# Post-focal compression as a prosodic cue for focus perception in Hindi

FRANK KÜGLER, *Goethe University Frankfurt*

Received APRIL 2019; Revised MARCH 2020

ABSTRACT

Focus in Hindi is prosodically marked by means of post-focal compression (PFC), and the present study examines whether PFC is a prosodic cue that is functionally used by listeners to perceive focus. In a production study with 30 native Hindi speakers uttering six syntactically ambiguous sentences, PFC occurred after a focused indirect object, but not after a focused direct object. Data from the production study were taken as stimuli for a forced-choice sentence-completion experiment, in which 18 native Hindi speakers listened to sentence fragments of the ambiguous sentences and were asked to decide which of two possible objects contrasts (direct object or indirect object) would correctly complete the sentence. Results show that if PFC was absent listeners were unable to choose the intended sentence completion. If PFC was present, correct sentence completion judgements increased significantly. Thus PFC is a cue for focus perception in Hindi. Based on its functional load in Hindi, we argue that pitch register with is compressed post-focally represents a further intonational category to consider, at least for languages like Hindi.

## 1 Introduction

For many languages, it is claimed that prosody plays a key role in encoding focus (e.g. Bolinger 1989). However, the cues languages use to prosodically express focus vary tremendously. Focus is understood here as a cognitive category of information structure that fulfils a communicative function to update a common ground shared by a speaker and hearer (Krifka 2008). To encode the cognitive category of focus, languages use either syntactic, morphological, and/or phonological means. Prosody is thus one means of focus marking. In a typology of prosodic focus-marking, Kügler & Calhoun (2020) propose three basic strategies into which most languages studied to date can be placed: (i) stress-based cues, (ii) phrase-based cues, or (iii) register-based cues.

The first group is certainly the most well-studied and typologically the most widespread group of languages, and their prosodic expression of focus—well-known from Germanic languages (e.g., Fanselow 2016, Féry & Kügler 2008)—is achieved by enhancing the prosodic cues of a phonological category associated with a focused word. These cues usually are higher fundamental frequency (f0), greater f0 movement, lengthening, increased intensity and higher spectral tilt on the focused word, which mainly but not exclusively relates to the stressed syllable of the focused word, as well as a drop in f0 after it (cf. overviews in Ladd 2008, Fletcher 2010, Turk 2011). The enhancing of cues marks the focused word in an utterance as prosodically most prominent.

The second group of languages, which uses phrase-based cues predominantly, has no lexical prominence, i.e. lexical stress, in their word-prosodic system. Prosodic phrasing is indicated by phrase tones and/or temporal and segmental cues. Focused constituents are marked by an insertion and/or deletion of a prosodic phrase boundary to increase the prominence of the focused constituent.

The third group of languages uses register-based cues. Focus affects the pitch register on, and/or after the focused word either by raising or lowering it. The pitch register defines reference lines relative to which local tonal targets are scaled (Clements 1979, Ladd 2008). A well-studied example is the tone language Mandarin Chinese (Xu 1999). On the focused word, the high tone of a lexical high (H), rising (LH), or falling tone (HL) is realized higher, while the low tone of a lexical low (L), rising, or falling tone is realized lower under focus. Thus, register expansion affects both the top-line and the bottom-line of the register (Xu 1999, p. 69). In addition to register expansion on the focused word in Mandarin, f0 after the focused word is lowered, named post-focal compression (PFC) (Xu 1999, 2011, Xu et al. 2012). The current study suggests that Hindi also belongs to the

group of languages using register-based cues to express focus.

According to Xu (2011), PFC as a cue of focus marking is widespread among the languages of the world. However, PFC is a non-universal cue of focus marking. It has been reported for a variety of typologically distinct as well as for closely related languages (cf. Xu 2011, and references therein). In particular, Xu et al. (2012) show data from closely related languages of the same language family —Taiwanese Min, Taiwan Mandarin, and Beijing Mandarin—of which Beijing Mandarin uses PFC as a prosodic cue of focus marking while the other two do not. All three languages are tone languages and they use similar morphosyntactic means to encode focus. PFC is thus a possible prosodic indication of focus, yet not necessarily a required cue or the only cue. As is discussed in the next section, PFC is present in Hindi as well.

The proposed typology of prosodic focus marking (Kügler & Calhoun 2020) and the classification of languages according to PFC (Xu 2011) are related to each other. The proposals however do not mutually exclude each other, but the theory of PFC should be viewed as orthogonal to the proposed typology of prosodic focus marking. A crucial difference between the two proposals concerns the position within the sentence of the prosodic cues involved. PFC is restricted to post-focal components and divides languages into those that have PFC and those that do not. This division is too simple a restriction to cover the broad variation of prosodic focus marking depicted in Kügler & Calhoun (2020), especially since the variation of focus marking on the focused constituent does not predict the presence or absence of PFC. In fact, PFC appears to be independent of prosodic properties of a language. As discussed above, a similar prosodic system of two closely related languages does not imply that both languages have PFC (Xu et al. 2012). Surveying the studies on PFC we observe that languages with and without PFC group together in each of the three proposed language strategies of prosodic focus marking proposed by Kügler & Calhoun (2020). PFC is found to also mark focus in languages that use stress-based cues (e.g., German or English, Fery & Kügler 2008, Breen et al. 2010, Xu & Xu 2005) and in languages that use register-based cues (e.g., Xu 2011, Xu et al. 2012). For languages using phrase-based cues, the situation is not as clear as for the other two groups. One example is Korean, for which the facts on PFC are disputed. While Jun & Kim (2007) and Lee & Xu (2010) show data for lower realization of tones in the post-focal area, Jeon & Nolan (2017) present data where post-focal tones were not phonetically reduced. Hence, more studies are needed to understand the correlation between phrase-based focus marking an occurrence of PFC. This lack of evidence however does not affect the general point that the presence or absence of PFC is orthogonal to prosodic focus marking strategies.

There are only few studies that examine the perceptual relevance of prosodic cues that encode focus. These studies show that if a prosodic cue that is systematically realized in speech production is present in stimuli for perception, listeners are still able to detect the focus structure of a sentence (e.g., Botinis et al. 1999, Liu & Xu 2005, Vainio & Järvikivi 2006, Krahmer & Swerts 2007, Xu et al. 2012, Kügler & Gollrad 2015). In an intonation language, these prosodic cues usually refer to pitch accents (Pierrehumbert 1980, Pierrehumbert & Beckman 1988, Ladd 2008, Gussenhoven 2004). The function and meaning of a pitch accent is to highlight relevant information in a sentence; in case of focus, the pitch accent highlights the information of that constituent which the pitch accent is associated with. In general, the interpretation of a sentence meaning is dependent on phrase boundaries (e.g., Lehiste 1973) and on pitch accents (e.g., Schafer et al. 1996, Carlson et al. 2001, Watson et al. 2008). Pitch accents and boundary tones (phrase boundaries) are phonological categories of intonation that functionally encode meaning; in relation to the prosodic expression of focus, these cues are found in languages using stress-based cues. To what extent pitch register also carries meaning remains an issue. We propose that Hindi, which employs register-based cues to focus expression, provides evidence that variation in pitch register is meaningful and thus that pitch register can be interpreted as a phonological category like a pitch accent or a boundary tone.

### 1.1 Background on Hindi intonation

To investigate the role of pitch register as a potential phonological category in prosody, we begin with an overview of the prosodic marking of focus in Hindi. Hindi is characterized by a series of repeated rising contours (RRC), which is argued to be an areal feature of South Asian languages in general (Khan 2016). Every prosodic word is associated with an f0 rise, except the last one in the intonational phrase. These f0 rises are systematically and sequentially downstepped, meaning that

in general, each f0 rise reaches a lower peak than the preceding rise (Moore 1965, Harnsberger 1996, 1999, Harnsberger & Judge 1996, Patil et al. 2008).

Earlier studies analyzed the f0 rise as a rising pitch accent L*+H (Harnsberger 1996, 1999, Harnsberger & Judge 1996). More recent proposals interpret the f0 rise as a pair of phrase tones, as there is no systematic alignment of the low (L) and high (H) tones of the f0 rise with a particular prominent (or stressed) syllable (Féry 2017). It is an open debate whether Hindi has stress or not. While some authors argue for a word stress system of Hindi (Nair 2001), others argue against the existence of word stress (Ohala 1991, 1994). The variable alignment of an f0 rise is taken as evidence that an f0 rise is not associated with metrical prominence as in intonation languages like English or German (Féry 2010, 2017). Therefore, Féry (2010, 2017) proposes that, in her terms, Hindi is a phrase-language from the point of view of prosodic typology. In her phrase-level prosodic typology, she distinguishes between the traditional division of intonation languages and tone languages, and—as the new proposal—phrase languages.

Different studies on Hindi intonation show that focus is realized with post-focal compression but without any prosodic effect on the focused word (Harnsberger & Judge 1996, Patil et al. 2008). In those studies, focus appeared in two conditions: Broad focus (or all-new) and narrow focus (see Ladd 1980, for this distinction). According to Ladd, broad focus refers to a whole sentence in focus while a narrow focus is defined as any constituent which can be identified by the placement of a pitch accent (in English). In an all-new sentence, no element has been mentioned in the preceding context or was especially prominent in the common ground of the speakers. Narrow focus can be induced by a context asking explicitly for a particular constituent. As in many languages, Hindi has a designated syntactic position for focused elements immediately before the verb (Kidwai 2000, p. 116). By structural means, this syntactic pre-verbal focus position expresses prominence as was shown in a prominence rating study. Constituents in pre-verbal position received consistently higher prominence scores (Luchkina et al. 2015). Given that the position expresses prominence for structural reasons there is no obvious reason to mark a constituent appearing in this position also as prosodically prominent.

