Interaction between prosody and focus types: Evidence from Bangla and Hindi*

Arunima Choudhury & Elsi Kaiser

University of Southern California

1. Introduction

This paper reports a series of psycholinguistic experiments investigating the prosodic encoding of information structure in Bangla and Hindi. We tested whether speakers of these languages encode the information-structural distinction between new-information focus and corrective focus prosodically, and if so, what prosodic dimensions (e.g. fundamental frequency (*F0*, duration, intensity) are used. We also tested to what extent listeners are able to perceive prosodic cues to information structure.

The term 'information structure' is used to refer to parts of the sentence that can be classified as given and topical information or as new/focused information (e.g. Chafe 1976, Vallduví 1990, Lambrecht 1994, Féry & Krifka 2008). The category of focus is often classified into two broad types: (a) New Information Focus, and (b) Contrastive Focus. New-information focus is primarily associated with new, non-presupposed information, and a commonly used diagnostic for new-information focus is the use of wh-questions, as in (1a), from Zimmermann & Onea (2011). The focus constituent (marked with square brackets and the subscript F) is the part of the sentence that corresponds to the answer to the wh-question (e.g. Jackendoff 1972, Gussenhoven 2008, Kanerva 1990).

(1) a. A: What color did Peter paint his bicycle? B: He painted it $[\text{blue}]_F$.

Following Zimmermann & Onea (2011), we define new-information focus as follows: A focused constituent (e.g. *blue* in 1a) expresses new information if the focused element introduces new information to the Common Ground (the mutually shared knowledge between speaker and addressee), when alternatives to the focused element had not been explicitly mentioned in the prior discourse. Contrastive focus, on the other hand, occurs when one or more of the alternatives to the focused constituent have been mentioned. For example, in (1b), one speaker claims that Peter painted his bicycle red.

* We are grateful to the audience at FASAL 5 (Yale, 4/2015) for thoughtful comments and suggestions. We also thank Prof. Andrew Simpson for his valuable feedback and support at various stages of this project.

The other speaker corrects this and says that out of the possible colors that Peter could have painted this bicycle, he painted it blue. The contrastively-focused element *blue* competes with another, explicitly-mentioned member in the set of possible bicycle colors (*red*) for introduction in the Common Ground (see Zimmermann & Onea 2011).

(1) a. A: Peter painted his bicycle red. B: No, he painted it $[\text{blue}]_F$.

The type of focus in (1b) is more specifically known as corrective focus, which is sometimes regarded as a subtype of the more general class of contrastive focus, and sometimes as distinct from contrastive focus (for discussion, see e.g. Zimmermann & Onea 2011, Dik 1997, Gussenhoven 2007). In this paper, we investigate the prosodic encoding of new-information focus and corrective focus, since these two focus types are intuitively distinct and reliably elicited by wh-questions and correction contexts.

Despite examples like (1a,b), researchers disagree whether contrastive focus constitutes an information-structural category of its own, distinct from new-information focus. Some researchers – largely with a syntactic focus – argue that contrastive focus is a focus type independent of new-information focus (e.g. Chafe 1967, Halliday 1967, Rochemont 1986, Kiss 1998, Valduví & Vilkuna 1998, Molnár 2002). On the other hand, others like Bolinger (1961), Rooth (1992), Krifka (1993) and Schwarzschild (1999) – largely with a semantic focus – have claimed that contrastive focus is not fundamentally different from information focus, as far as its semantics are concerned.

Work on prosody has identified different intonational properties for new-information focus and contrastive/corrective focus. Zubizarreta (1998) believes new-information focus and contrastive focus to be information-structurally distinct and to have different intonational realizations in Romance languages like Italian and Spanish. Pierrehumbert (1980) and others working within auto-segmental metrical phonology posit distinct pitch accents in English for new-information focus (H*) and contrastive focus (L + H*).

Psycholinguistic work suggests that speakers encode different focus types with different prosodic cues and that listeners are sensitive to this, but the mapping between focus types and prosody may not be straightforward. For example, a comprehension study by Watson et al. (2008) on English showed that L+H* accents are interpreted as having a contrastive interpretation, whereas H* accents can mark either contrastive referents or new-information referents. A series of production-and-perception experiments by Breen et al. (2010) found that speakers distinguish corrective and non-corrective focus reliably only when they are aware of the prosodic ambiguity present across different information structures. In that case, speakers produced contrastively focused elements with greater intensity, longer duration, and (perhaps surprisingly) lower mean and maximum *F0* than non-contrastively-focused elements.

On the perception side, Breen et al. observed that even when speakers' productions distinguished corrective and non-corrective focus, listeners did not successfully identify focus type. (Listeners' performance on focus types improved when an attributive phrase "I heard that" preceded the critical SVO sentence, which Breen et al. attribute to speakers prosodically marking "I" when the sentences were contrastive.) Recent production work by Katz & Selkirk (2011) found that contrastive focus and new-information focus are prosodically distinct, but Katz and Selkirk did not test perception.

In sum, within theoretical linguistics there is an on-going debate about whether we

should have a grammatical representation of contrastive focus that is distinct from that of new-information focus. Current psycholinguistic work exploring the distinctness (or lack thereof) of new-information focus and corrective focus has led to mixed results, especially when we look at both production and perception.

In the current study, we investigate the production and perception of new-information and corrective focus with native Bangla and Hindi speakers, to see how two languages which are typologically related to each other but which differ from English in some key respects, can contribute to this debate. As will become apparent, looking at two related languages turns out to be highly informative with respect to the question of what prosodic/acoustic dimensions languages use for encoding focus.

1.1 Background: Bangla and Hindi

Bangla and Hindi are a part of the modern-day Indic/Indo-Aryan branch of the Indo-European family of languages. Their canonical word order is SOV, but word order is relatively flexible. Both Bangla and Hindi have a 'default focus position' which immediately precedes the verb (e.g. Choudhury 2010 on Bangla, Kidwai 2000 on Hindi), and this position has been suggested to be the default position for new information focus in both languages (S $[O]_{new}$ V). Furthermore, native speaker judgments indicate that contrastively-focused elements can also occur in this position (S $[O]_{corr}$ V). Furthermore, in canonical SOV word order, the subject can be in new-information focus ($[S]_{new}$ O V) or contrastively focused ($[S]_{corr}$ O V).[1] Hence, the canonical SOV word order in both Bangla and Hindi can have the following interpretations:

(i) S $[\mathbf{O}]_{\mathbf{new}}$V - New-information focus on **object**; subject is unfocused
(ii) S $[\mathbf{O}]_{\mathbf{corr}}$V - Corrective focus on **object;** subject is unfocused
(iii) $[\mathbf{S}]_{\mathbf{new}}$ O V - New-information focus on **subject**; object is unfocused
(iv) $[\mathbf{S}]_{\mathbf{corr}}$ O V - Corrective focus on **subject**; object is unfocused

