

# The Phonology of Contrastive Focus in Standard Colloquial Assamese

Asim. I. Twaha<sup>1</sup>, Shakuntala Mahanta<sup>1</sup>

<sup>1</sup>*Phonetics and Phonology Lab, Indian Institute of Technology Guwahati*

## 1 Introduction

The present study investigates how Standard Colloquial Assamese (henceforth SCA) underlines contrastive focus (henceforth CF) phonologically, and what are the phonetic cues it employs in doing so. Assamese belongs to the Eastern Indo-Aryan language area of the Indo European language family (Goswami, 1982; Goswami & Tamuli, 2003) with SOV as the canonical word order. SCA variety is mostly spoken in the eastern districts of Assam: Tinsukia, Dibrugarh, Lakhimpur, Dhemaji, Sibsagar, Jorhat, Golaghat and Sonitpur (Moral, 1992).

The present paper has been arranged into five sections: first section (§2) elaborates CF and the perspective in which the concept has been used in this paper, the second section (§3) deals with the post-lexical prosody of SCA. Subsequently in the next section (§4) phonological manifestation of CF in SCA has been explained. The following section (§5) concentrates on how cross-linguistically attested phonetic correlates of CF such as pitch and duration values interact with CF in SCA. Finally the conclusion (§6) consolidates the entire discussion with respect to the findings of the present study.

## 2 Contrastive Focus (CF)

CF has been considered the strongest type of focus ‘as the speaker asserts something which may contradict the expectations of the hearer’ (Féry, 2013). It has been described variously as identificational focus (Kiss, 1998), alternatives focus (Rooth, 1992) and contrastive focus (Selkirk, 2002; Zubizarreta, 1998; Kratzer, 2004). Although CF has been differentiated from the instances of focus created out of correction (Tomioka, 2009; Zimmermann & Onea, 2011), in this chapter both types of foci will be treated interchangeably as both of them generate a set of alternatives out of which the focused alternant receives contrastive focus (Rooth, 1992; Vallduví & Vilkuna, 1998; Kiss, 1998).

Rooth (1992) defines CF in terms of a set of alternatives; for her when a constituent receives CF it generates a set of alternatives which constitutes its focused meaning. This alternatives set includes the ordinary meaning of the focused constituent within its focused meaning. In (1) when *Romen* is focused it creates an alternatives set of ordinary meanings: [Romen killed the cat, Ram killed the cat, Shyam killed the cat, etc.] of which the focused meaning is also a part. It is this alternatives set which differentiates a focused constituent from non-focused ones: constituents which are not focused do not generate an alternative set of meanings (Rooth, 1992; 1997).

- (1) A. Ram killed the cat.  
B. No, Ramen killed the cat.

According to Zubizarreta (1998), CF makes its realisation in relation to the context; it is the preceding statement which provides the context for CF. Zubizarreta talked about two-fold function of contrastive focus: apart from negating “the value assigned to a variable” in the preceding statement, contrastive focus provides an alternate value for the variable. In the following example, the context for contrastive focus is created by (2A), and (2B) which bears contrastive focus on *Red* performs two simultaneous functions: first it negates *John is wearing a blue shirt today* i.e. John is not wearing a blue shirt today, and second it induces an alternate value for the constituent which has been negated, here it is *Red* which is introduced in contrast to *blue* in the previous context statement: *Ram is wearing a red shirt today*.

- (2) A. John is wearing a blue shirt today.  
B. John is wearing a Red shirt today (not a blue shirt).

From the cross-linguistic literature it is apparent that while some languages employ phonological and phonetic cues in order to mark focus, some others rely solely upon phonetic correlates. Languages like English (Silverman & Pierrehumbert, 1990), German (Féry, 1993; Féry & Kügler, 2008), Dutch (Gussenhoven, 1983), Bengali (Hayes & Lahiri, 1991) and Korean (Jun & Lee, 1998) phonologically distinguish contrastive focus from broad or wide focus (henceforth WF). In these languages CF is marked by either placing sentential or nuclear accent<sup>1</sup> on the focused constituent or by demarcating a prosodic boundary after focus or both. For example, in English the tonal pattern of a sentence changes with focus change: the nuclear or sentence stress falls on the most prominent word or constituent within the IP and the post-focus constituents undergo de-accenting. There are other languages like French (Féry, et al., 2010) and Bengali (Hayes & Lahiri, 1991) where focus exercises a demarcating function at the phrase boundary of the focused constituent.

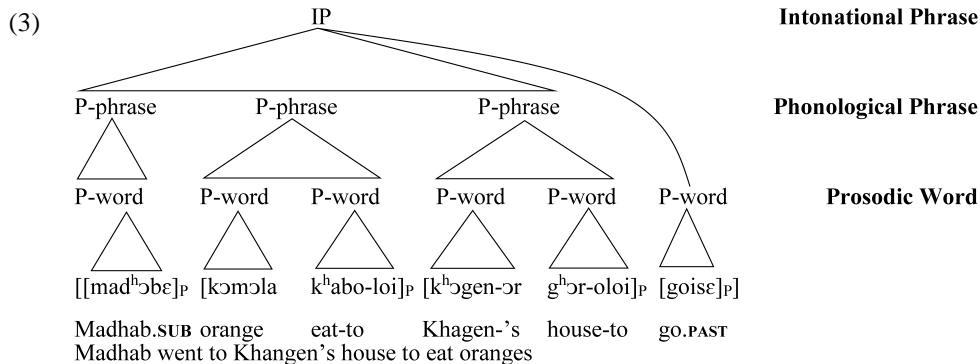
The above discussed languages employ pitch accents and phrasing in a phonologically significant way in order to highlight the focused status of a constituent. However, there are languages like Romanian (Manolescu, et al., 2009), Catalan (Borras-Comes, et al., 2014), Spanish (Prieto, 2004) where greater pitch movement, longer duration and larger pitch range mark CF. These languages advocate for the inevitability of Pitch range to be included in the phonological representation. In these languages, as has been revealed by results of the perception experiments undertaken in the referred studies, pitch range is employed in a categorically distinct manner in contradiction to the traditional view (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986) which holds that pitch range is gradient and beyond the scope of phonology.