In a production study, Patil et al. (2008) investigated the effects of word order and information structural context on the prosodic realization of declarative sentences in Hindi. The analysis of Subject-Object-Verb (SOV) and Object-Subject-Verb (OSV) sentences in three information structural conditions (broad focus, narrow focus on the subject, or narrow focus on the object) revealed that constituents are in a strict downstep relationship regardless of word order and focus, and focus is mainly characterized by lowering the post-focal pitch register rather than pitch raising of the element in focus. This means that the characteristic f0 rises are still realized post-focally, though in a compressed pitch register. This is different from English and a number of other languages where no tonal cues are realized after a focused constituent (cf. Cruttenden 2006), which is commonly known as 'deaccentuation' (Ladd 2008). Similar to Hindi, but contrary to a deaccentuation view, Kügler & Féry (2017) showed for German that post-focal pitch accents are realized in a compressed pitch register. The amount of compression differs though between Hindi and German, where it is almost completely compressed. A compressed pitch register after focus is also well attested in some tone languages, e.g., Mandarin Chinese (Xu 1999) or Akan (Kügler & Genzel 2012). See the discussion on this cue in the Introduction above and Xu et al. (2012) for an overview. Comparing the facts on Hindi intonation with those of other languages discussed, we assume that Hindi uses register-based cues to express focus prosodically (cf. Kügler & Calhoun 2020).

Nonetheless, all previous studies suggest that PFC is a prosodic cue that is realized consistently. This is unlike other cues which may mark focus on the focused constituent. For instance, Genzel & Kügler (2010) showed that both the L tone and the H tone of the f0 rise were realized significantly lower and higher, respectively, on focused adjectives in Hindi. The interpretation was that the pitch register on the focused word was expanded. We can only speculate how the diverging results come about. One possible explanation is that different types of focus are expressed differently prosodically. Narrow information focus elicited by wh-questions in previous studies does not seem to require any prosodic focus marking on the focused constituent (e.g., Harnsberger & Judge 1996, Patil et al. 2008). Contrastive focus, as used in Genzel & Kügler (2010), seems to affect the prosodic marking of the focused constituent. It is possible that correcting an element from a previous discourse may require more prominence such that the corrected constituent is marked prosodically (see Kügler & Genzel 2012, for similar findings in Akan). An alternative view as an explanation for the divergent

findings could be that the prosodic cues of focus marking on the focused constituent in Hindi are not obligatory (Moore 1965). Whatever reason there might be, we can safely conclude that PFC is a cue that appears to be realized consistently in Hindi.

The goal of the present study therefore is to test whether PFC serves as a prosodic cue that may influence the interpretation of a sentence. To this end, we first run a production study to test whether speakers systematically realize PFC in a given sentence with a structure that differs from previous studies. Second, we run a sentence completion study with stimuli from the production study that either contained or did not contain PFC to test whether PFC serves as a perceptual cue for focus interpretation.

## 2 The production study

The production study was designed to test for the presence of PFC in a particular syntactic structure that would serve as an eligible structure for a perception study. For this reason, the structure of a contrastive ellipsis as given in (1a) was chosen, which consist of a subject (S), indirect object (IO), direct object (DO), a verb (V), the negation particle (NEG) plus a complementizer (C), and the following contrasted object of the remnant. In the example, a bracketed segment with a subscripted F stands for the focused element which contrasts with the object of the remnant.

(1) a. S IO DO V NEG C Contrast

b. Rāhul=ne mā=ko [davāī]$_F$ dī nā ki [ghaṛī]$_F$
Rahul=ERG mother=DAT medicine gave NEG C watch
'Rahul gave the medicine to mother and not the watch.'

c. Rāhul=ne [mā=ko]$_F$ davāī dī nā ki [nānī=ko]$_F$
Rahul=ERG mother=DAT medicine gave NEG C grandmother=DAT
'Rahul gave the medicine to mother and not to grandmother.'

Contrastive ellipsis in Hindi is particularly effective for testing the influence of PFC since (i) the matrix clause is structurally ambiguous, and (ii) focus appears in the matrix clause contrasting with an element in the remnant. In (1), either the indirect object (IO), i.e. *mother* (1c), or the direct object (DO), i.e. *the medicine* (1b), contrasts with the object in the remnant. According to Drubig (1994), contrastive ellipsis, or *replacives* as he coins the term, have parallel foci: the remnant contrasts with a focused argument in the matrix clause. According to Repp (2010) both foci of a parallel construction can be considered contrastive.

Given this particular focus structure, the expectations based on Patil et al. (2008) are that speakers would realize PFC after the focus within the matrix clause. For (1b), with focus on the direct object, we expect PFC on the verb and up to the complementizer. For (1c), with focus on the indirect object, we expect PFC on the following direct object, the verb, and the complementizer. The aim of the production study is to gain quantitative data on PFC in contrastive ellipsis, as the speech production data are used as stimuli for the sentence-completion study.

### 2.1 Methods

Here we provide an overview of the methods used for the production study. We introduce the design of the study, the speech materials used (2.1.1), information about the participants (2.1.2), the procedure of data elicitation (2.1.3), and information about data pre-processing and statistical analysis (2.1.4).

#### 2.1.1 Design and Materials

The experiment involved the factor REMNANT CONTRAST: the object in the remnant contrasts either with the direct object (1b) or on the indirect object (1c). Six different sentence pairs of the type in (1) were constructed (listed in the Appendix), resulting in a total of 12 test sentences. Each sentence was presented without context. The 12 test sentences were presented to each speaker in a pseudo-

randomized manner; items from three other unrelated experiments were interspersed as fillers. Four pseudo-randomized lists were prepared to minimize order effects.

### 2.1.2 Participants

Thirty native speakers of Hindi participated in the experiment. All were female students at the University of Delhi, India, and were residents of Delhi and surrounding areas. Each speaker was paid Rs. 150 for participation and took approximately 45 minutes to complete the experiment.

### 2.1.3 Procedure

The experiment was carried out using presentation software. First, participants were equipped with a set of headphones and a microphone headset, and familiarized with the task through written and verbal instructions, followed by two practice trials. Each trial consisted of a presentation of the experimental sentence on the computer screen, written in Devanagari script. Participants were instructed to speak out the sentence displayed on the screen as naturally as possible. If the sentence was uttered without any hesitations or false starts, the next trial was presented. If there were hesitations, participants were asked to repeat the sentence. Presentation flow was controlled by the experimenter, and participants were allowed to take a break whenever they wanted. The sentences produced by participants were recorded at the University of Delhi on a DAT tape recorder.

### 2.1.4 Data pre-processing and statistical analysis

The recordings were re-digitized from DAT at a sampling frequency of 44.1 kHz and 16 bit resolution. In total, 360 sentences (6 items x 2 conditions x 30 speakers) were recorded. The sentences were labeled by hand at the level of the constituent, as shown in (2). The vertical lines mark constituent boundaries. Standard cues for segmental labeling were employed, and boundaries were set at zero crossings (Turk et al. 2006).

(2) | Rāhul=ne | mā=ko | davāī | dī | nā | ki | ghaṛī |
Rahul=ERG mother=DAT medicine gave NEG C watch
'Rahul gave the medicine to mother and not the watch.'

The pitch analysis was conducted using a Hanning window of 0.4 seconds length with a default 10 ms analysis frame. The pitch contour was smoothed using the Praat (Boersma & Weenink 2018) smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations.

For each constituent in (2), the f0-maximum, the f0-minimum, and the duration were detected using a Praat script. In each constituent, only those f0-maxima were measured that followed the f0-minimum in that constituent; this was done in order to exclude maxima due to transitions from preceding H tones. The f0-maximum after the L tonal target represents the H tone in the rising pitch gesture.

The statistical analysis relied on the dependent variables, f0-maximum, and duration measured at the location of the indirect object, the direct object and the verb. All dependent measures were log-transformed to meet the assumption of the regression model. A multilevel model (Bates et al. 2015) was fit, using crossed random factors SPEAKER and ITEM, and REMNANT CONTRAST (IO, DO) as fixed factor.

## 2.2 Results

Figure 1 shows the aggregated mean f0 over all speakers and items for the two conditions. For each constituent given on the x-axis, the scaling of the L and H tones are plotted. The solid line represents the condition with the contrast on the direct object (1b), the dashed line with the contrast on the indirect object (1c). The stylized f0 contour is remarkably similar across the two conditions except for the scaling of the H tone on the direct object. Comparing the individual H tones, a clear downstep pattern can be observed. Each H tone on each constituent up to the verb is realized lower than the corresponding previous one. This prominent downstep pattern is in line with Moore (1965) and Patil et al. (2008).

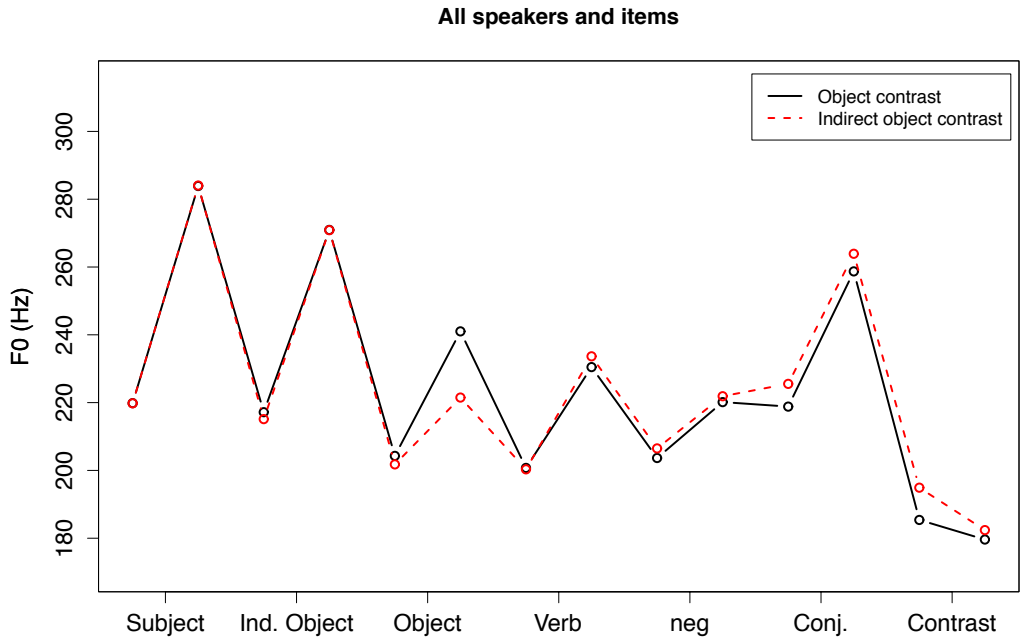


Figure 1. f0 plot of L and H tone per prosodic word averaged across all speakers (n = 30) and items (n = 6). The black solid line represents the object contrast condition, the dashed red line the indirect object contrast condition.