Now, let us turn to what is known about the prosodic properties of Hindi and Bangla. Previous studies (Hayes & Lahiri 1991, Féry 2010, Patil et al 2008, Moore 1965, Harnsberger 1994, Khan 2007) have shown that both languages have an L^*H_P pitch accent on the focused constituent (based on the autosegmental metrical phonology framework, see Pierrehumbert 1980). In both languages, each content word forms its independent phonological phrase and each phonological phrase receives a low tone and a high phrase boundary associated with the right edge of the prosodic word. Both Bangla and Hindi clauses also exhibit a downstep intonation pattern. However, not much is known about the quantitative values of fundamental frequency, duration and intensity of the focused constituents in Hindi and Bangla. Also, we do not know much about the prosodic differences between focus types in Hindi and Bangla, and whether Bangla and

[1] Noncanonical OSV order does not seem to allow the sentence-initial argument to be in new-information focus (* $[O]_{new}$ S V) although contrastive focus seems to be possible ($[O]_{corr}$ S V). See Choudhury (2015) for further discussion and experiments regarding OSV order in Hindi and Bangla. Here, we focus on SOV order. It is also worth noting that Bangla and Hindi have focus particles (e.g. Bhatt 1994, Sharma 2003 on Hindi, Choudhury 2010 on Bangla), but they are only used in certain contexts, and only with corrective focus. We do not investigate them here. Our experimental stimuli do not require focus particles.

Hindi encode the focus types distinctly using prosody. Our work aims to contribute to these questions.

1.2 Research Questions

The first research question that our experiments address is whether *speakers* of Bangla and Hindi encode the distinction between new-information focus and corrective focus prosodically, and if so, what prosodic dimensions (e.g. fundamental frequency *F0*, duration, intensity) are used, and are they the same in both languages? The second research question is whether *listeners* of these two languages are able to perceive the prosodic cues used to signal focus types. Are listeners able to differentiate between new information focus and corrective focus, when word order provides no cues?

These two questions will provide new data on the encoding and perception of information structure in Bangla and Hindi, and they also relate to the debate regarding the relationship between new-information focus and contrastive/corrective focus. If speakers of Hindi and Bangla distinguish prosodically between the two focus types and listeners are sensitive to these prosodic cues, this will provide us with further crosslinguistic evidence that focus types are encoded as separate categories.

Conducting a parallel investigation of two closely-related languages will allow us to gain insights into how variable languages are in terms of the specific prosodic dimensions that they use to encode information structure. Several studies on Chinese languages suggest that even typologically closely related language may employ different prosodic cues to encode focus (Xu 1999 for Mandarin, Wu & Xu 2010 for Cantonese, Chen et al 2009 for Beijing Mandarin, Taiwanese Mandarin and Taiwanese). Thus, we should not assume an overly simplistic view regarding the 'division of labor' between different prosodic dimensions such as *F0*, duration and intensity.

We conducted three sets of experiments. All three experiments consist of a production study followed by a perception study, where the production study recordings were used as stimuli for the perception study. Experiment 1 and Experiment 2 were conducted with native Bangla speakers, and Experiment 3 was conducted with native Hindi speakers. Experiment 2 (Bangla) and Experiment 3 (Hindi) use sentences with Adverb-Subject-Object-Verb word order. This was done based on the results of Experiment 1 (Bangla), which used Subject-Object-Verb sentences. In Experiments 2 (Bangla) and 3 (Hindi), the subject is no longer at the sentence-initial position, but the object is still at the default focus position. This allows us to see whether the results of Experiment 1 could be due to specific prosodic properties of sentence-initial elements. For all three studies, we first describe the design and methods used in the production phase, followed by the perception study and the results of the perception study, and finally the acoustic analyses of the stimuli.

2. Experiment 1: SOV word order in Bangla

2.1 Production phase

In the production phase, we elicited Bangla SOV sentences with new-information or corrective focus on the subject or the object. The data from this study was analyzed and

used as the stimuli for the perception study phase of Experiment 1. Five adult native Bangla speakers (3 female, 2 male) participated in the production phase. All were originally from Kolkata, India, and had been in the U.S. for at most 3 years.

Participants produced SOV sentences in contexts designed to elicit either new-information focus or corrective focus. We used wh-questions to elicit new-information focus (following Gussenhoven 2008, Kanerva 1989 and others), and yes/no questions to elicit corrective focus (following Breen et al 2010), as in (2a)-(2d). We manipulated focus type (new-information focus vs. contrastive focus) and the grammatical role of the focused element (subject vs. object). This resulted in four conditions: Sub-New, Obj-New, Sub-Corr, Obj-Corr (Table 1). The word order of the answer is the same in all conditions (SOV). The verbs were in the simple past tense, and nouns were 2.15 syllables long on average. All subjects were [+human], and all objects were inanimate.

Table 1. Experiment 1: Production study design. (Underlining indicates focus)

	Subject	Object
New-information Focus	S O V	S O V
Contrastive Focus	S O V	S O V

(2a) Subject wh-question to elicit new-information focus (Sub-New information)
Q: ke gari kinlo? A: **[baba]**$_{\textbf{newinfo}}$ gari kinlo
Who car bought father car bought
Who bought a car? 'Father bought a car'

(2b) Object wh-question to elicit new-information focus (Obj-New information)
Q: baba ki kinlo? A: baba **[gari]**$_{\textbf{newinfo}}$ kinlo
Father what bought father car bought
'What did father buy?' 'Father bought a car'

(2c) Subject yes/no question to elicit corrective focus (Sub-Corrective)
Q: protibeshi gari kinlo ki? A: **[baba]**$_{\textbf{Corr}}$ gari kinlo
Neighbor car bought Q father car bought
Did neighbor buy a car? 'Father bought a car'

(2d) Object yes/no question to elicit corrective focus (Obj-Corrective)
Q: baba kompyutar kinlo ki? A: baba **[gari]**$_{\textbf{Corr}}$ kinlo
Father computer bought Q father car bought
Did father buy a computer? 'Father bought a car'

Participants saw a question-answer pair on the screen, and said the answer out loud. They were encouraged to speak naturally, as if in a conversation.[2] The responses were recorded. This study was run on a PC using Paradigm software (Perception Research Systems). We created 20 target items and 32 filler items. Each speaker produced four targets per condition – for a total of 16 targets – as well as 16 fillers. We did two rounds of recordings for each speaker (all 32 trials in an initial round, and then all 32 trials again), to minimize disfluencies. The recordings from the second round were used for all

[2] When answering the yes/no questions, participants were instructed to avoid saying 'no' explicitly.

subsequent analyses. Each speaker produced 16 targets out of the full 20-item target set, and 16 fillers out of the full 32-item filler set; this kept the study at a reasonable duration. The full target and filler sets become relevant for the design of the perception study. In the next section, we present the perception study that used these sound files as stimuli. In Section 2.3, we present the acoustic analyses for these sound files.