In case of SCA, CF is marked by phonological phrasing, post-focus compression and increased pitch range. A constituent which receives CF forms a phonological phrase (henceforth P-phrase) overriding the syntactic rules of phrasing, and all the constituents following the focused constituent undergo a F<sub>0</sub> compression. In this variety it has been observed that a focused constituent is marked with higher pitch range value.

### 3 Post-lexical Prosody of SCA

Assamese, like most other South Asian languages (Bengali (Hayes & Lahiri, 1991), Hindi (Patil, et al., 2008), Tamil (Keane, 2014), etc.), is a head final language where verb concludes a declarative utterance. In our inquiry it has been found that SCA declarative utterances maintain rising contour on each of the pre-verbal constituents where each succeeding rise shows lower (downstepped) pitch escalation compared to the previous peak. The final constituent, which is mostly a verb, lacks prominence.

**3.1 Prosodic hierarchy** The intonational framework adopted in this study is based on the pioneering works by Liberman (1975), Bruce (1977), Pierrehumbert (1980), Beckman and Pierrehumbert (1986), Pierrehumbert and Beckman (1988), and Hayes and Lahiri (1991). Intonation is a post-lexical phenomena (Ladd, 1996) which adheres to a hierarchical structure of domains, which are phonological in nature (Ito & Mester, 2012). It has been assumed in this work that SCA also shows a hierarchically arranged prosodic structure. Here we have adopted the prosodic hierarchy theory propagated and developed by Selkirk (1978; 1986; 2009), Nespor and Vogel (1986) and others.



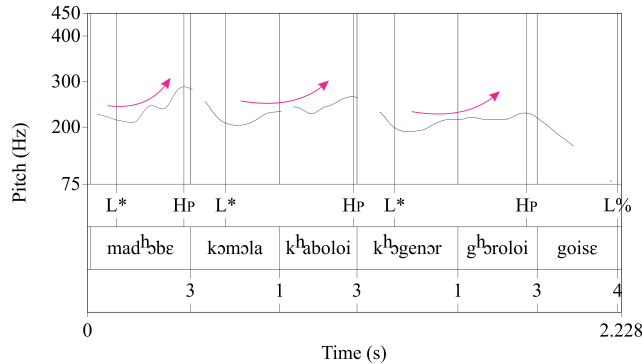
<sup>1</sup> A sentential or nuclear accent refers to 'the last major intonational movement in an utterance' (Xu, 2011)

The highest node in the hierarchy is Intonational Phrase (henceforth IP), which corresponds mostly to a sentence. The next node below IP is that of P-Phrase. A P-Phrase is motivated by junctural phenomena rather than by the intonation pattern; it is marked by an obligatory pitch accent and often a boundary tone. The lowest node in the hierarchy adopted by us is that of prosodic word (henceforth P-word). In SCA, a prosodic word phonologically maps a syntactic word, it creates a prosodic domain parallel to the syntactic word. The idea behind considering P-words as the terminal nodes in the hierarchy is motivated by the assumption that intonation involves phrase-level phonological processes<sup>2</sup>, and a P-word is not associated with tone assignment (pitch accent and boundary tone). It is the P-Phrase which is the minimal unit of tone assignment in Assamese.

**3.2 Intonational structure** Even though phrases are supposed to be governed by the syntax of a language, increasing evidence shows that prosodic phrasing is governed more by information structuring imperatives rather than syntactic constituency. If we consider the schema displayed above (3) and the syntactic structure demonstrated below (4) we see that each pre-verbal P-phrase corresponds to a syntactic phrase.

- (4) [[mad<sup>h</sup>ɔbɛ]<sub>NP</sub> [[kɔmɔla k<sup>h</sup>abo-loi]<sub>PP</sub> [k<sup>h</sup>ɔgen-ɔr g<sup>h</sup>ɔr-oloi]<sub>PP</sub> [goisɛ]<sub>VP</sub>]<sub>VP</sub>  
 Madhab.SUB orange eat-to Khagen-'s house-to go.PAST  
 Madhab went to Khagen's house to eat oranges

In (4) *mad<sup>h</sup>ɔbɛ*, *kɔmɔla k<sup>h</sup>abo-loi* and *k<sup>h</sup>ɔgen-ɔr g<sup>h</sup>ɔr-oloi* are three syntactic phrases (one noun phrase and two postpositional phrases respectively) which form P-phrases at the prosodic level. When we look at the intonational contour, as Figure-1 shows, each syntactic phrase corresponds to a pitch rise highlighting its manifestation as a single intonational unit, and the entire sentence is delimited by a low IP final boundary tone (in this paper L<sub>I</sub> and L% have been used interchangeably).