In order to evaluate the amount of post-focal lowering, we compare the f0-maximum and duration of the object constituents and the verb in turn (Tables 1–6). There is no significant difference in f0 nor in duration on the indirect object (Tables 1 and 2). A significant difference in f0 and in duration is found on the direct object (Tables 3 and 4). In case of contrast on the indirect object, the following direct object is realized with significantly longer duration and with a significantly lower f0 on the H target, which meets our prediction of post-focal pitch register compression.

| | Estimate | SE | t value | Significance |
|---|---|---|---|---|
| (Intercept) | 5.603747 | 0.022571 | 248.27 | |
| Condition=IO | 0.015450 | 0.008016 | 1.93 | n.s. |

Table 1. Difference of f0-scaling in terms of f0-maximum on IO, baseline is condition DO.

| | Estimate | SE | t value | Significance |
|---|---|---|---|---|
| (Intercept) | 0.53529 | 0.04529 | 11.819 | |
| Condition=IO | 0.01310 | 0.01079 | 1.213 | n.s. |

Table 2. Difference of duration on IO, baseline is condition DO.

| | Estimate | SE | t value | Significance |
|---|---|---|---|---|
| (Intercept) | 5.51883 | 0.02491 | 221.55 | |
| Condition=IO | -0.03287 | 0.00694 | -4.74 | * |

Table 3. Difference of f0-scaling in terms of f0-maximum on DO, baseline is condition DO.

| | Estimate | SE | t value | Significance |
|---|---|---|---|---|
| (Intercept) | 0.418799 | 0.043654 | 9.594 | |
| Condition=IO | -0.048867 | 0.005335 | -9.160 | * |

Table 4. Difference of duration on DO, baseline is condition DO.

On the verb, there is no significant difference in f0 or duration (Tables 5 and 6), which goes against our expectation since PFC after the focused objects was expected. In particular, for the direct object focus condition, we expected lower scaling of the verb compared to the indirect focus condition. The scaling of the H tone on the verb is identical between conditions, and remarkably, it is higher than the compressed one on the direct object.

| | Estimate | SE | t value | Significance |
|---|---|---|---|---|
| (Intercept) | 5.46828 | 0.02856 | 191.47 | |
| Condition=IO | -0.01611 | 0.01907 | -0.84 | n.s. |

Table 5. Difference of f0-scaling in terms of f0-maximum on the verb, baseline is condition DO.

| | Estimate | SE | t value | Significance |
|---|---|---|---|---|
| (Intercept) | 0.326627 | 0.033614 | 9.717 | |
| Condition=IO | -0.010851 | 0.007492 | -1.448 | n.s. |

Table 6. Difference of duration on the verb, baseline is condition DO.

The data discussed so far represent mean values. Since the statistical analysis applies SPEAKER as a random factor, the variation between speakers is taken into account. However, we observe a considerable amount of speaker variation. The majority of speakers clearly express the difference between conditions prosodically. PFC was realized after the focused indirect object, and PFC was not realized after the focused direct object. Speaker variation arises in two different ways. First, the degree of PFC varies: on average, speakers realized a PFC between 10 and 30 Hz. Second, some speakers did not realize PFC at all after the focused indirect object; for these speakers, there was no prosodic difference observed between the conditions. One such example is shown in Figure 2, in which there is no difference in f0 on the direct object between conditions. This speaker had an overall flat intonation contour.

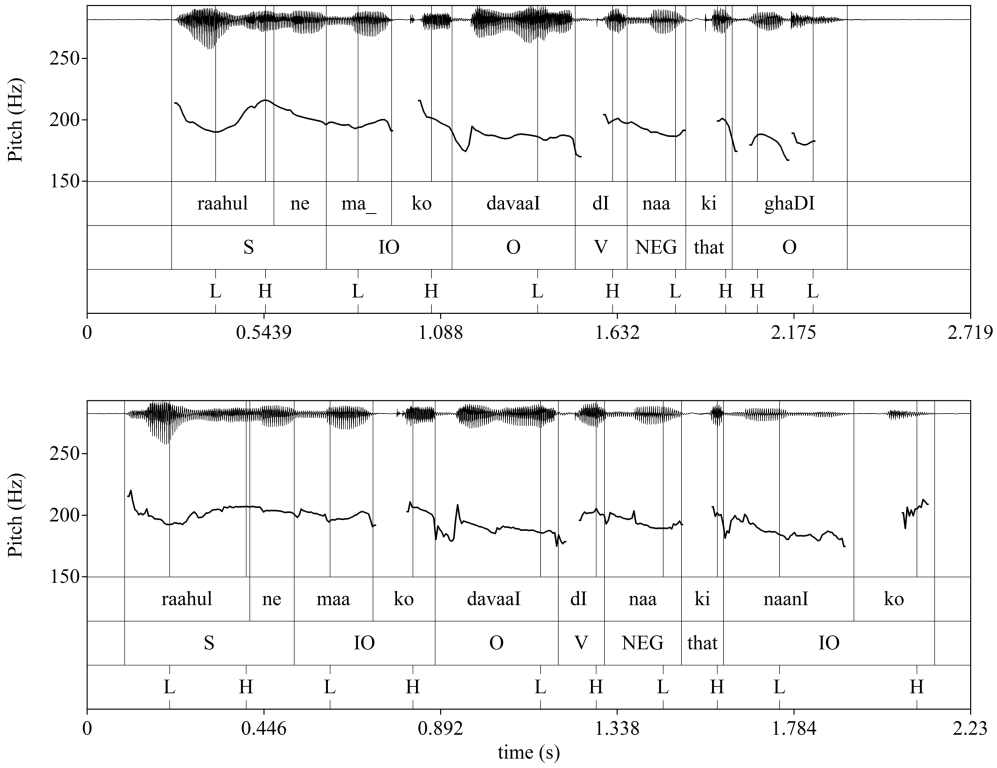


Figure 2. A comparison between conditions for a speaker who does not express PFC on the direct object: the object contrast condition is shown in the upper panel, and the indirect object contrast condition in the lower panel.

### 2.3 Discussion

In the production data, the f0 maximum measurement taken on each constituent turned out to be the acoustic cue of the constituent-final H tone that is proposed as a tonal property of Hindi (e.g., Harnsberger 1996, Féry 2017). Comparing the f0 maxima within each condition, we take the f0 maximum as a reference of the H tone. In both conditions, the scaling of subsequent H tones is considerably lower than the previous H tone, and we conclude that downstep between consecutive H tones occurs, in line with Harnsberger (1996), Patil et al. (2008), Féry et al. (2016b), and Féry (2017).

Comparing the f0 maxima between conditions, we observe no difference in scaling of the H tones on the first constituent, the subject, nor on the second constituent, the indirect object. In one condition, the indirect object is in pre-focal position, in the other condition, the indirect object is in focus. The lack of f0 difference on the indirect object leads thus to the conclusion that focus is not prosodically marked on the focused constituent, in line with Patil et al. (2008). The difference between conditions occurs on the third constituent, i.e. the direct object. If the direct object is in focus, the scaling of the H tone is higher than if it occurs in post-focal position. Given that a constituent in focus is not prosodically marked on that H tone, we conclude that the difference in scaling of the H tone on the direct object is a consequence of it being in post-focal position.

The striking result is that PFC is only realized in the indirect object condition, i.e. on the following direct object. In the direct object condition, no PFC was observed on the following verb. It may be the case that because the direct object appears in a syntactic position that is associated with focus (cf. Kidwai 2000) speakers do not realize any prosodic cues of focus. If this were the case, we would expect that any constituent appearing in the pre-verbal position bears positional prominence,

and no prosodic cues were needed.

As an alternative explanation, it may be the case that the direct object and the verb are phrased together such that no PFC effect occurs. In fact, Bali et al. (2009) observed variable phrasing in Hindi based on chunking experiments where native speakers had to indicate which parts of speech belong together. Their conclusion was that there might be "an underlying notion of verb phrase in Indian languages", which could point to some kind of integration of the pre-verbal constituent and the verb. The interaction between a constituent in pre-verbal focus position and variation in the presence of PFC needs further investigation and is left for future research.

If we compare Hindi intonation to that of Mandarin, Hindi appears to be similar to Mandarin as it shows PFC. However, in our data no register expansion was found on the focused constituent. The similarities between Hindi and Mandarin, as well as a lack of clear evidence of stress in Hindi point to the fact that Hindi uses register-based cues to express focus (Kügler & Calhoun 2020).

The fact that Hindi does not mark focus prosodically on the focused constituent raises the question of whether listeners may recognize focus if it is not prosodically marked as it is in English or German (cf., Breen et al. 2010, Kügler & Gollrad 2015). However, focus does have a prosodic effect in Hindi; it is simply found on phrases that follow the focused constituent itself. Hence, listeners may rely on the cue PFC to identify focus in Hindi. This will be tested in a sentence completion study in the next section.