2.2 Perception phase

The perception phase tests whether Bangla speakers perceive a distinction between new-information focus and corrective focus in the sentences generated during the production phase, and to see if the grammatical role of the focused element impacts perception. Twelve adult native speakers of Bangla participated (all originally from India, all had been in the U.S. for less than five years). None of these participants took part in the production study. The materials for this study were the audio files from the production phase. This study had 20 targets and 32 fillers, in a Latin-Square design. Every list had an equal distribution of target sentences spoken by all five speakers in all four conditions. The study was run using Paradigm software (Perception Research Systems).

On critical trials, participants saw a wh-question and a yes/no question on the screen (Figure 1). The grammatical role focused by the questions matched the grammatical role focused in the sound files: When participants saw a subject wh-question and a subject yes/no question, they heard a sound file elicited by (a) a subject wh-question or by (b) a subject yes/no question. Conversely, when they saw an object wh-question and an object yes/no question, they heard a sound file elicited by (a) an object wh-question or by (b) an object yes/no question. This allows us to test whether listeners can distinguish new-information focus and contrastive focus. Left and right locations of question types were counterbalanced. (The questions were shown in Bangla script.) Participants' task was to choose whether the sound file is an answer to the wh-question or the yes/no corrective question – i.e., which question is most appropriate for the sentence that they heard.

Figure 1. Experiment 1: Schematic example (Sub-New condition)

Ke gari kinlo?	**Protibeshi gari kinlo ki?**
Who bought a car?	Did neighbor buy a car?
(Subject wh-question)	*(Subject yes/no question)*

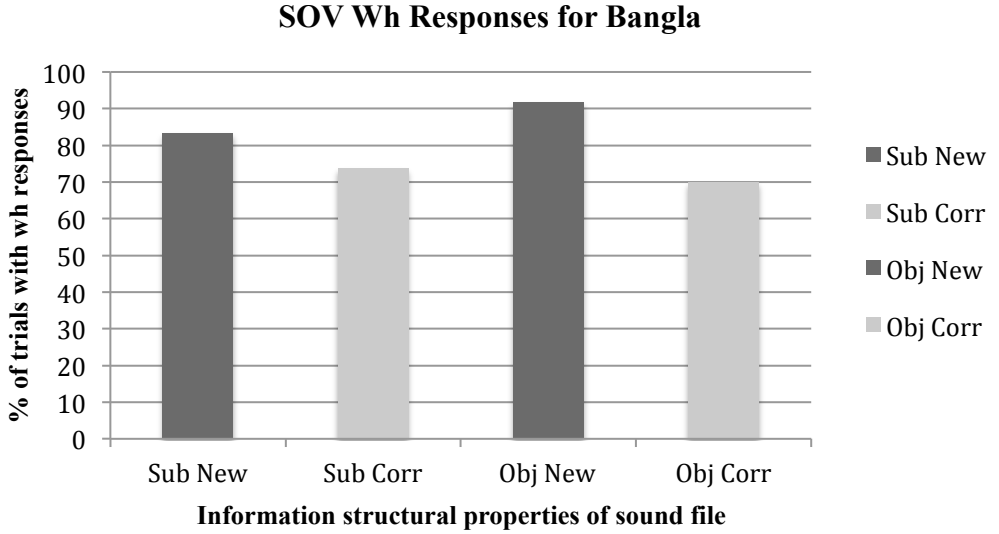
Presented auditorily (example of Sub-New condition)
[**baba**]$_{\text{NewInfo}}$ gaRi kinlo
father car bought

2.2.1 Results for perception phase

We present the results in terms of how often the participants chose the wh-question. Since wh-questions elicit new-information focus, this can be thought of as the *rate of new-information interpretations*. Figure 2 shows the rate of wh-question responses for each condition. Because participants only had two choices (wh-question, yes/no question), this means that whenever they did not choose a wh-question, they chose a yes-

no-question. The rate of yes/no-question responses can thus be inferred from Figure 2. (The wh-question is the 'right' answer for sounds elicited in new-information focus contexts, but the 'wrong' answer for sound files elicited in corrective focus contexts.)

Figure 2: Experiment 1: Rate of wh-question choices (indicating new-information focus)



As can be seen in Figure 2, the percentage of wh-question responses – i.e., the percentage of the time participants perceived the focus as being new-information focus – is well over 50% in all conditions. In other words, participants have an overall preference for a new-information interpretation (for discussion of why this might be the case, see Choudhury 2015). What is relevant for us here is that the preference for wh-questions is weakened for sound files elicited in a corrective context, when compared to sound files elicited in a new-information context. In object-focus conditions, the rate of wh-question choices is 92% for sound files elicited in a new-information focus context (i.e., preceded by a wh-question), but only 70% for sound files been elicited in a corrective focus context (i.e., preceded by a yes/no question). The same numerical pattern is observed in subject-focus conditions, but at a smaller magnitude: 83% of wh-question choices with new-information focus sound files; 73% wh-question choices with corrective focus sound files. Linear mixed-effects regressions confirm that the difference is significant in the case of object focus ($p<.05$) but not significant in the case of subject focus ($p=.14$).

We also tested for main effects and interactions here and in the other statistical analyses reported in this paper. However, due to space limitations, we mostly focus on the planned comparisons between the two different focus types on subjects and objects. Please see Choudhury (2015) for additional details and full results of all analyses.

Thus, our results indicate that Bangla listeners are able to distinguish corrective focus and new-information focus using only prosodic information when the focused element is the object, but not when it is the subject.

2.3. Acoustic analyses of production phase

We now turn to the prosodic acoustic properties of the focused constituents. We report on

two main prosodic cues, namely fundamental frequency (*F0* and duration). (We also looked at intensity, normalized over time, but found no significant effects of focus type, so we will not discuss the intensity data further. It seems that intensity does not play any role in encoding focus types in Bangla. Please see Choudhury 2015 for details.) The main purpose of the acoustic analyses is to determine which of these prosodic cues is facilitating the perception of the focus types, and to see if we can learn more about why listeners are not able to distinguish focus type differences at the subject position.

The production data was analyzed using Prosody Pro (Xu 2005-2012) and Praat (Boersma & Weenink 2009). We analyzed the time-normalized fundamental frequency and relative duration. For the time-normalized measures, we used Prosody Pro to divide each constituent into 10 equal time points/time segments. For the *F0* analyses reported throughout this paper, we focus on (the average of) the last 5 of these time points, because of the pitch accenting alignment of both Hindi and Bangla (Hayes & Lahiri 1991, Féry 2010, see Choudhury 2015 for additional discussion). We also analyzed the relative duration of the focused elements. To account for differences in speech rate, we followed Ito et al. (2006) and Kaland et al. (2011), and computed the relative duration of a constituent by dividing its absolute duration with the absolute duration of the sentence.