**Figure 1** Here the three preverbal syntactic phrases: *mad<sup>h</sup>ɔbɛ*, *kɔmɔla k<sup>h</sup>aboloi* and *k<sup>h</sup>ɔgenɔr g<sup>h</sup>ɔroloi* are intonationally marked by a rise on each. The left most syllable in each phrase bears low pitch accent (L\*) and the rightmost syllable demarcates the boundary with a high boundary tone (HP).

A P-word in SCA is not tonally specified unless it forms a phrase at the prosodic level. For instance, in (4) the phrase *kɔmɔla k<sup>h</sup>aboloi* contains two P-words: *kɔmɔla* and *k<sup>h</sup>aboloi* which lack tonal specification individually. The first syllable of *kɔmɔla* receives low pitch accent (L\*) of the phrase *kɔmɔla k<sup>h</sup>aboloi* and the final syllable of *k<sup>h</sup>aboloi* marks the termination of the phrase with high boundary tone (HP).

In SCA the stress placement rule is predictable: P-words, being sensitive to weight-to-stress principle, place prominence on a heavy syllable, though the default position for stress assignment is the first syllable. Heavy syllables do not attract primary stress beyond the second syllable (Goswami, 1982; Mahanta, 2001). It is syllable duration, rather than pitch value, that serves as the cue to this prominence pattern which is greater in case of heavy (bimoraic) syllables compared to light (monomoraic) syllables. However, irrespective of P-word level prominence pattern, the pitch accent of a P-phrase is always found to be assigned to the phrase

<sup>2</sup> These three categories (P-word, P-Phrase and IP) have been described as *interface* categories by Ito and Mester. **Invalid source specified.**





## 5 Phonetic cues to CF

With an aim to investigate the effects CF wields on the focused constituent and on the surrounding (pre- and post-focus) constituents, we designed an experiment which observes pitch and duration values of the said constituents in two focus conditions – WF and CF. The methodology followed in the experiment is elaborated below.

**5.1 Methodology** In the experiment we compared two sets of utterances: one in WF realisation and the other in CF context. The experimented sentence sets comprised of identical and equal numbers of sentences, and the sentence frames were in unmarked SOV word order. Renderings in WF condition were novel sentences uttered in response to the question *ki hol?* (What happened?), whereas in the second set, CF was placed on the word in object position (see below in (7) and (8)). In order to create the CF environment, speakers were asked to utter compound sentences with two independent IPs connected with the conjunction *kintu* (but). The word in the object position of the first IP is in contrastive relation with the corresponding word in the first clause. The speakers were explicitly instructed to maintain contrast only between the objects (leaving the complementiser) in the two IPs as the subjects in both the IPs are different, and there is a possibility that they may also induce a contrastive effect at the IP initial position. In addition to that the subjects (proper nouns) in both the IPs are repetitive, only two names (Ramen and Nagen) have been repeatedly used, which makes the speakers to take them for granted, and hence not focus them. The experiment is conducted for both di- and tri-syllabic words in two different focus conditions. The schema representing the data collection process is given below.

(7) *The First set of sentences (WF)*

ki hol ?  
what happen.PAST

- a. [[rɔməŋɛ]P [mamɔr buli]P kolɛ]I  
Ramen.SUB rust.OBJ that.COMP say.PAST  
Ramen said mamɔr (rust).
- b. [[rɔməŋɛ]P [mɔɔna buli]P kolɛ]I  
Ramen.SUB threshing.OBJ that.COMP say.PAST  
Ramen said mɔɔna (threshing).

(8) *The Second set of sentences (CF in the second IP)*

- a. [[rɔməŋɛ]P [matal buli]P kolɛ]I [[kintu]P [nɔgɛŋɛ]P [mamɔr]CF buli kolɛ]I  
Ramen.SUB drunk.OBJ that.COMP say.PAST but Nagen.SUB rust.OBJ that.COMP say.PAST  
Ramen said matal (drunk) but Nagen said mamɔr (rust).
- b. [[rɔməŋɛ]P [mɔmɔta buli]P kolɛ]I [[kintu]P [nɔgɛŋɛ]P [mɔɔna]CF buli kolɛ]I  
Ramen.SUB drunk.OBJ that.COMP say.PAST but Nagen.SUB threshing.OBJ that.COMP say.PAST  
Ramen said mɔmɔta (affection) but Nagen said mɔɔna (threshing).

The specimen sentences given in (7) constitute the WF data set, 7(a) disyllabic and 7(b) tri-syllabic words in object position. In the second set of exemplified sentences (8), the second IP after *kintu* (but) form the data set with CF on the object; here also we have words with two 8(a) and three 8(b) syllables. The reason behind using two different lengths of focused words is to explore any influence word length has on focus realisation.

**5.2 Subjects** For the data three (3) male and two (2) female speakers (20 to 30 years old) from Sivasagar District of Assam were recorded in the recording booth of Phonetics and Phonology Lab, Indian Institute of Technology Guwahati.

