journal of Phonetics 32, 313–348.

[32] Beckman, Mary E. (1986). Stress and Non-stress Accent. Berlin: Mouton.

[33] Baumann, S., Becker, J., Grice, M., Mücke, D. (2007). Tonal and Articulatory Marking of Focus in German. Proceedings of the 16th International Congress of Phonetic Science. Saarbrücken, Germany. 1029-1032.

[34] Xu, Y. (1999). "Effects of tone and focus on the formation and alignment of f0 contours". Journal of Phonetics 27, 55-105.