2.4 Results of production phase

Figure 3 shows the **fundamental frequency *(F0)*** patterns for the four conditions in Experiment 1, on SOV sentences in Bangla. As can be seen in the figure, it looks like the subject in all conditions ends with a high *F0* at the right edge of the word, regardless of whether or not it is focused. Objects, on the other hand, have high *F0*s at the right edge when they are focused and lower *F0*s when they are unfocused. Furthermore, we see that numerically, correctively focused objects (triangles) have a higher *F0* than new-information focused objects (diamonds). Statistical analyses (two-tailed paired t-tests) confirm that *F0*'s are significantly higher for corrective objects than new-information objects ($t(4)=4.019$; $P<0.05$), but corrective subjects vs. new-information subjects show no significant effect of focus type ($t(4)= -1.849$; $P=0.138$).

Figure 4 shows the **mean relative duration** of the focused constituent (subject or object) in all four conditions in Experiment 1. Correctively-focused constituents appear to be considerably longer than constituents in new-information focus. Statistical analyses (two-tailed paired t-tests) confirm that the mean durations are significantly higher for elements in corrective focus than for elements in new-information focus (subjects: $t(4)=-4.958$ $p<0.05$; objects: $t(4)=-7.217$ $p<0.01$). In sum, the duration analyses show that correctively-focused constituents are indeed significantly longer than constituents in new-information focus.

2.5 Discussion of Experiment 1 (Bangla SOV sentences)

Experiment 1 consisted of a production and perception study looking at SOV sentences in Bangla to test whether and with what prosodic dimension speakers encode the difference between new-information focus and corrective focus on subjects and objects. The acoustic analyses show that *F0* in Bangla encodes a difference between focus types on objects but not on subjects, whereas duration encodes a difference between focus types

on both subjects and objects. (Intensity does appear to reflect focus types.) However, in the perception phase of this study (which used the stimuli from the production phase as its stimuli), listeners were able to distinguish the two focus types reliably on objects but not on subjects. This is in line with the acoustic analyses for *F0*, but seems surprising in light of the duration data.

Before considering the implications of these findings further, it is important to mention a possible complication in Experiment 1: Could the asymmetrical *F0* results for subjects and objects be an artifact of the sentence-initial position of the subject, which is known to be associated with prosodic prominence. Could it be that this was masking potential effects of focus type on the subject? Experiment 2 addresses this issue, by testing SOV sentences that have an adverb in sentence-initial position.

Figure 3. Experiment 1: Time-normalized F0 contours of the SOV target sentences

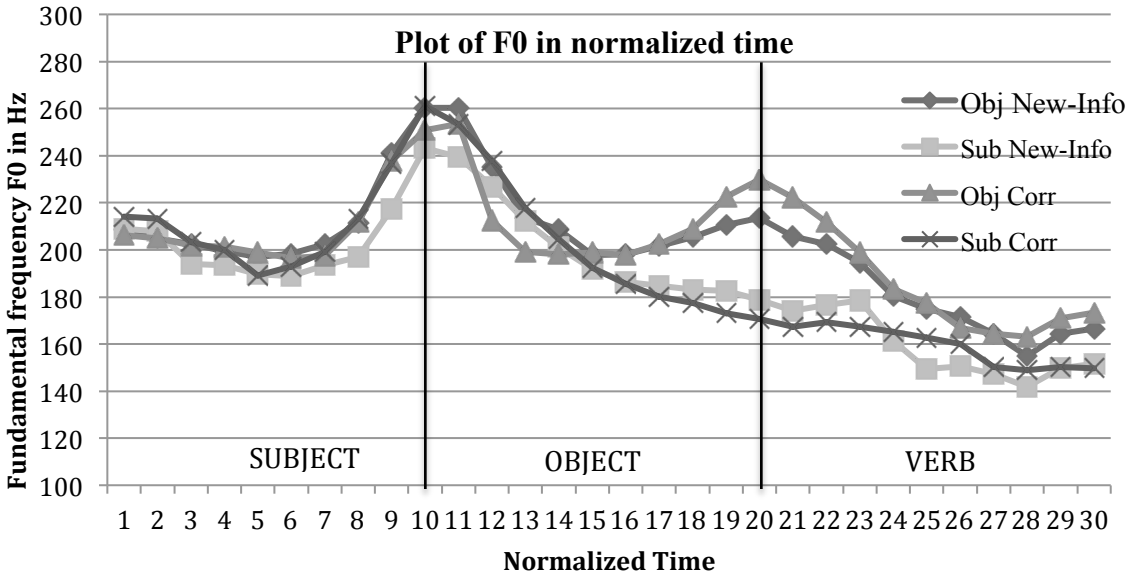
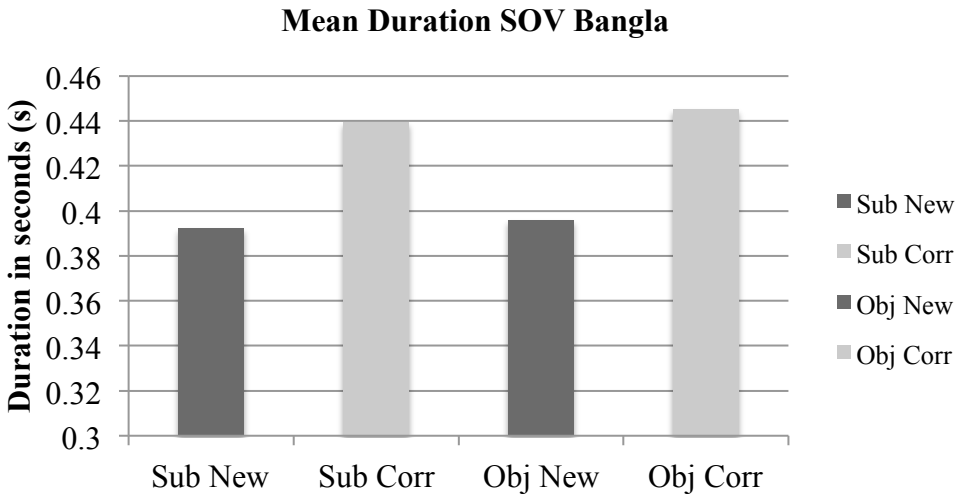


Figure 4: Experiment 1: Mean relative duration of the focused constituent (in seconds)



3. Experiment 2: Adv-SOV word order in Bangla

Experiment 2 tests if the absence of focus type effects on subjects exists even if the subject is not in sentence-initial position. Bangla is known to have a down-step pitch pattern (Hayes & Lahiri 1999, Féry 2010). This raises the question of whether the sentence-initial prominence could be masking potential differences between new-information focus and contrastive focus on the subject. Experiment 2 uses sentences where the subject is no longer in sentence-initial position: We added a sentence-initial adverb in front of the subject, as in example (3).

Similar to Experiment 1, this study is divided into two parts: a production study, followed by a perception study. The sentences generated during the production study were used as stimuli for the perception study.