The observed speaker variation opens a further puzzling issue here. If we assume that focus is perceived only after the focused constituent by means of PFC the question arises if listeners are able to perceive focus from speakers that produce hardly any post-focal compression. We observed different degrees of post-focal compression ranging from zero compression to slight compression (10 Hz on average) to full compression (30 Hz on average). Consequently, we tested the perception of focus in a sentence completion study.

## 3 The completion task

In order to test listeners' ability to recognize prosodic focus in Hindi, we set up a forced-choice auditory sentence completion study using a contrastive ellipsis structure. In this study we investigated listeners' perception of focus to the differing degrees of PFC found in the production data.

### 3.1 Methods

This section provides information about the methodological aspects of the auditory sentence completion study. It introduces the speech materials used as stimuli for perception (3.1.1), gives information about the participants of the study (3.1.2), the procedure of the sentence completion study (3.1.3), and the predictions (3.1.4).

#### 3.1.1 Speech materials

For the auditory sentence completion experiment, fragments of a contrastive ellipsis structure as in (3a) were used. The matrix clause of the contrastive ellipsis is ambiguous up to the complementizer of the remnant. Only the remnant contains the relevant information which of the two objects of the matrix clause (the direct object in (3b), or the indirect object in (3c)) are contrasted. If we were to run a reading task, we may expect a tendency for the direct object to be contrasted (cf. (3b)) because it appears in pre-verbal position which is assumed to be the syntactic focus position (Kidwai 2000, Luchkina et al. 2015). Another possible outcome may be that readers have no clue for the interpretation of the contrast (chance-level recognition of about 50%).

(3) a. Rāhul=ne mā=ko davāī dī nā ki …
Rahul=ERG mother=DAT medicine gave NEG C …
'Rahul gave the medicine to the mother and not…'

b. ghaṛī
watch
'the watch.'

c. nānī=ko
grandmother=DAT
'to the grandmother.'

For the completion study, three naturally spoken sets of stimuli from the production experiment by six speakers were chosen, which contained maximally and minimally informative prosody. The criterion for the selection of the speakers was the degree of PFC realized by the speakers. One set of items consisted of realizations with (almost) no PFC, the second set consisted of realizations of about 10 Hz PFC, and the third set consisted of realizations of about 30 Hz PFC. Each set contained the condition with a contrast of the indirect object (3c) and the condition with a contrast on the direct object (3b). Together, five items per set were used. The experiment thus contained 60 stimuli (5 items x 2 conditions x 6 speakers). All stimuli were cut after the complementizer *ki* and before the remnant contrast, either the indirect object or the direct object. These sentence fragments were then used as stimuli for the completion study.

### 3.1.2 Participants

Eighteen (13 M, 5 F) native Hindi listeners took part in the experiment. Participants were recruited in and around Berlin, Germany. Their mean age was 27 years. In an interview about their social and language background, all participants declared that they were native speakers of Hindi and had no speech or hearing impairment. The participants were naïve as to the purpose of the experiment. Each participant was paid €8.

### 3.1.3 Procedure

The experiment took place in a quiet room at participants' homes and at the University of Potsdam. The completion experiment was carried out using Praat's ExperimentMFC function (Boersma & Weenink 2018). As instructions, participants were asked to listen to the sentence fragment and then to choose between one of two possible sentence completions by clicking either on the button containing the indirect object or on the button containing the direct object. The objects were presented in Devanagari script.

The 60 sentence fragments were randomly played through Sennheiser HD 25 headphones. The rectangles containing the response categories were displayed on the left-hand side and right-hand side in the lower half of the screen. The appearance of the position of the buttons containing the direct and indirect object were counterbalanced. Each stimulus had an initial duration of silence of 0.5 seconds. In total, 1080 trials were collected (18 listeners x 5 sentences x 2 conditions x 6 speakers). The experiment lasted approximately 15 minutes.

### 3.1.4 Predictions

We applied a within-subject design for the forced choice sentence completion task. The predictions were the following: If PFC is present, listeners will identify the object contrast correctly. If PFC is absent, listeners will show a chance level identification of the contrast. These hypotheses include the observed speaker variation in condition IO. When presented with recordings of speakers who produced no PFC, it was expected that listeners would not be able to perceive the object contrast, resulting in a chance-level identification.

## 3.2 Results

We ran a linear mixed effects model (Bates et al. 2015) with CONDITION (with the two levels direct object DO and indirect object IO) as fixed factor and LISTENER and ITEM as random factors. The reference level in the model was DO. We applied random slopes and intercepts for LISTENER and random slopes for ITEM. The analysis relied on the choice of answer (counted as correct sentence completion) as a dependent variable. The model reveals a significant effect for CONDITION, as reported in Table 7. Hence, we observe that the two conditions differ significantly in terms of correct sentence completion. While in condition IO, listeners identified the correct sentence completion on

average in 70.9% of the cases, in condition DO, listeners identified the correct sentence completion on average only in 44.1% of the cases.

| | Estimate | SE | z value | Pr(> \|z\|) | Significance |
|---|---|---|---|---|---|
| (Intercept) | -0.5933 | 0.4618 | -1.285 | 0.1988 | |
| Condition=IO | 1.9398 | 0.8316 | 2.333 | 0.0197 | * |

Table 7. Report of the linear mixed effects model applying CONDITION as fixed factor with correct sentence completion ratings as dependent variable.

The difference in the identification of the correct sentence completion between conditions is shown in Table 8 for each of the five sentences (items) used. On average, correct sentence completion identification is higher for condition IO. However, one of the sentences, sentence 5, shows similar completion rates for condition DO and condition IO.

| Sentence | Condition DO | Condition IO |
|---|---|---|
| 1 | 0.287 | 0.824 |
| 2 | 0.519 | 0.639 |
| 3 | 0.370 | 0.769 |
| 4 | 0.472 | 0.741 |
| 5 | 0.556 | 0.574 |

Table 8. Mean identification of correct sentence completion split by test sentences.

Since we found speaker variation with respect to the amount of PFC in condition IO in the production study, in the sentence completion study we included items from six speakers who produced different degrees of PFC. This allows us to investigate whether the correct sentence completion identification differs between speakers that realize different degrees of PFC. The hypothesis was that the presence of PFC leads to correct sentence completion, which was overall confirmed by the significant effect of CONDITION reported in Table 7. Table 9 displays the mean identification of correct sentence completion split by speakers. Speakers 5 and 21 realized no PFC (given as 0 Hz) in the production study. Speakers 26 and 30 produced clear PFC of about 30 Hz on average. Speakers 8 and 18 produced PFC, though to a lesser degree of 10 Hz on average.

| Speaker | Degree of PFC | Condition DO | Condition IO |
|---|---|---|---|
| 5 | 0 Hz | 0.489 | 0.611 |
| 21 | 0 Hz | 0.422 | 0.622 |
| 8 | 10 Hz | 0.478 | 0.800 |
| 18 | 10 Hz | 0.422 | 0.678 |
| 26 | 30 Hz | 0.456 | 0.767 |
| 30 | 30 Hz | 0.378 | 0.778 |

Table 9. Mean identification of correct sentence completion split by speaker.

In order to test for speaker variation we ran a linear mixed effects model with CONDITION and SPEAKER as fixed factors, and LISTENER and ITEM as random factors. The factor SPEAKER was binary coded: The two speakers who did not produce PFC at all were grouped together against the other four speakers who produced PFC. The reference level in the model for CONDITION was DO, and for SPEAKER it was the speaker group with no PFC. We applied random slopes and intercepts for LISTENER and random slopes for ITEM. The results of the model including speaker variation is given in Table 10. We observe that the significant effect of CONDITION from the simple model (Table 7) disappears. We did not find a significant effect for SPEAKER. However, we found a significant interaction of CONDITION and SPEAKER (Table 10). This suggests that speaker variation indeed mattered for the correct sentence completion. The interaction plot in Figure 3 shows a clear effect of CONDITION for speakers who produced PFC: with recordings of speakers who used PFC, we found on

average a higher (75.6%) correct sentence completion identification. With recordings of speakers who did not produce PFC, lower completion rates (61.6%) on average were obtained.

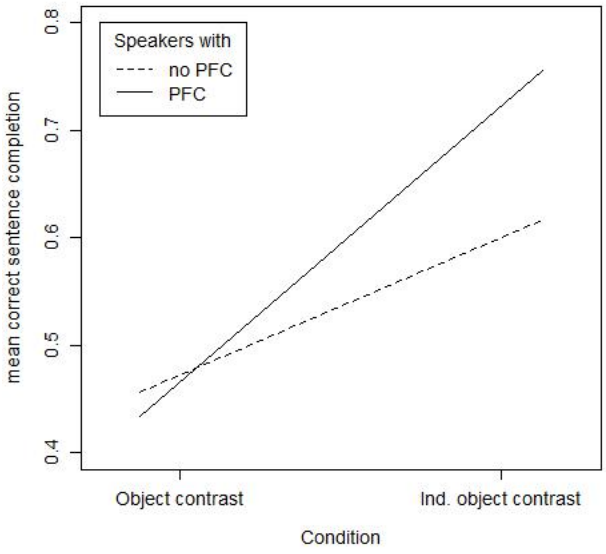


Figure 3. Interaction plot: responses to recordings of speakers without PFC (dotted line) compared to those of speakers with PFC (solid line), split by condition (contrast on DO vs. contrast on IO).