**5.3 Data analysis** All the pre-focused, focused and post-focused constituents from the compared clauses are measured for their pitch and duration values both at the syllable and word level using PRAAT (Boersma & Weenink, 2015). The values were then analysed running one-way ANOVA test in StataMP13 (StataCorp, 2013). A total of five speakers (two female and three male) are recorded using a Tascam DR-100mkII

recorder with a Shure SM10A-CN head-worn microphone in wav format at the frequency of 44.1 kHz with 16 bit resolution. In order to tackle inter-speaker variation, pitch and duration values are normalised using the z-score normalisation method (Disner, 1980; Rose, 1987; 1991) before running the statistical tests. In our study, for the statistical analysis each  $F_0$  value is calculated by using the formula demonstrated below.

$$(9) \quad F_{0\text{ norm}} = (F_{0i} - F_{0\text{ aver}}) / s$$

Where

$F_{0i}$	=	$F_0$ value of an individual point
$F_{0\text{ aver}}$	=	average of all the $F_0$ values in a P-phrase
$s$	=	standard deviation of all $F_0$ values in a P-phrase

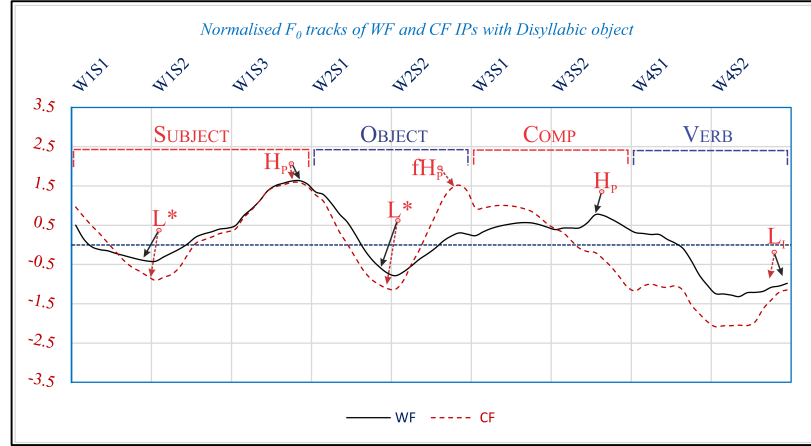
A sum total of [20(expressions) x 5(speakers) x 2(focus conditions) x 3(iterations)] six hundred (600) utterances comprise the current data set, which is equally distributed into di- and tri-syllabic words in focused position: three hundred (300) utterances with di-syllabic and three hundred (300) with tri-syllabic words in object position. In case of focused and pre-focused constituents,  $F_0$  max and min are measured at the right and left half of the constituents respectively. It was done so because if we refer to Figures-3 and 4 we see that both pre-focused and focused constituents show rising contours with low pitch accent manifested within the first half and high boundary tone realised in the second half of the constituent. In case of the post-focus constituent *buli* (that),  $F_0$  was measured at three points: onset, mid and offset so that the significance level of post-focus compression may be explored.

In order to increase the accuracy of the pitch normalised contours displayed in Figure-3 and 4, pitch values are measured for each syllable at ten (10) points with equal intervals.

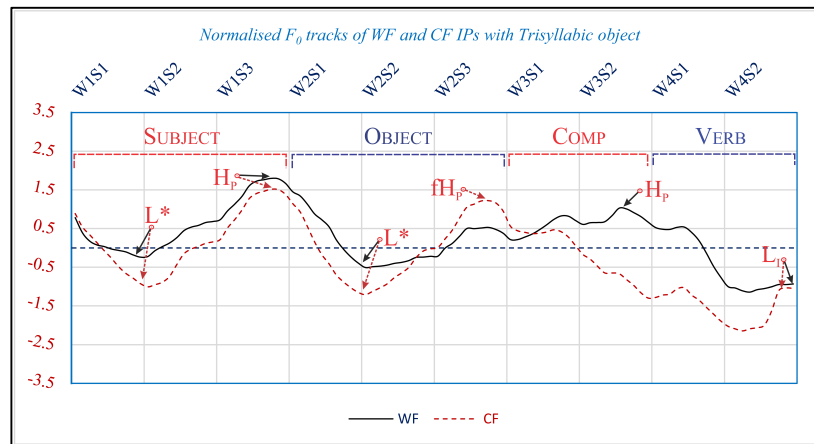
**5.4 Findings** The findings of the experiment will be discussed below with reference to pitch and durational changes introduced by CF. This section is divided into two sections (§5.4.1 and §5.4.2) where the first part deals with the impact of CF on the pitch values of pre-focus, focused and post-focus constituents in declarative utterances. The second portion illustrates how CF interacts with the durational measurements of the said constituents. In both the sections along with the focused constituent, preceding and succeeding environment of a contrastively focused constituent have been statistically compared. In half of the occurrences the focused constituent is disyllabic and in the other half it is trisyllabic. It was necessitated by the need to observe the consistency of the effect CF has on the surroundings. We did not measure the influence of the length of the word which receives focus on the surrounding environment. Apart from the focused constituent, the effect of CF on the pre- and post-focus constituents has been separately reported for both di- and trisyllabic focused words and upon casual conversation there did not seem to be any difference. Since this is outside the purview of this very short paper we will not elaborate on this aspect any more.

**5.4.1 Pitch ( $F_0$ )** Contours generated by z-score normalised pitch values provide us with a fair idea of the impact of CF on the overall pitch contour of an IP as against its WF variant. The diagrams below (Figures-3 and 4) demonstrate that in WF condition prosodic phrasing adheres to syntactic phrasing: the subject forms the first P-phrase marked by a pitch rise ( $L^*H_P$ ), the object together with the complementiser *buli* (that) constitutes the second P-phrase symbolised by another rise ( $L^*H_P$ ). The low pitch accent of the second rise also performs as the nuclear accent of the IP. The final constituent *kolē* (said) is left unaccented. When the object receives CF it forms a P-phrase demarcated by a focus induced pitch peak ( $fH_P$ ) on its final syllable after which the pitch contour maintains a smooth fall until the end of the IP. Here the focused constituent, through the formation of a P-phrase, bears the nuclear accent of the utterance. Though in both the focus conditions nuclear accent falls on the first syllable of the object, placement of the P-phrase boundary tone differs in the two contexts. In WF context  $H_P$  falls on the second syllable of *buli* (that-COMP), whereas in CF condition the focus high boundary tone ( $fH_P$ ) falls on the final syllable of the focused word. If we refer to Figure-3 and Figure-4, we clearly see that in WF condition, following the realisation of low pitch accent ( $L^*$ ) on the first syllable of the object the pitch track continues to rise and attains its highest value on the last syllable of *buli* establishing Object+*buli* as a single prosodic constituent and also demarcating the intonational boundary of the constituent ( $H_P$ ). However, when CF falls on the object it induces a focus high boundary tone ( $fH_P$ ) right after the object which is followed only by low IP boundary tone ( $L_I$ ). As such CF initiates a P-phrase out of the focused constituent.