(3a) Subject wh-question to elicit new-information focus (Sub-New)
Q: gotokal ke gari kinlo? A: gotokal [baba]$_{newinfo}$ gari kinlo
Yesterday who car bought yesterday father car bought
'Who bought a car yesterday? 'Father bought a car yesterday'

(3b) Object wh-question to elicit new-information focus (Obj-New)
Q: gotokal baba ki kinlo? A: gotokal baba [gari]$_{newinfo}$ kinlo
Yesterday father what bought yesterday father car bought
'What did father buy yesterday?' 'Father bought a car yesterday'

(3c) Subject yes/no question to elicit corrective focus (Sub-Corr)
Q: gotokal protibeshi gari kinlo ki? A: gotokal [baba]$_{Corr}$ gari kinlo
Yesterday neighbor car bought Q yesterday father car bought
Did neighbor buy a car yesterday? 'Father bought a car yesterday'

(3d) Object yes/no question to elicit corrective focus (Obj-Corr)
Q: gotokal baba kompyutar kinlo ki? A: gotokal baba [gari]$_{Corr}$ kinlo
Yesterday father computer bought Q yesterday father car bought
Did father buy a computer yesterday? 'Yesterday father bought a car'

3.1 Production phase

Similar to Experiment 1, the data from the production phase was acoustically analyzed and used as stimuli for the perception study. Five native Bangla speakers (3 female, 2 male; all originally from West Bengal, India, allhad been in the U.S. for at most 5 years) participated in the production phase. None of the participants participated in the previous experiment. The design was the same as in the production phase of Experiment 1, with four conditions (Sub-New, Obj-New, Sub-Corr, Obj-Corr). We used the same 16 targets as in Experiment 1. However, now participants produced Adv-SOV sentences instead of SOV sentences on target trials. The SOV part of the sentences was the same as in Experiment 1, and we now added a sentence-initial adverb (ex.3). The adverbs were all single-word expressions in Bangla and were of three types: time adverbs (e.g., *yesterday, last_night*), manner adverbs (e.g., *immediately, quickly*), and place adverbs (e.g., *inside, there*). All of these can naturally and felicitously occur in sentence-initial position in

Bangla. A total of ten common adverbs were used, repeated twice, of which four were time adverbs, four manner adverbs, and three place adverbs. The adverb was also mentioned in the question, as shown in ex.(3), to maximize naturalness. The adverbs were single words and 2-3 syllables in length. The study also included 32 fillers, which were similar to those in Experiment 1 but some had adverbs in different positions – sentence-medial and sentence-initial – because adverbs were also added to the targets. The production phase was conducted in the same was as in Experiment 1; participants saw question-answer pairs and were instructed to say the answer aloud, and the second round of recordings was used for all analyses.

3.2 Perception phase

In this phase, similar to Experiment 1, we test whether participants can perceive differences between subjects and objects in new-information focus vs. Contrastive focus. The design of this perception study was the same as the perception study in Experiment 1. Twenty adult native speakers of Bangla from India participated (all had been in the U.S. for less than 5 years). None of them took part in the production phase of this experiment or in the perception phase of Experiment 1, but four of the participants had participated in the production phase of Experiment 1. This was primarily due to the lack of native Bangla speakers in and around University of Southern California. However, an average of 24 months passed between the time when these four participants completed the production phase of Experiment 1 and the perception phase of Experiment 2. Thus, we do not expect their participation to distort the data.

As in Experiment 1, the sound files from the production phase were used as the stimuli for the perception phase. There were 20 targets and 32 fillers, in a Latin-Square design. The lists were created such that every list had an equal distribution of target sentences spoken by all five speakers in all four conditions. The procedure was the same as the perception phase of Experiment 1.

3.2.1 Results for perception phase

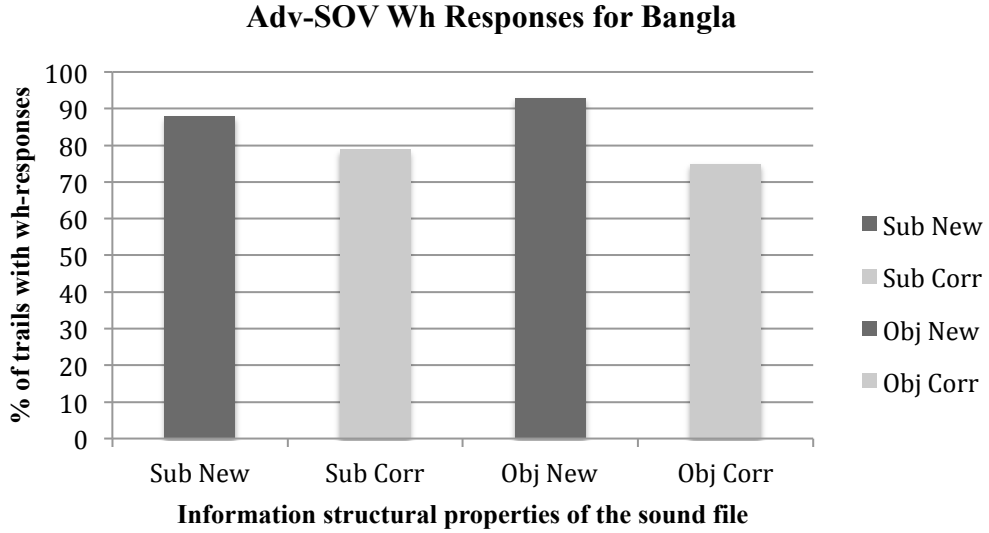
Similar to Experiment 1, we plot the results in terms of the percentage of 'wh-responses', shown in Figure 5. The rate of wh-responses can thus be thought of as the rate of new-information focus interpretations.

As a whole, Experiment 2 replicates the outcomes of Experiment 1. In object-focus conditions, the rate of wh-question choices is 93% for sound files elicited in a new-information focus context (wh-question), but only 75% for sound files been elicited in a corrective focus context (yes/no question). Similarly, in subject-focus conditions, the rate of wh-question choices for new-information focus sound files is 88%, and 79% for wh-question choices with corrective focus sound files. In other words, the overall wh-preference is again weakened for sound files elicited in a corrective context.

Linear mixed-effects regressions confirm this asymmetry: The difference in the rate of wh-question responses is significant in the case of object focus ($p<.05$) but not significant in the case of subject focus ($p=.11$). These results are consistent with the perception phase of Experiment 1. Thus, even with an adverb at the sentence-initial position, Bangla speakers are unable to perceive the prosodic difference between focus

types on subjects, although they can do so on objects. This suggests that the difference between subjects and objects is not an artifact of the subject's sentence-initial position.

Figure 5. Experiment 2: Rate of wh-question choices (indicating new-information focus)



Because the focus-types results for subjects are approaching marginal significance (commonly defined as $0.1 \leq p > 0.05$), we wanted to check whether we might be overlooking a meaningful result due to lack of statistical power. To test this, we decided to conduct a combined analysis on Experiments 1 and 2. However, before combining the perception data of Experiments 1 and 2, we first compared the experiments with repeated measures ANOVA with 'experiment' as a between-subjects factor, and found no significant effects of experiment ($p=0.419$). We then combined the perception data from Experiments 1 and 2 to see if effects that are almost marginal in the individual studies would reach significance with more data. However, mixed-effects logistic regression analyses yield the same results as we obtained for each experiment individually: With focused objects, we find significantly more wh-question choices when the object had been elicited with new-information focus than with corrective focus ($p<0.05$). However, for focused subjects there is no significant difference in the rate of selection of wh questions ($p=0.931$). These results corroborate our conclusion that native Bangla listeners can distinguish focus type on the (immediately preverbal) object, but not on the subject.