A post-hoc paired t-test confirmed this pattern ($t = –8.39$, $df = 359$, $p < 0.001$). A model comparison between the simple model and the speaker variation model showed significant improvement of the speaker variation model measured as the difference in deviance (Bates et al. 2015) ($\Delta D = 14.2$, $p < 0.001$).

| | Estimate | SE | z value | Pr(> \|z\|) | Significance |
|---|---|---|---|---|---|
| (Intercept) | -0.5075 | 0.4835 | -1.050 | 0.29393 | |
| Condition=IO | 1.3397 | 0.8648 | 1.549 | 0.12135 | |
| Speaker=PFC | -0.1276 | 0.2164 | -0.590 | 0.55552 | |
| Interaction | 0.9584 | 0.3096 | 3.096 | 0.00196 | ** |

Table 10. Report of the linear mixed effects model including CONDITION and SPEAKER as fixed factors with correct sentence completion ratings as dependent variable.

### 3.3 Discussion

The auditory sentence completion experiment revealed that the presence or absence of PFC matters for the identification of focus. PFC was present in some of the experimental stimuli of condition IO. As was shown in the production data of the contrastive ellipsis, on average, speakers produced a rising tonal pattern on a post-focal constituent, in this case on the direct object which followed the focused indirect object (Fig. 1). When PFC was present in the stimuli of the sentence completion experiment, listeners were able to correctly identify the sentence completion with a contrast of the indirect object. Listeners thus interpreted the focus structure of the elliptical sentence on the basis of the prosodic cue of PFC.

PFC was consistently not present in condition DO when the contrastive ellipsis contained a contrast on the direct object. In the production data, on average, speakers did not produce PFC on the following verb (which was unexpected though, cf. Section 2.3 above). For the sentence completion experiment, the absence of PFC resulted in listeners performing roughly at chance. Note that

the direct object was not perceived as more prominent due to its preverbal position (contrary to ratings in Luchkina et al. 2015). If it were, listeners should have identified the correct sentence completion in condition DO, which they did not.

If we compare the results of the present study with results on focus perception in general, we may conclude that native Hindi listeners are able to recognize an intended focus structure (e.g., Kügler & Gollrad 2015, del Mar Vanrell 2011). However, while Kügler & Gollrad (2015) and del Mar Vanrell (2011) manipulated phonetic cues on the focused word, the current study presents evidence that the prosodic realization of the post-focal area matters for focus perception.

With respect to speaker variation in the production of PFC, we can safely conclude that speaker variation has an impact on focus perception: sentences from speakers who did not realize PFC in condition IO were completed roughly by chance (the correct sentence completion rate of 61.6% did not differ significantly from the DO condition because CONDITION and SPEAKER did not show a significant effect in the model). Sentences from speakers who did produce PFC in condition IO were completed correctly (75.6% on average). PFC is thus a functional cue to focus identification in Hindi. The prosodic information of pitch register is a cue that resolves local ambiguities.

## 4 General discussion

The current study was motivated by experimental findings on the prosodic expression of focus in Hindi, which showed that the focused word or constituent itself is not prosodically marked by native speakers of Hindi (Harnsberger & Judge 1996, Patil et al. 2008). The question investigated in the current study was whether other prosodic cues such as PFC indicate the focus structure of a sentence. Independent of prosody, syntactic means such as the use of a preverbal focus position (Kidwai 2000) may also indicate focus prominence of a constituent in Hindi (Luchkina et al. 2015). Instead of prosodically marking the focused constituent itself, the prosodic marking of focus in Hindi is achieved by a compression of the pitch register on post-focal constituents. South Asian languages in general, independent of their language family origin, appear to constitute a group of languages which share similarities in their intonation systems (Féry 2010, 2017, Khan 2016). These languages are characterized by 'repeating rising contours (RRCs)' (Khan 2016). However, a considerable amount of cross-linguistic variation in terms of the alignment of the L and H tone of an f0 rise and of the number H tones within a prosodic domain have been observed, and Khan (2016) concludes that this observation clearly calls for more carefully controlled studies on the intonational systems of South Asian languages.

According to Féry (2010), Indian languages share not only the characteristic RRCs but some of the languages, at least, also use similar means to prosodically mark focus. Post-focal compression is found in Hindi, and dephrasing which goes hand-in-hand with post-focal compression is found in Bengali (cf. also Hayes & Lahiri 1991, Khan 2008). Crucially and in contrast to Hindi, in Bengali dephrasing is accompanied by deletion of post-focal tones (Khan 2016). For some other languages the facts around focus marking are not so clear. In Tamil, for instance, Keane (2014) shows some data where post-focal compression may occur similar to the Hindi case. However, the example given is a very short sentence and the scaling of the f0 rise is not directly compared to a broad focus. Therefore one cannot draw a clear conclusion about the prosodic marking of focus in Tamil except for a clear case of phrasing: focus appears to require a prosodic boundary that is indicated by an f0 rise on the focused constituent.

In sum, pitch compression patterns have been described in other South Asian languages, and there is some uncertainty as to how they should be analyzed. In his analysis of the intonation of Bangladeshi Standard Bengali, Khan (2008, 282ff) observes that post-focal pitch accents and boundary tones are frequently deleted. In order to account for the post-focal effects he suggests four possible interpretations why post-focal tones are deleted in Bengali — leaving a conclusive proposal open to further investigation: (i) a loss of metrical prominence on post-focal material, (ii) prevention of pitch accent assignment to post-focal accentual phrases, (iii) pitch range compression, and (iv) deletion of post-focal accentual phrases. The variation in metrical structure (i) and (ii) or phrasing (iv) leads to the deletion of post-focal tones due to reduced or no structure with which tones could be associated. Variation of pitch range (iii) leads to a phonetic effect that existing post-focal tones are simply completely reduced.

A comparison with Hindi and with the data of our study shows a crucial difference which becomes relevant for interpreting the post-focal effects in Hindi: in Bengali, speakers use tonal deletion as a strategy to compose the post-focal area prosodically. Contrary, in Hindi, speakers do realize post-focal tones (Fig. 1). The f0 rises seem to be obligatory, though realized in a compressed pitch register. This fact leads to a crucial difference with respect to the nature of post-focal compression: In Bengali, the compression appears to be almost complete while in Hindi, pitch register compression is partial, leaving enough space for the realization of post-focal f0 rises. This difference in the amount of pitch register reduction is similar to other languages. For instance, both in English and German, post-focal pitch register compression is complete to the effect of there being no (or very reduced) pitch accents (e.g., Kügler & Féry 2017, for German) while in Mandarin, partial PFC occurs with the effect that lexical tonal distinctions are maintained post-focally (Xu 1999).

If we were to apply Khan's (2008) analytical options to Hindi, presumably three of them would naturally fail because they relate to the effect of post-focal tonal deletion, which obviously does not happen in Hindi (Fig. 1). Only option (iii) would remain to discuss for Hindi, which interprets range compression as an epiphenomenon of focus. This option proposes that any metrical structure of post-focal constituents remains and post-focal tones exist. However, phonetically their realization is reduced due to range compression (Khan 2008, p. 284). If we were to assume this solution for Hindi, there is a conceivable objection of PFC being an epiphenomenon of focus. First, it is unclear for Hindi if any metrical structure is present at all given the unclear facts about word stress. Let us assume for now Hindi has a metrical structure as given in (4).

(4)

| | | | | | |
|---|---|---|---|---|---|
| ( | | X | | | )$_{\iota}$ |
| (L$^{X}$ H)$_{\phi}$ | (L$^{X}$ H)$_{\phi}$ | (L $^{X}$H)$_{\phi}$ | (L$^{X}$H)$_{\phi}$ | (L$^{X}$ H)$_{\phi}$ | |
| (Rāhul=ne) | (mā=ko$_{F}$) | (davāī) | (dī) | (nā ki) | |

Instead of projecting stress-based metrical prominence per word, each $\phi$-phrase receives a metrical beat and is assumed to be tonally marked by a rise represented as an initial L and final H phrase tone. The focused constituent *mā=ko* receives a further metrical beat at the $\iota$-phrase level to represent the highest prominence in the phrase. Focus in English, Bengali, or Hindi for this matter, metrically falls out as the most prominent element. This metrical prominence is expressed in terms of prosodic cues on the focused constituent to signal the focus as being prominent. It is a higher scaled pitch accent in English (Breen et al. 2010) or a focal H boundary tone in Bengali (Khan 2008). In Hindi, no such effect is realized on the focused constituent, which makes it hard to assume any metrical prominence on the focused word as given in (4). Second, if we were further to assume that post-focally, metrical structure remains intact after focus we run into problems explaining the difference between pre-focal and post-focal metrical prominences, since each $\phi$-phrase received its particular metrical beat. If the focused constituent bears the highest prominence, and pre-focally and post-focally, metrical structure is present, the PFC effect does not fall out. On metrical grounds, the difference in pitch register between pre-focal and post-focal constituents cannot be explained.

Based on these facts we assume that PFC is not an epiphenomenon in Hindi. The proposal we would want to advocate for is that PFC is a phonological category of the prosodic system of Hindi. Any prosodic category known so far, e.g. a pitch accent or boundary tone, expresses a function and its phonological form can be related to an interpretative function. In Bengali or English, register compression may be interpreted as a phonetic epiphenomenon of focus because compression is complete. Partial register compression however is meaningful in that tones can still be realized in that domain and the function these tones have can thus be expressed. In order to interpret the results of our study in the following sections, we relate our results to findings of other languages in terms of the perception of focus, with respect to speaker variation and concerning the role of processing of prosodic categories. Since the effect of PFC is comparable to effects of pitch accents or boundary tones, we conclude that PFC in Hindi should be viewed as a phonological category. This would mean that pitch register is part of the abstract phonological representation of an utterance just as pitch accents, boundary tones, and prosodic constituents are.