Apart from the phonological aspect of focus marking in terms of P-phrase formation of the focused constituent, CF is also marked by pre- and post-focus pitch compression. In order to see whether this compression is statistically significant or not we conducted one-way ANOVA test on the values extracted from the different constituents, the result of which is reported and discussed immediately after the following Figures (3 and 4).



**Figure 3** Normalised pitch contours of identical sentences in WF and CF conditions. The black solid line represents WF utterances and red dashed line represents the same sentences uttered with CF on the disyllabic object.



**Figure 4** Normalised pitch contours of identical sentences in WF and CF conditions. The black solid line represents WF utterances and red dashed line represents the same sentences uttered with CF on the tri-syllabic object.

**5.4.1.1  $F_0$  values of pre-focus constituent** The measurement of  $F_0$  values of pre-focus constituent for both the syllable types show that this constituent undergoes pitch compression. While the  $F_0$  max does not confirm consistent significant difference between the two focus realisations, the  $F_0$  min value drops significantly. As it is reported in Table-1 and displayed in Figure-3, the  $F_0$  max of a CF pre-focus constituent does not show significant departure from the WF baseline ( $p > .05$ ,  $F [1, 298] = 1.28$ ,  $p = 0.26$ ) when the focused constituent is di-syllabic, though it shows significant decrease ( $p < .05$ ,  $F [1, 298] = 5.88$ ,  $p = 0.02$ ) when a tri-syllabic constituent is in focus. In case of  $F_0$  min, its value regularly decreases:  $p < .05$  when the focus is di-syllabic ( $F [1, 298] = 19.87$ ,  $p = 0.00$ ) and when the focus is tri-syllabic ( $F [1, 298] = 150.05$ ,  $p = 0.00$ ). Here the claim is not about whether the length of sentence medial focused constituent exercises any influence on the pre-focus constituent, but to observe and affirm the consistency of pre-focus pitch variation.

Category	Variables	<i>Pre-focus Constituent</i>				F	p-value
		WF		CF			
		Mean	SD	Mean	SD		
<i>Disyllabic focus</i>	F <sub>0</sub> max	-.06	1.01	.06	.97	(1, 298) = 1.28	0.26
	F <sub>0</sub> min	.24	1.09	-.24	.81	(1, 298) = 19.87	0.00
<i>Trisyllabic focus</i>	F <sub>0</sub> max	.14	.95	-.14	1.02	(1, 298) = 5.88	0.02
	F <sub>0</sub> min	.57	.85	-.57	.77	(1, 298) = 150.05	0.00

**Table 1** Comparison of the pitch values of pre-focus constituents in (di- and tri-syllabic) CF and WF conditions

**5.4.1.2 *F<sub>0</sub> values of focused constituent*** When a constituent is focused, as we can see in Figure-3 and 4, it forms a P-phrase and bears high focus boundary tone (fH<sub>P</sub>) on its last syllable. In WF contours the word in object position does not form a phrase independently at the prosodic level, and hence remains tonally unspecified at its boundary. In CF condition, on the other hand, focus induces a prosodic boundary right aligning with the focused constituent, as a result the last syllable of the word gets tonally associated with a high focus boundary tone (fH<sub>P</sub>). Since the object final syllable gets associated with a high focus boundary tone fH<sub>P</sub>, it is marked by a higher pitch value compared to when it is not focused. In Table-2 all the F<sub>0</sub> values – max, min and range relating to the focused constituent are presented. The F<sub>0</sub> max value of focused constituent displays statically significant increase, where  $p < 0.05$  for both disyllabic (F [1, 298]=287.67,  $p=0.00$ ) and tri-syllabic (F [1, 298]=46.39.67,  $p=0.00$ ) constituents. CF further forces low pitch accent to drop to new minimum for both the word lengths; F<sub>0</sub> min is always higher in WF conditions compared. ANOVA results show significant F<sub>0</sub> fall in the disyllabic ( $p < .05$ , F [1, 298]=40.7,  $p=0.00$ ) and tri-syllabic ( $p < .05$ , F [1, 298]=191.29,  $p=0.00$ ) focused words. Thus we see that in CF condition F<sub>0</sub> min (on the first syllable) of the focused constituent lowers significantly and F<sub>0</sub> max (at the right boundary) gets a boost. These changes in pitch min and max values are supplemented by an expanded F<sub>0</sub> range value: in both disyllabic words ( $p < .05$ , F [1, 298]=493.41,  $p=0.00$ ) and tri-syllabic words ( $p < .05$ , F [1, 298]=427.43,  $p=0.00$ ).