3.3. Acoustic analyses of production phase

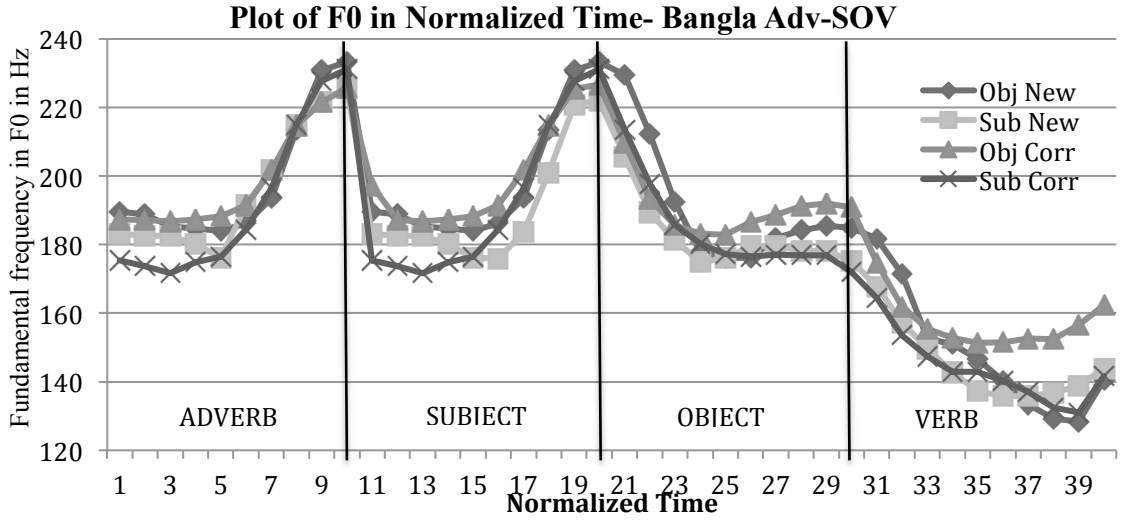
The acoustic data was analyzed in the same way as in Experiment 1, except that because the sentences contained an adverb, for the time-normalized analyses we now used 40 segments (segments 1-10: adverb, segments 11-20: subject, segments 21-30: object, segments 31-40: verb). (Similar to Experiment 1, we also looked at intensity, normalized over time, but found no significant effects of focus type, so we will not discuss the intensity data further in this paper.)

3.4 Results of production phase

Figure 6 shows the **fundamental frequency (*F0)*** patterns for Experiment 2, on Adverb-SOV sentences in Bangla. As can be seen in the figure, it looks like even with the presence of a sentence initial adverb, the subject still has a prosodic prominence, which may be related to prosodic resetting.[3] However, similar to SOV sentences in Experiment 1, we see that numerically, correctively focused objects (triangles) have a higher *F0* than new-information focused objects (diamonds). Statistical analyses (two-tailed paired t-tests) confirm that focused objects have significantly higher *F0* when they are contrastively focused than when they are in new-information focus ($t(4)=-4.850$; $P<0.05$), but no such difference is found on focused subjects ($t(4)= -1.946$; $P=0.124$). In sum, even when the subject is not at the sentence-initial position, the prosodic distinction between new-information focus and contrastive focus that we see on objects is still not reliably present on subjects.

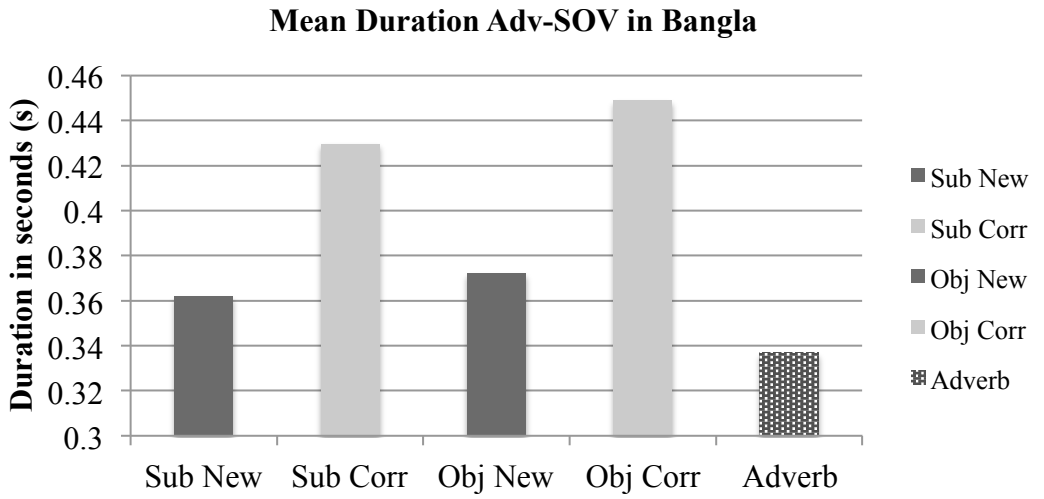
Figure 7 shows the **mean relative duration** of the focused constituent (subject or object) in all four conditions in Experiment 2. Similar to SOV sentences in Experiment 1, correctively-focused constituents appear to be considerably longer than constituents in new-information focus. Statistical analyses (two-tailed paired t-tests) confirm that the mean durations are significantly higher for elements in corrective focus than for elements in new-information focus (subjects: $t(4)=-3.111$, $p<0.05$; objects: ($t(4)=-3.420$ $p<0.05$). Thus, correctively-focused constituents are indeed significantly longer than constituents in new-information focus, regardless of grammatical role.

Figure 6. Experiment 2: Time-normalized F0 contours of the Adv-SOV target sentences



[3] A question that comes up regarding Figure 6 concerns the potential occurrence of an *F0* reset after the adverb and before the subject. It may be that the SOV part of the sentence forms a separate phonological phrase. This would explain why the subject (the first constituent of the hypothesized phonological phrase) has high *F0* and the downstep begins after the subject. We leave this as a question for future work.