### 4.1 On the perception of focus

There are far fewer studies that investigate the perception of focus than there are studies on the

production of the prosodic expression of focus. From these studies looking at the interplay of prosodic cues and their relevance to the perception of focus structure, we can observe that if certain prosodic cues of focus are present in the stimuli, in general, listeners are still able to detect the focus structure of a sentence (Botinis et al. 1999, Liu & Xu 2005, Vainio & Järvikivi 2006, Krahmer & Swerts 2007, Xu et al. 2012, Kügler & Gollrad 2015). This finding holds for typologically diverse languages independent of their distinct prosodic systems and independent of the prosodic cues these languages employ in the expression of focus (cf. Kügler & Calhoun 2020). For instance in German, speakers predominantly use an increase in f0 on the focused constituent (Baumann et al. 2007, Féry & Kügler 2008) but also duration and intensity (Baumann et al. 2007). For focus perception in a semantic congruency task, Kügler & Gollrad (2015) clearly showed that an increase of f0 on the focused constituent was highly accepted in a contrastive focus context.

Contrary to German, in Mandarin speakers produce an expanded pitch register on the focused constituent and PFC on following post-focal constituents (Xu 1999). In perception, both cues seem to play an important role for the identification of a focus structure (Liu & Xu 2005, Xu et al. 2012). Interestingly, the correct focus identification rates were high in case of sentence-initial or sentence-medial focus while the correct identification rates drop down considerably in case of sentence-final focus (Liu & Xiu 2005, Xu et al. 2012, Wang et al. 2012). This fact may speak in favor of PFC as a relevant additional cue for focus perception in Mandarin. Interestingly, Botinis et al. (1999) report a similar drop of final focus identification rate in Swedish, English, and Greek. Given the fact that final focus is less clearly identified in the three languages, there is some indirect evidence that the f0 drop from the focused constituent and the following post-focal register compression may play a considerable role in the identification of the focus structure of a sentence because all of these languages show some degree of PFC in case of non-final focus. In final focus, these cues are absent and thus final focus is less informative to listeners (Botinis et al. 1999, p. 1560).

A different perspective on the perception of focus with respect to PFC comes from the study by Wang et al. (2012). They investigated whether speakers of a language who produce PFC (in this case Mandarin) and speakers of a language who do not produce PFC (in this case Tsat, a Malayo-Polynesian language spoken on the island of Hainan, in southern China) perceive the focus structure of a sentence depending on whether the prosodic cue of PFC is present or not. All Tsat speakers were also L2 speakers of Mandarin. In fact, they found that the non-PFC speakers (Tsat speakers) perceived focus with much lower accuracy than PFC speakers (native Mandarin speakers) when listening to Mandarin sentences. Wang et al. (2012, p. 666) conclude that "PFC is probably hard to be noticed by speakers from a non-PFC language". This finding suggests that the prosodic cue to identify focus needs to be functionally relevant for listeners in their native language grammar. It suggests further that PFC is an instance of a phonological category if listeners use this cue to identify a focus structure. From L2 research it is known that sounds of a native language are perceived more easily than those acquired later in life of an L2 (Hume & Johnson 2003, and references therein). Usually, listeners match the speech signal to phonological categories of their L1, and if their L1 does not have an appropriate category listeners choose a category close enough to the speech signal (Hamann 2009). If the cue is not phonologically functional in a language, as in the case of speakers from a language without PFC (Tsat speakers in this case), these speakers are unable to identify focus in a second or foreign language because there is no such a phonological category in their L1. PFC for them would be phonetic noise in the signal. Hence, PFC could be viewed as a phonological category to be present in the grammar in one language, though absent in another. This would speak for the fact that PFC is a phonological category in the intonational grammar of Hindi. A reviewer suggested further interpretation which we think is relevant here: if we assume that Tsat speakers lack the intonational category of pitch register, they may have more difficulty in perceiving it. The difficulty in perception is comparable to processing other non-native prosodic phenomena such as stress; Peperkamp & Dupoux (2002) showed that French speakers were unable to identify stress simply because stress is not a component of French prosody. They interpreted the perceptual effect of the absence of a category as "stress-deafness". Similarly, we could interpret the absence of pitch register in Tsat prosody and the perceptual effect thereof as "register-deafness".

### 4.2 The role of speaker variation

A note is due on the issue of speaker variation. Many studies have found speaker variation with respect to the prosodic expression of focus or prominence. For instance, Baumann et al. (2007) observe that speakers of Standard German vary in their use of prosodic cues to express focus. German is a language that uses stress-based cues to express focus such as f0, duration, and intensity (cf. Kügler & Calhoun 2020). According to Baumann et al. (2007), some speakers use higher f0 while some others use longer duration for the expression of focus in the same task. The speakers thus achieve the perceptual impression of prosodic prominence by means of different prosodic cues.

In terms of f0, Gussenhoven (2002, p. 52) reports on a compensating strategy to express prosodic focus: f0-peaks may be later aligned or higher scaled. In both cases, they signal strong prominence. The effect arises because both realizations need more time to be produced compared to neutral or broad focus. In a perception study, Ladd & Morton (1997) argued for this relationship. They presented speech stimulus continua to listeners with higher or later aligned f0 peaks, and listeners were asked to indicate whether they perceived the stimulus sentence as in an "everyday occurrence" or as an "unexpected occurrence". The results of the their experiment confirmed the hypothesis: late f0 peaks are "intrinsically more emphatic" (Ladd & Morton 1997, p. 332), and the category boundary in the identification task shows up where the f0 peak itself is not that high, in their data at 132 Hz. In the early aligned stimuli (i.e., stimuli with higher f0 peaks), the category shift from "everyday occurrence" to "unexpected occurrence" only started at a stimulus with higher f0, in their stimuli at 144 Hz.

The different strategies that speakers employ to express focus prosodically appear to have in common that they aim for marking a focused constituent as prosodically prominent. This is different from Hindi where we observed speaker variation rather to be a matter of presence or absence of a cue to prominence. In Hindi, there is no combination of cues that together signal prominence as in German or English (or presumably as in other stress-based cue languages), but rather only one cue, i.e. f0 expressed as post-focal compression. The Hindi perception data then suggest that if speakers do not produce the prosodic cue to express focus, listeners get no prosodic cue at all to retrieve the intended focus structure. Hence, compared to other languages, speaker variation in Hindi rather seems to be concerned with whether speakers identify a focus structure and use the relevant cue to express it or not. We may thus conclude that those speakers who did not produce PFC may not have paid attention to the intended focus structure, although we cannot tell for certain. As a suggestion from a reviewer, a possible alternative explanation could be that some speakers simply do not use any prosodic means to signal focus, because focus structure can also be signaled by word order and from context. It is also well known that languages vary in whether or not certain contrasts are signaled morpho-syntactically (e.g., the subject-object asymmetry of focus marking, Fiedler et al. 2010), and perhaps this is just another case of such variation involving prosody.

### 4.3 The role of cues of intonation in sentence processing

In languages like English, the phonological parts of an intonation contour consist of pitch accents and boundary tones, which convey particular meanings (e.g., Ladd 2008). Phonetic cues signaling a boundary tone are first of all a significant f0 target accompanied with pauses, phrase-final lengthening, and/or segmental cues such as laryngealization (e.g., Martin 1970, Lehiste 1972, Klatt 1975, Wightman et al. 1992). These cues have been shown to matter in speech processing (Lehiste 1973, Warren et al. 1995, Schafer et al. 1996, Carlson et al. 2001, Watson et al. 2008). For instance, a longer duration of the words *greeted the girl* in (5) indicates a phrase break after the verb *greeted* leading to the interpretation that the girl is smiling and not the hostess. Lehiste relates the longer duration measured over these words to the presence of a phrase break between the verb and the complement. This phrase break creates the reading that the girl smiled. Note that this interpretation should also come about if measuring the duration on the verb alone, which Lehiste did not do.

(5) The hostess greeted the girl with a smile. (Lehiste 1973, p. 108)

Pitch accents are expressed in terms of changes of f0. The relative clause in (6) modifies either *the propeller* or *the plane*. The interpretation of the whole utterance depends on the placement of pitch

accent. Hence, the relative clause is attached to that constituent that carries the pitch accent (Schafer et al. 1996). This phenomenon has been called the *focus attraction hypothesis* put forward by Schafer et al. (1996). It states that a phrase or complement rather attaches to a focused (accented) phrase than to an unaccented one. The site of attachment then matters for the interpretation of the sentence. A pitch accent indicates the focus of the matrix clause. The constituent carrying the pitch accent will be the likely one that a following constituent such as a relative clause in (6) would attach to. Hence, the mechanic is examining the propeller if *the propeller* carries an accent, else the plane if *the plane* carries the accent.

(6) The sun sparkled on the propeller of the plane that the mechanic was so carefully examining. (Schafer et al. 1996, p. 142)

Studies on the role of prosody in language processing have investigated the two basic functions of intonation, grouping and highlighting of information, which are expressed by pitch accents and boundary tones, respectively. Given that these phonological categories play a significant role in sentence disambiguation in intonation languages, the question remains if similar categories exist in typologically different languages. Given that Hindi belongs to the group of phrase languages (Féry 2017) and thus does not have pitch accents, the fact that focus is prosodically realized by variations in the pitch register points to the fact the pitch register is such a phonological cue. Based on the results of the present study that PFC is functionally used to identify the focus structure of a sentence we propose that pitch register indeed represents a phonological category. If register compression would be a mere epiphenomenon of focus we would not expect it to be functionally used. In Akan question intonation, Genzel & Kügler (2018) identified pitch register as a phonetic epiphenomenon besides sentence-final cues that signal yes/no questions. They showed that raised pitch register played no role in the identification of a yes/no question in perception. Hence, variation in pitch register in Akan has no function and can be interpreted as a phonetic effect.

Whether pitch register plays a functional role in other, closely related South Asian languages has yet to be shown. Given the areal similarity in prosody between these languages proposed by Khan (2016), we might expect some South Asian languages to show a similar function of pitch register. However, as Xu (2011) discussed, the presence or absence of PFC is not a matter of typological relatedness nor a matter of areal relatedness. In addition, the fact that we propose PFC to be a prosodic category in Hindi does not imply that PFC must be a prosodic category universally. In particular, in languages like German or English there are more prosodic cues to express focus than just PFC; in these languages PFC may not carry the same function it does in Hindi.