Category	Variables	<i>Focused Constituent</i>				F	p-value
		WF		CF			
		Mean	SD	Mean	SD		
<i>Disyllabic focus</i>	F <sub>0</sub> max	-.69	.68	.69	.74	(1, 298) = 287.67	0.00
	F <sub>0</sub> min	.34	.92	-.34	.94	(1, 298) = 40.7	0.00
	F <sub>0</sub> range	-.78	.59	.78	.63	(1, 298) = 493.41	0.00
<i>Trisyllabic focus</i>	F <sub>0</sub> max	-.36	.80	.36	1.04	(1, 298) = 46.39	0.00
	F <sub>0</sub> min	.62	.78	-.62	.77	(1, 298) = 191.29	0.00
	F <sub>0</sub> range	-.76	.60	.76	.67	(1, 298) = 427.43	0.00

**Table 2** Comparison of the pitch values of focused constituent (di- and tri-syllabic) in CF and WF conditions

**5.4.1.3 *F<sub>0</sub> values of post-focus constituent*** The normalised pitch contours displayed in Figure-3 and 4 visibly testify the post focus pitch compression. In order to verify whether this compression is statistically significant F<sub>0</sub> values are measured and compared at three points in the post focus constituent *buli*, which otherwise forms a single P-phrase together with the word preceding it. At initial, medial and final positions pitch value is recorded in order to measure the difference between the two normalised pitch contours of *buli*. As already mentioned, in WF context, object+*buli* form a single P-phrase with a high boundary tone (H<sub>P</sub>) right aligned with *buli* inducing a pitch peak on the final syllable of the word. With the formation of a P-phrase by the focused constituent (object) in CF condition, the high boundary tone gets deleted as an effect of post focus dephrasing. As such, our hypotheses are: 1) in CF condition, word initially, *buli* should display a significantly higher pitch value; 2) word medially, the pitch contours may or may not maintain an important difference; 3) finally, as the high tone (H<sub>P</sub>) disappears we should come across a radical pitch compression at the offset of *buli*. The statistical results demonstrated in Table-3 establishes our hypotheses to be true. F<sub>0</sub> value at the onset of *buli* is significantly higher in CF condition as it is measured at a position which immediately follows high boundary tone of the focused constituent:  $p < 0.05$  when *buli* follows both a disyllabic (F [1, 298]=94.78,  $p=0.00$ ) and a tri-syllabic F [1, 298]=6.16,  $p=0.01$ ) focused constituent.

However, at the mid position, the two contours do not always maintain a significant difference of values:  $p > 0.05$  when followed by a disyllabic focused constituent ( $F [1, 298] = 0.13$ ,  $p = 0.72$ ) and  $p < 0.05$  when followed by a tri-syllabic focused constituent ( $F [1, 298] = 130.19$ ,  $p = 0.00$ ). This is in conformity with our second hypothesis. Finally at the offset of *buli* we see a drastic pitch compression when preceded by either disyllabic ( $p < 0.05$ ,  $F [1, 298] = 594.15$ ,  $p = 0.00$ ) or tri-syllabic ( $p < 0.05$ ,  $F [1, 298] = 1384.22$ ,  $p = 0.00$ ) focused word.

**Post-focus Constituent**

Category	Variables	WF		CF		F	p-value
		Mean	SD	Mean	SD		
Disyllabic focus	F <sub>0</sub> onset@ <i>buli</i>	-.49	.82	.49	.91	(1, 298) = 94.78	0.00
	F <sub>0</sub> mid@ <i>buli</i>	.02	.88	-.02	1.10	(1, 298) = 0.13	0.72
	F <sub>0</sub> offset@ <i>buli</i>	.81	.61	-.81	.53	(1, 298) = 594.15	0.00
Tri-syllabic focus	F <sub>0</sub> onset@ <i>buli</i>	-.14	.94	.14	1.03	(1, 298) = 6.16	0.01
	F <sub>0</sub> mid@ <i>buli</i>	.55	.72	-.55	.93	(1, 298) = 130.19	0.00
	F <sub>0</sub> offset@ <i>buli</i>	.90	.48	-.90	.34	(1, 298) = 1384.22	0.00

**Table 3** Comparison of pitch realisation on post-focus *buli* (that-Complementiser) at three positions – onset (starting), middle and offset (final) in two focus realisation

**5.4.2 Duration** Similar to pitch values, duration values are also measured for the pre-focus, focused and post-focus constituents with an aim to examine the durational effect of CF on these constituents. The results of the one-way ANOVA test conducted reveal that in SCA duration does not play a conclusive role in CF marking. The results of one-way ANOVA test have been discussed below.

**5.4.2.1 Duration of pre-focused constituent** Duration of the pre-focus constituent shows an important increase in its value when it precedes a constituent with CF. Table-4 demonstrates that length of the focused constituent does not seem to exercise any effect on this durational expansion: when the focused constituent is disyllabic  $p < 0.05$  ( $F [1, 298] = 65.27$ ,  $p = 0.00$ ) and also when it is trisyllabic  $p < 0.05$  ( $F [1, 298] = 1384.22$ ,  $p = 0.00$ ).