Figure 7: Experiment 2: Mean relative duration of the focused constituent (and adverb)



3.5 Discussion of Experiment 2 (Bangla Adv-SOV)

Experiment 2 addressed a question left open by Experiment 1, namely whether the lack of significant focus type effects on the F0 measurements of subjects could be due to the sentence-initial position of the subject in Experiment 1. To test this, in Experiment 2 we used adverb-initial sentences. The results show that even when the subject is no longer sentence-initial, (i) Bangla speakers use *F0* to encode a difference between focus types on objects but *not* on subjects, and (ii) listeners are able to distinguish the two focus types reliably on objects but *not* on subjects. In sum, the results are very much in line with what we found in Experiment 1. Furthermore, it is also important to point out that in both Experiments 1 and 2, we find effects of focus type on the relative duration of both subjects and objects (correctively-focused elements are longer than elements in new-information focus). In light of this, it is intriguing that Bangla listeners do not seem to be sensitive to these duration cues. We return to this in the General Discussion section.

4. Experiment 3: Adv-SOV word order in Hindi

We conducted a study parallel to Experiment 2 in Hindi. As we will see, even two closely related languages do not pattern alike in terms of the prosodic encoding of focus types.

4.1 Production phase

The production phase was parallel to Experiments 1 and 2, except that the study was now conducted in Hindi. Five adult native Hindi speakers (3 female, 2 male; all living in Delhi, India at the time of testing and had lived there for most of their lives) participated. The participants did not speak any other Indian language apart from Hindi. The design was the same as in Experiments 1 and 2, with four conditions (Sub New, Obj New, Sub Corr, Obj Corr). Also, similar to Experiment 2, sentence-initial adverbs were used (4).

The design, methods and procedure were the same as Experiment 2, except for the language of the stimuli.

(4a) Subject wh-question to elicit new-information focus (Sub-New)
Q: parso kisne gari kharidi? A: parso [bhaiyya-ne]$_{newinfo}$ gari kharidi
day-before-yesterday who car bought day-before-yesterday brother car bought
'Who bought car day before yesterday? 'Brother bought car day before ystday'

(4b) Object wh-question to elicit new-information focus (Obj-New)
Q: pasro bhaiyya-ne kya kharida? A: parso bhaiyya-ne [gari]$_{newinfo}$ kharidi
day-before-yesterday brother what bought day-before-yesterday brother car bought
'What did brother buy day before yesterday? 'Brother bought car day before ystday'

(4c) Subject yes/no question to elicit corrective focus (Sub-Corr)
Q: parso papa-ne gari kharida kya? A: parso [bhaiyya-ne]$_{Corr}$ gari kharidi
day-before-yesterday father car bought Q day-before-yesterday brother car bought
'Did father buy car day before yesterday? 'Brother bought a car day before ystday'

(4d) Object yes/no question to elicit corrective focus (Obj-Corr)
Q: parso bhaiyya-ne skutar kharida kya? A: parso bhaiyya-ne [gari]$_{Corr}$ kharidi
day-before-yesterday brother scooter bought Q day-before-yesterday brother car bought
'Did brother buy scooter day before ystday?' 'Day before ystday brother bought car'

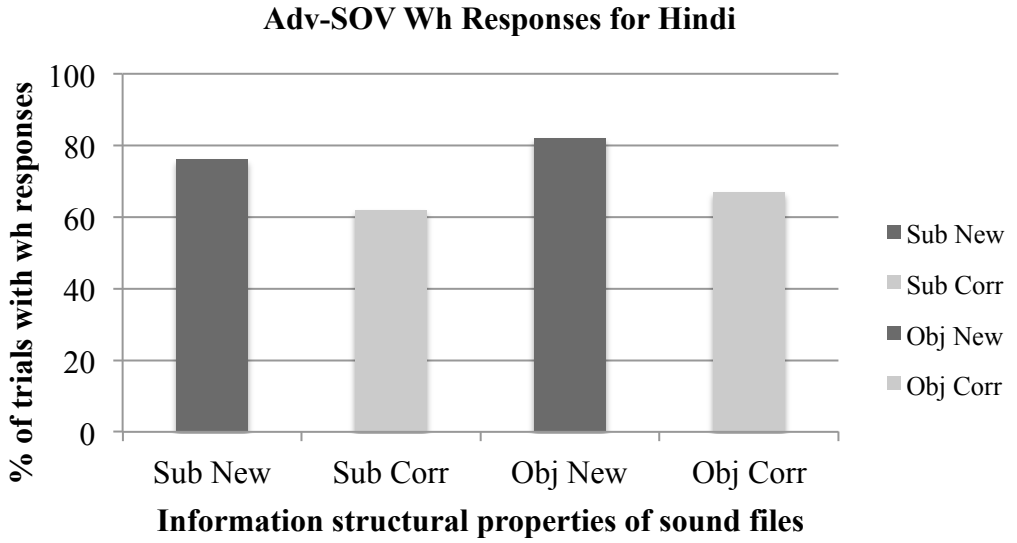
4.2 Perception Phase

The perception phase was again parallel to Experiment 2. Twenty adult native speakers of Hindi (all living in Delhi at the time testing, and had lived there for most of their lives) participated in the study. The participants did not speak any other Indian language. None of these participants took part in the production phase of this experiment. The design and procedure was he same as the perception phase of Experiment 2, except that the stimuli were now in Hindi.

4.2.1 Results for perception phase

We again present the results in terms of the percentage of wh-responses, i.e., the percentage of trials where the participants chose a wh-question (indicative of a new-information focus interpretation). In object-focus conditions, the rate of wh-question choices is 82% for sound files elicited in a new-information focus context (preceded by a wh-question), but only 67% for sound files elicited in a corrective focus context (preceded by a yes/no question). Similarly, in subject-focus conditions, the rate of wh-question choices for new-information focus sound files is 76%, and 62% for wh-question choices with corrective focus sound files. Linear mixed-effects regressions confirm that there is a main effect of focus type ($p<.05$), no effect of grammatical role ($p=.2$) and no interaction ($p=.76$). Thus, participants are able to perceive the difference between the two focus types equally well for subjects and objects. This is different from Bangla, where the participants were only able to distinguish the focus types for objects, and not subjects.

Figure 8. Experiment 3: Rate of wh-question choices (indicating new-information focus)



4.3. Acoustic analyses of production phase

For the acoustic analyses, the data was analyzed as in Experiment 2. (Similar to Experiments 1 and 2, we also looked at intensity, normalized over time, but found no significant effects of focus type, so we do not discuss the intensity data further here.)

4.4 Results of production phase

Figure 9 shows the **fundamental frequency (*F0)*** patterns for each of the four conditions in Experiment 3. We clearly see a general down-step pattern after the subject. However, the adverbs (segments 1-10) and subjects (segments 10-20) both have almost equally high (peak) *F0*'s. The *F0* starts to gradually fall after subjects such that the objects (segments 21-30) have a much lower (peak) *F0* than the subjects followed by the verbs (segments 31-40), which have the lowest *F0*. Statistical analyses show that there is no effect of focus type on either subjects or objects. Thus, unlike Bangla, where we found an effect of focus type on *F0*s for the object, in Hindi we do not find any *F0* differences in the focus types for either subjects or objects.