## 5 Conclusion

In line with previous studies on Hindi intonation, we have shown that focus is prosodically expressed by means of post-focal compression (PFC), and that no prosodic cues are realized on the focused constituent itself. This was shown on production data of 30 female native speakers of Hindi reading contrastive ellipsis sentences. The matrix clause of a contrastive ellipsis contained two objects, an indirect object and a direct object. We examined whether changes in the placement of focus on the objects were accompanied by prosodic cues, namely PFC. The production data revealed that if the indirect object is in contrast, PFC was realized on the following direct object. The opposite case, contrast of the direct object did not yield any realization of PFC, which was unexpected.

The production data served as basis for a auditory sentence completion study in order to examine whether the prosodic cue PFC guides listeners to identify the focus structure of a sentence. The results of the sentence completion study showed clearly that if PFC was present in the stimuli, listeners were able to detect the focus structure. They could successfully complete the sentence with the correctly contrasted object. The conclusion is that PFC carries functional load in Hindi in that speakers can rely on PFC to identify focus placement, even when other prosodic cues are absent. PFC helps to disambiguate the local ambiguity in contrastive ellipsis structures. In languages that use different strategies to express focus prosodically, pitch accents and/or boundary tones serve to cue syntactic information. We argued that the prosodic cue of PFC showed similar behavior as the intonational categories pitch accents and boundary tones, and that pitch register therefore is a phonological category in Hindi.

## Acknowledgements

This work was supported by the German Research Foundation DFG with projects KU 2323/2-1 and KU 2323/4-1. Special thanks to the Co-PI's of the project 'Prosody in Parsing' at Potsdam University, Caroline Féry and Shravan Vasishth, and to the project members Gerrit Kentner, Anja Gollrad and Stella Gryllia as well as our student assistant, at that time, Dinah Baer-Henney. Thanks also to Susanne Genzel for discussion of different issues of Hindi prosody. Thank you to two anonymous reviewers and the editor for very helpful comments on an earlier version of the paper.

## References

Bali, Kalika, Monojit Choudhury, Diptesh Chatterjee, Arjit Maheswari, and Sankalan Prasad. 2009. Correlates between performance, prosodic and phrase structures in Bangla and Hindi: Insights from a psycholinguistic experiment. In *Proceedings of ICON-09,* Hyderabad, India, pages 101-110.

Bates, Douglas, Martin Mächler, Benjamin Bolker, and Steven Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Baumann, Stefan, Johannes Becker, Martine Grice, and Doris Mücke. 2007. Tonal and articulatory marking of focus in German. In J. Trouvain and W. J. Barry, eds., *Proceedings of the 16th International Congress of Phonetic Sciences,* pages 1029–1032. Saarbrücken, Germany.

Boersma, Paul and David Weenink. 2018. Praat: doing phonetics by computer [computer program]. Retrieved from http://www.praat.org/

Bolinger, Dwight. 1989. Intonation and its uses: Melody in grammar and discourse. Stanford, Calif.: Stanford University Press. ISBN 0804715351.

Botinis, Antonis, Marios Fourakis, and Barbara Gawronska. 1999. Focus identification in English, Greek and Swedish. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey, eds., *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS),* pages 1557–1560. San Francisco: University of California. ISBN 978-1563968990.

Breen, Mara, Evelina Fedorenko, Michael Wagner, and Edward Gibson. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes* 25(7-9):1044–1098.

Carlson, Katy, Charles Clifton, JR., and Lyn Frazier. 2001. Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45:58–81.

Clements, George Nick. 1979. The description of terraced-level tone languages. *Language* 55(3):536–558.

Cruttenden, Alan. 2006. The de-accenting of given information: A cognitive universal. In G. Bernini and M. L. Schwartz, eds., *Pragmatic organization of discourse in the languages of Europe*, pages 311–355. Berlin, Mouton de Gruyter.

del Mar Vanrell, Maria. 2011. The phonological relevance of tonal scaling in the intonational grammar of Catalan. Phd thesis, Universitat de Barcelona, Barcelona.

Drubig, Hans Bernhard. 1994. Island constraints and the syntactic nature of focus and association with focus. In *Arbeitspapiere des Sonderforschungsbereichs 340: Sprachtheoretische Grundlagen der Computerlinguistik, Vol. 51*. Tübingen/Stuttgart.

Fanselow, Gisbert. 2016. Syntactic and prosodic reflexes of information structure in Germanic. In C. Féry and S. Ishihara, eds., *Oxford Handbook of Information Structure*, pages 621–641. Oxford: Oxford University Press. ISBN 9780199642670.

Féry, Caroline. 2010. Indian languages as intonational 'phrase languages'. In I. Hasnain and S. Chaudhury, eds., *Problematizing language studies. Festschrift for Rama Agnihotri*, pages 288–312. Delhi: Aakar Books. ISBN 978-9350020791.

Féry, Caroline. 2017. *Intonation and Prosodic Structure.* Cambridge: Cambridge University Press. ISBN 978-1-107-00806-9.

Féry, Caroline and Frank Kügler. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36(4):680–703.

Féry, Caroline, Pramod Pandey, and Gerrit Kentner. 2016. The prosody of focus and givenness in Hindi and Indian English. *Studies in Language* 40(2):302–339.

Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz, and Malte Zimmermann. 2010. Subject focus in West African languages. In M. Zimmermann and C. Féry, eds., *Information structure*, pages 234–257. Oxford: Oxford University Press. ISBN 978-0-19-957095-9.

Fletcher, Janet. 2010. The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, and F. E. Gibbon, eds., *The Handbook of Phonetic Sciences*, pages 523–602. Oxford: Wiley-Blackwell. ISBN 978-1-4051-4590-9.

Genzel, Susanne and Frank Kügler. 2010. The prosodic expression of contrast in Hindi. In *Proceedings of the Fifth International Conference on Speech Prosody 2010*, pages 100143:1–4. Chicago, USA.
Genzel, Susanne and Frank Kügler. 2018. Production and perception of question prosody in Akan. *Journal of the International Phonetic Association* pages 1–32. doi: 10.1017/S0025100318000191
Gussenhoven, Carlos. 2002. Intonation and interpretation: Phonetics and phonology. In B. Bel and I. Marlien, eds., *Proceedings of the Speech Prosody 2002 Conference*, pages 47–57. Aix-en-Provence: Laboratoire Parole et Langage.
Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press. ISBN 9780521012003.
Hamann, Silke. 2009. Variation in the perception of an L2 contrast: A combined phonetic and phonological account. In F. Kügler, C. Féry, and R. van de Vijver, eds., *Variation and gradience in phonetics and phonology, Phonology and phonetics*, pages 71–98. Berlin: Mouton De Gruyter. ISBN 978-3-11-021931-9.
Harnsberger, James D. 1996. Towards an intonational phonology of Hindi. In *Fifth Conference on Laboratory Phonology*. Chicago: Northwestern University.
Harnsberger, James D. 1999. The role of metrical structure in Hindi intonation. Paper at South Asian Language Analysis Roundtable.
Harnsberger, James D. and Jasmeet Judge. 1996. Pitch range and focus in Hindi. *The Journal of the Acoustical Society of America* 99:2493.
Hayes, Bruce and Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9(1):47–96.
Hume, Elizabeth and Keith Johnson. 2003. The impact of partial phonological contrast on speech perception. In M.-J. Solé, D. Recasens i Vives, and J. Romero, eds., *Proceedings of the 15th International Congress of Phonetic Sciences*, pages 2385–2388. Barcelona: Univ. Autònoma de Barcelona. ISBN 9781876346492.
Jeon, Hae-Sung and Francis Nolan. 2017. Prosodic marking of narrow focus in Seoul Korean. *Laboratory Phonology* 8(1):390.
Jun, Sun-Ah and Hee-Sin Kim. 2007. VP focus and narrow focus in Korean. In J. Trouvain and W. J. Barry, eds., *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 1277–1280. Saarbrücken, Germany, 4-10 August 2007.
Keane, Elinor. 2014. The intonational phonology of Tamil. In S.-A. Jun, ed., *Prosodic typology II*, pages 118–153. Oxford: Oxford University Press. ISBN 978-0-19-956730-0.
Khan, Sameer ud Dowla. 2008. Intonational Phonology and Focus Prosody of Bengali. Phd thesis, UCLA, Los Angeles.
Khan, Sameer ud Dowla. 2016. The intonation of South Asian languages: towards a comparative analysis. In M. Menon and S. Syed, eds., *Proceedings of FASAL 6,* pages 23–36.
Kidwai, Ayesha. 2000. *XP-adjunction in Universal Grammar: Scrambling and binding in Hindi-Urdu.* Oxford: Oxford University Press. ISBN 0-19-513251-3.
Klatt, Dennis H. 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3(3):129–140.
Krahmer, Emiel and Marc Swerts. 2007. Perceiving focus. In C. Lee, M. K. Gordon, and D. Büring, eds., *Topic and Focus: Cross-linguistic perspectives on Meaning and Intonation*, pages 121–138. Dordrecht: Springer. ISBN 978-1-4020-4795-4.
Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3):243–276.
Kügler, Frank and Sasha Calhoun. 2020. Prosodic encoding of information structure: A typological perspective. In C. Gussenhoven and A. Chen, eds., *The Oxford Handbook of Language Prosody*, Chapter 31. Oxford: Oxford University Press.
Kügler, Frank and Caroline Féry. 2017. Postfocal downstep in German. *Language and Speech* 60(2):260–288.
Kügler, Frank and Susanne Genzel. 2012. On the prosodic expression of pragmatic prominence: The case of pitch register lowering in Akan. *Language and Speech* 55(3):331–359.
Kügler, Frank and Anja Gollrad. 2015. Production and perception of contrast: The case of the rise-fall contour in German. *Frontiers in psychology* 6(1254):1–18.
Ladd, D. Robert. 1980. *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.
Ladd, D. Robert. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press, 2nd edn.
Ladd, D. Robert and Rachel Morton. 1997. The perception of intonational emphasis: continuous or categorical? *Journal of Phonetics* 25(3):313–342.
Lee, Yong-cheol and Yi Xu. 2010. Phonetic realization of contrastive focus in Korean. In *Proceedings of the Fifth International Conference on Speech Prosody 2010*, pages 100033:1–4. Chicago.
Lehiste, Ilse. 1972. The timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America* 51(6,2):2018–2024.
Lehiste, Ilse. 1973. Phonetic disambiguation of syntactic ambiguity. *Glossa* 7(2):107–122.