**Pre-focus Constituent**

Category	Variables	WF		CF		F	p-value
		Mean	SD	Mean	SD		
Disyllabic focus	Duration	-.42	.93	.42	.87	(1, 298) = 65.27	0.00
Trisyllabic focus	Duration	-.65	.65	.65	.83	(1, 298) = 227.72	0.00

**Table 4** Comparison of the duration values of pre-focus constituents in (di- and tri-syllabic) CF and WF conditions

**5.4.2.2 Duration of focused constituent** A contrastively focused constituent is characterised by an overall lengthening; this lengthening is also accompanied by an increase in the final syllable duration. The overall durational increase is significant for both disyllabic ( $p < 0.05$ ,  $F [1, 298] = 47.65$ ,  $p = 0.00$ ) and trisyllabic ( $p < 0.05$ ,  $F [1, 298] = 100.02$ ,  $p = 0.00$ ) constituents. Further our claim that CF induces a phrasing effect on the focused constituent is phonetically evident in the final syllable lengthening. In both the types of focused words this lengthening is statistically significant ( $p < 0.05$ ): disyllabic ( $F [1, 298] = 6.07$ ,  $p = 0.01$ ) and trisyllabic ( $F [1, 298] = 17.16$ ,  $p = 0.00$ ).

**Focused Constituent**

Category	Variables	WF		CF		F	p-value
		Mean	SD	Mean	SD		
Disyllabic focus	DUR@object	-.37	.93	.37	.92	(1, 298) = 47.65	0.00
	Duration@last $\sigma$	-.14	.94	.14	1.01	(1, 298) = 6.07	0.01
Trisyllabic focus	DUR@object	-.50	.83	.50	.89	(1, 298) = 100.02	0.00
	Duration@last $\sigma$	-.23	.90	.23	1.03	(1, 298) = 17.16	0.00

**Table 5** Comparison of the duration values of focused constituent (di- and tri-syllabic) in CF and WF conditions

**5.4.2.3 Duration of post-focus constituent** Post-focus constituents, though undergo substantial pitch compression, do not experience ample durational change. After disyllabic ( $F [1, 298]=2.57, p=0.11$ ) and trisyllabic ( $F [1, 298]=0.48, p=0.49$ ) focused word  $p>0.05$  suggesting a trivial durational difference between WF and CF realisations (Table-6).

Category	Variables	Post-focus Constituent				F	p-value
		WF		CF			
		Mean	SD	Mean	SD		
Disyllabic focus	DUR@post-focus	-.09	.98	.09	1.01	(1, 298) = 2.57	0.11
Tri-syllabic focus	DUR@post-focus	.04	.95	-.04	1.04	(1, 298) = 0.48	0.49

**Table 6** Comparison of the duration values of pre-focus constituent (di- and tri-syllabic) in CF and WF conditions

In this section we have seen that pitch and duration value mark CF in SCA: the focused constituent shows greater pitch range and increased length. Pre-focus constituent also shows a significantly augmented pitch range compared to when the entire IP is in focus, and post-focus constituent shows a drastic compression of the pitch value. However, duration values of both the pre- and post-focus constituent do not seem to play a significant role in focus marking.

## 6 Conclusion

Speakers of SCA highlight a sequence with CF at both phonological and phonetic levels. A contrastively focused constituent always receives the nuclear accent of the domain it occurs (IP). This constituent demonstrates a rising contour ( $L^*fH_P$ ) associated at two points with the focused constituent: the low pitch accent ( $L^*$ ) is aligned with first syllable and the focus high boundary tone ( $fH_P$ ) to the final syllable. Focused nucleus is characterised by greater pitch range and increased duration value. Apart from initiating a phrasing effect on the focused constituent, CF also exercises a dephrasing effect on the post-focus sequence. Post-focus dephrasing is supported by the results of the phonetic experiments conducted and reported in the current study; results reveal that post-focus pitch compression is drastic. As far as the pre-focus constituent is concerned, it is not always marked by a significant pitch peak compression as it is the case with the post-focus string; duration does not seem to play the role of a reliable cue to CF status. Contrastive focus is highlighted in three ways: it forms a P-phrase, increases the pitch value of the focused constituent and it significantly shrivels the pitch value of the sequence following it.

## References

- Patil, U. et al., 2008. Focus, Word Order and Intonation in Hindi. *JSAL*, 1(1), pp. 55-72.
- Beckman, M. E. & Pierrehumbert, J. B., 1986. Intonational structure in English and Japanese. In: *Phonology Yearbook* 3. s.l.:s.n., pp. 255-310.
- Boersma, P. & Weenink, D., 2015. *Praat: doing phonetics by computer [Computer Program]*. [Online].
- Borras-Comes, J., Vanrell, M. d. M. & Prieto, P., 2014. The role of pitch range in establishing intonational contrasts. *Journal of the International Phonetic Association*, 44(1).
- Bruce, G., 1977. *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.
- Disner, S., 1980. Evaluation of vowels normalization procedures. *Journal of the Acoustical Society of America*, Volume 67, pp. 253-261.
- Féry, C., 1993. *German intonational patterns*. Niemeyer ed. s.l.:Tübingen.
- Féry, C., 2013. Focus as prosodic alignment. *Natural Language and Linguistic Theory*, pp. 683-734.
- Féry, C., Hörnig, R. & Pahaut, S., 2010. Correlates of phrasing in French and German from an experiment with semi-spontaneous speech. In: C. Gabriel & C. Lleó, eds. *Intonational Phrasing in Romance and Germanic: Cross-linguistic and bilingual studies*. University of Hamburg: John Benjamins Publishing Company, pp. 11-41.
- Féry, C. & Kügler, F., 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, October, 36(4), pp. 680-703.
- Goswami, G., 1982. *Structure of Assamese*. First Edition ed. Guwahati: Department of Publication, Gauhati University.

- Goswami, G. C., 1982. *Structure of Assamese*. First Edition ed. Guwahati: Department of Education, Gauhati University.
- Goswami, G. C. & Tamuli, J., 2003. Asamiya. In: G. Cardona & D. Jain, eds. *The Indo-Aryan Languages*. London: Routledge, pp. 391-443.
- Gussenhoven, C., 1983. Testing the reality of focus domains. *Language and Speech*, Volume 26, pp. 61-80.
- Hayes, B. & Lahiri, A., 1991. Bengali Intonational Phonology. *Natural Language & Linguistic Theory*, 9(1), pp. 47-96.
- Ito, J. & Mester, A., 2012. Recursive prosodic phrasing in Japanese. In: T. Borowsky, S. Kawahara, T. Shinya & Mariko Sugahara, eds. *Prosody Matters: Essays in Honor of Elisabeth Selkirk*. London: Equinox, pp. 280-303.
- Jun, S.-A. & Lee, H.-J., 1998. *Phonetic and phonological markers of contrastive focus in Korean*. Sydney, Australia, s.n., pp. 1295-1298.
- Keane, E., 2014. The Intonational phonology of Tamil. In: S. Jun, ed. *Prosodic Typology II: the Phonology of Intonation and Phrasing*. New York: Oxford University Press, pp. 119-153.
- Kiss, K. E., 1998. Identificational focus versus information focus. *Language*, 74(2), pp. 245-273.
- Kratzer, A., 2004. Interpreting focus: Presupposed or expressive meanings? A comment on Geurts and van der Sandt. *Theoretical Linguistics* 30, pp. 123-136.
- Ladd, D. R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lieberman, M. Y., 1975. *The intonational system of English*. s.l.:Ph.D. thesis--Massachusetts Institute of Technology.
- Mahanta, S., 2001. *Some Aspects of Prominence in Assamese and Assamese English*, Hyderabad: s.n.
- Manolescu, A., Olson, D. & Llebaria, M. O., 2009. Cues to contrastive focus in Romanian. In: *Phonetics and Phonology: Interactions and interrelations*. Austin: University of Texas, pp. 71-90.
- Moral, D., 1992. *A phonology of Asamiya Dialects : Contemporary Standard and Mayong*. PhD Thesis ed. Pune: PhD Thesis, Deccan College.
- Nespor, M. & Vogel, I., 1986. *Prosodic phonology*. Dordrecht, Holland ; Riverton, N.J.: Foris.
- Pierrehumbert, J. B., 1980. *The phonology and phonetics of English intonation*. The phonology and phonetics of English intonation. PhD thesis, MIT. Distributed 1988, Indiana University Linguistics Club: PhD thesis, MIT. Distributed 1988, Distributed 1988. Indiana University Linguistics Club.
- Pierrehumbert, J. B. & Beckman, M. E., 1988. *Japanese tone structure*. Cambridge, Mass.: MIT Press.
- Prieto, P., 2004. The search for phonological targets in the tonal space:H1 scaling and alignment in five sentence-types in Peninsular Spanish. *Laboratory approaches to Spanish phonology*, pp. 29-59.
- Rooth, M., 1992. A Theory of Focus Interpretation. *Natural Language Semantics*, Volume 1, pp. 75-116.
- Rooth, M., 1997. Focus. In: S. Lappin, ed. *The Handbook of Contemporary Semantic Theory*. Blackwell Reference Online ed. s.l.:Blackwell Publishing.
- Rose, P. J., 1987. Considerations on the normalization of the fundamental frequency of linguistic tone. *Speech Communication*, Volume 6, pp. 343-351.
- Rose, P. J., 1991. How effective are long term mean and standard deviation as normalization parameters for tonal fundamental frequency. *Speech Communication*, Volume 10, pp. 229-247.
- Selkirk, E., 1978. On Prosodic structure and its relation to syntactic structure. In: T. Fretheim, ed. *Nordic Prosody II*. Trondheim: TAPIR: Indiana University Linguistics Club.
- Selkirk, E., 1986. On derived domains in sentence phonology. *Phonology Yearbook*, Volume 3, pp. 371-405.
- Selkirk, E., 2002. *Contrastive FOCUS vs. presentational focus: Prosodic evidence from English*. s.l., s.n., pp. 643-46.
- Selkirk, E., 2009. On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu (Journal of the Linguistics Society of Japan)*, 136(Special Issue on Linguistic Interfaces), pp. 1-39.
- Silverman, K. & Pierrehumbert, J., 1990. The timing of prenuclear high accents in English. In: J. Kingston & M. Beckman, eds. *Papers in laboratory phonology*. Cambridge: Cambridge University Press, pp. 72-106.
- StataCorp, 2013. *Stata Statistical Software: Release 13*. College Station, TX(Texas): StataCorp LP.
- Tomioka, S., 2009. Contrastive Topics Operate on Speech Acts. In: M. Zimmermann & C. Fery, eds. *Information Structure: Theoretical, Typological, and Experimental Perspectives*. New York: Oxford University Press, pp. 115-138.
- Vallduvi, E. & Vilkuina, M., 1998. On Rheme and Kontrast. *Syntax and Semantics*, Volume 29 The Limits of Syntax, pp. 79-108.
- Xu, Y., 2011. *Post-focus compression: cross-linguistic distribution and historical origin*. Hong Kong, s.n., pp. 152-155.
- Zimmermann, M. & Onea, E., 2011. Focus marking and focus interpretation. *Lingua*, September, 121(11), pp. 1651-1670.
- Zubizarreta, M. L., 1998. *Prosody, Focus and Word Order*. London: The MIT Press.