Liu, Fang and Yi Xu. 2005. Parallel encoding of focus and interrogative meaning in mandarin intonation. *Phonetica* 62:70–87.

Luchkina, Tatiana, Vandana Puri, Preethi Jyothi, and Jennifer S. Cole. 2015. Prosodic and structural corelates of perceived prominence in Russian and Hindi. In The Scottish Consortium for ICPhS 2015, ed., *Proceedings of the 18th International Congress of Phonetic Sciences*, paper 793. Glasgow, UK: University of Glasgow. ISBN 978-0-85261-941-4.

Martin, James G. 1970. On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior* 9(1):75–78.

Moore, Robert Ripley. 1965. A Study of Hindi intonation. Phd thesis, University of Michigan, Ann Arbor.

Nair, Rami. 2001. Acoustic correlates of lexical stress in Hindi. In A. Abbi, R. S. Gupta, and A. Kidwai, eds., *Linguistic Structure and Language Dynamics in South Asia*, pages 123–143. Delhi: Motilal Banarsidas Publishers Pvt. Ltd.

Ohala, Manjari. 1991. Phonological areal features of some Indo-Aryan languages. *Language Sciences* 13(2):107–124.

Ohala, Manjari. 1994. Hindi. *Journal of the International Phonetic Association* 24(1):35–38.

Patil, Umesh, Gerrit Kentner, Anja Gollrad, Frank Kügler, Caroline Féry, and Shravan Vasishth. 2008. Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics* 1(1):53–67.

Peperkamp, Sharon and Emmanuel Dupoux. 2002. A typological study of stress 'deafness'. In C. Gussenhoven and N. Warner, eds., *Laboratory Phonology* 7, pages 203–240. Berlin: Mouton. ISBN 3110170868.

Pierrehumbert, Janet B. 1980. The phonology and phonetics of English intonation. Phd thesis, MIT, Cambridge, Massachusetts.

Pierrehumbert, Janet B. and Mary E. Beckman. 1988. Japanese Tone Structure. Cambridge: MIT Press.

Repp, Sophie. 2010. Defining 'contrast' as an information-structural notion in grammar. *Lingua* 120(6):1333–1345.

Schafer, Amy, Juli Carter, Charles Clifton, JR, and Lyn Frazier. 1996. Focus in relative clause construal. *Language and Cognitive Processes* 11(1/2):135–163.

Turk, Alice. 2011. The temporal implementation of prosodic structure. In A. C. Cohn, C. Fougeron, and M. K. Huffman, eds., *The Oxford Handbook of Laboratory Phonology*, pages 242–253. Oxford: Oxford University Press. ISBN 978-0-19-957503-9.

Turk, Alice, Satsuki Nakai, and Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: A practical guide. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, and J. Schließer, eds., *Methods in empirical prosody research*, pages 1–27. Berlin: De Gruyter. ISBN 978-3-11-018856-1.

Vainio, Martti and Juhani Järvikivi. 2006. Tonal features, intensity, and word order in the perception of prominence. *Journal of Phonetics* 34(3):319–342.

Wang, Bei, Chenxia Li, Qian Wu, Xiaxia Zhang, and Yi Xu. 2012. Production and perception of focus in Huihui language: Compared with Beijing Mandarin. In *Proceedings of the 13th International Conference on Speech Communication and Technology*, pages 663–666. Portland.

Warren, Paul, Esther Grabe, and Francis Nolan. 1995. Prosody, phonology, and parsing in close ambiguities. *Language and Cognitive Processes* 10(5):457–486.

Watson, Duane G., Michael K. Tanenhaus, and Christine A. Gunlogson. 2008. Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science* 32(7):1232–1244.

Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf, and Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America* 91(3):1707–1717.

Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27(1):55–105.

Xu, Yi. 2011. Post-focus compression: Cross-linguistic distribution and historical origin. In E. Zee and W.-S. Lee, eds., *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, pages 152–155.

Xu, Yi, Szu-Wei Chen, and Bei Wang. 2012. Prosodic focus with and without post-focus compression: A typological divide within the same language family? *The Linguistic Review* 29:131–147.

Xu, Yi and Ching X. Xu. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33(2):159–197.

## Appendix

*Speech materials*

E03-C01-01-30-079-

| राहुल ने | मां को | दवाई | दी | न | कि | घड़ी। |
|---|---|---|---|---|---|---|
| Rāhul=ne | mā=ko | davāī | dī | nā | ki | ghaṛī |
| Rahul=ERG | mother=DAT | medicine | gave | NEG | C | watch |

'Rahul gave the medicine to mother and not the watch.'

E03-C02-01-30-107-

| राहुल ने | मां को | दवाई | दी | न | कि | नानी को। |
|---|---|---|---|---|---|---|
| Rāhul=ne | mā=ko | davāī | dī | nā | ki | nānī=ko |
| Rahul=ERG | mother=DAT | medicine | gave | NEG | C | grandmother=DAT |

'Rahul gave the medicine to mother and not to grandmother.'

E03-C01-02-30-153-

| कपील | दादाजी से | बैग | लाया | न | कि | पेन। |
|---|---|---|---|---|---|---|
| Kapīl | dādājī=se | baig | lāyā | nā | ki | pen |
| Kapil | grandfather=ABL | bag | brought | NEG | C | pen |

'Kapil brought the bag from grandfather and not the pen'

E03-C02-02-30-214-

| कपील | दादाजी से | बैग | लाया | न | कि | नीरव से। |
|---|---|---|---|---|---|---|
| Kapīl | dādājī=se | baig | lāyā | nā | ki | Nīrav=se |
| Kapil | grandfather=ABL | bag | brought | NEG | C | Nirav=ABL |

'Kapil brought the bag from grandfather and not from Nirav.'

E03-C01-03-30-072-

| मुरली ने | चिराग से | मोबाइल | लिया | न | कि | हेलमेट। |
|---|---|---|---|---|---|---|
| Muralī=ne | Cirāg=se | mobāil | liyā | nā | ki | helmeṭ |
| Murali=ERG | Chirag=ABL | mobile | took | NEG | C | helmet |

'Murali took the mobile from Chirag and not the helmet.'

E03-C02-03-30-034-

| मुरली ने | चिराग से | मोबाइल | लिया | न | कि | अभय से। |
|---|---|---|---|---|---|---|
| Muralī=ne | Cirāg=se | mobāil | liyā | nā | ki | Abhay=se |
| Murali=ERG | Chirag=ABL | mobile | took | NEG | C | Abhay=ABL |

'Murali took the mobile from Chirag and not from Abhay.'

E03-C01-04-30-197-

| मामी ने | रवी को | कहानी | सुनाई | न | कि | शायरी। |
|---|---|---|---|---|---|---|
| māmī=ne | Ravī=ko | kahānī | sunāī | nā | ki | šāyarī |
| auntie=ERG | Ravi=DAT | story | told | NEG | C | poem |

'Auntie told a story to Ravi and not a poem.'

E03-C02-04-30-076-

| मामी ने | रवी को | कहानी | सुनाई | न | कि | रमन को। |
|---|---|---|---|---|---|---|
| māmī=ne | Ravī=ko | kahānī | sunāī | nā | ki | Raman=ko |
| auntie=ERG | Ravi=DAT | story | told | NEG | C | Raman=DAT |

'Auntie told a story to Ravi and not to Raman.'

E03-C01-05-30-132-

| परीक्षक ने | अमोल से | नाम | पुछा | न | कि | रोल नंबर। |
|---|---|---|---|---|---|---|
| parīkšak=ne | Amol=se | nām | puchā | nā | ki | rol nambar |
| examiner=ERG | Amol=ABL | name | asked | NEG | C | roll number |

'The examiner asked Amol his name and not the roll number.'

E03-C02-05-30-012-

| परीक्षक ने | अमोल से | नाम | पुछा | न | कि | सोनाली से। |
|---|---|---|---|---|---|---|
| parīkšak=ne | Amol=se | nām | puchā | nā | ki | Sonālī=se |
| examiner=ERG | Amol=ABL | name | asked | NEG | C | Sonali=ABL |

'The examiner asked Amol his name and not Sonali.'

E03-C01-06-30-135-

| गाइड ने | पर्यटक को | ताजमहल | दिखाया | न | कि | कुतुब मीनार। |
|---|---|---|---|---|---|---|
| gāiḍ=ne | paryaṭak=ko | Tājmêhêl | dikhāyā | nā | ki | Qutub Mīnār |
| guide=ERG | tourist=DAT | Taj Mahal | showed | NEG | C | Qutub Minar |

'The Guide showed the Taj Mahal to the tourist and not Qutub Minar.'

E03-C02-06-30-092-

| गाइड ने | पर्यटक को | ताजमहल | दिखाया | न | कि | भिखारी को। |
|---|---|---|---|---|---|---|
| gāiḍ=ne | paryaṭak=ko | Tājmêhêl | dikhāyā | nā | ki | bhikhārī=ko |
| guide=ERG | tourist=DAT | Taj Mahal | showed | NEG | C | beggar=DAT |

'The Guide showed the Taj Mahal to the tourist and not to the beggar.'