Figure 10 shows the **mean relative duration** of the focused constituent (subject or object) in all four conditions in Hindi. Similar to Bangla in Experiments 1 and 2, correctively-focused constituents appear to be considerably longer than constituents in new-information focus. Statistical analyses (two-tailed paired t-tests) confirm that the mean durations are significantly longer for elements in corrective focus than for elements in new-information focus (subjects: $t(4)=-3.679$ $p<0.05$; objects: $t(4)=-3.302$ $p<0.05$). Thus, contrastively-focused subjects and objects are produced with reliably longer duration than subjects and objects in new-information focus in Hindi. Thus, similar to Experiments 1 and 2 in Bangla, Hindi speakers encode the difference between the focus types using duration, and these duration cues are equally strong for subjects and objects.

Figure 9: Experiment 3: Time-normalized F0 contours of the Adv-SOV target sentences

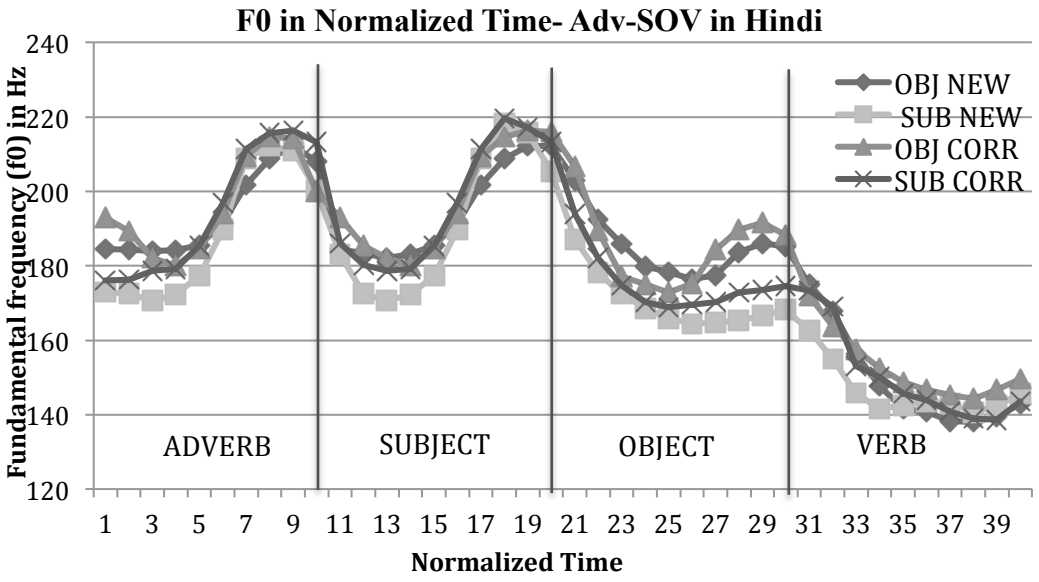
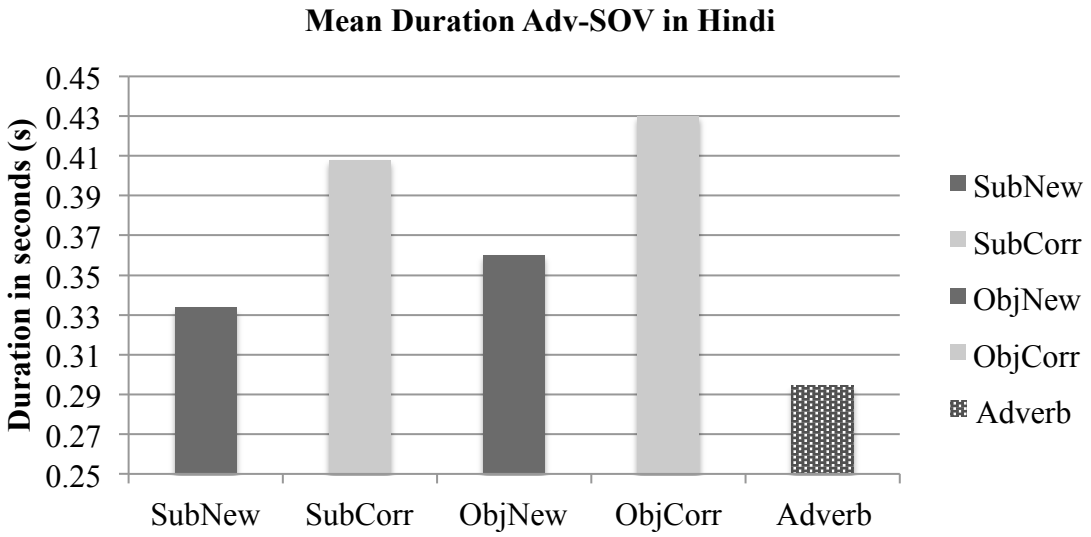


Figure 10: Experiment 3: Mean relative duration of the focused constituent (and adverb)



4.5 Discussion of Experiment 3 (Hindi Adv-SOV)

The acoustic analyses for Hindi indicate that speakers (i) are not using *F0* to encode focus types, but (ii) are using duration, because words in corrective focus are significantly longer than words in new-information focus. The results of the perception study show that Hindi listeners can reliably differentiate between the two focus types on both subjects and objects. In light of the acoustic analyses, this leads us to conclude that Hindi listeners are sensitive to duration cues but not to *F0* cues.

5. General Discussion

We conducted three production-perception studies to investigate whether speakers of Bangla and Hindi encode the distinction between new-information focus and corrective focus prosodically in sentences with canonical SOV order, and if so, what prosodic dimensions (e.g. *F0*, duration, intensity) are used, whether they are the same in both languages. We also tested whether listeners of these two languages are able to perceive the prosodic cues used to signal focus types (in the absence of any word order cues). These studies provide new empirical information about Hindi and Bangla and also contribute to the longstanding debate regarding the status of the focus types, i.e. whether new-information and corrective focus should be regarded as distinct notions, or as two subtypes that do not differ in their basic semantics.

Our results for Bangla show that Bangla speakers use *F0* to encode the distinction between new-information focus and corrective focus on objects but not on subjects. Interestingly, however, Bangla speakers use duration to distinguish focus types on both subjects and objects. Hindi speakers, on the other hand, do not seem to be using *F0* at all to encode distinctions in focus type on either subjects or objects. Like Bangla speakers, however, they produce correctively-focused elements with reliably longer duration than elements in new-information focus.

The finding that speakers in these two languages reliably distinguish new-information focus and corrective focus provides support for the idea that these two focus types are distinct categories. As discussed in section 1, there has been a long-standing debate regarding whether different focus types are information-structurally distinct. On the one hand, linguists like Kiss (1998), Vallduví &Vilkuna (1998), Hartmann & Zimmermann (2006) and Zimmermann & Onea (2011), distinguish between contrastive and new-information focus. On the other hand, others like Rooth (1992), Schwarzschild (1999) and Krifka (1993) argue that contrastive/corrective focus is not fundamentally different from information focus, as far as its underlying semantics are concerned. We find that speakers of both Bangla and Hindi produce elements in new-information focus and in contrastive focus with reliable prosodic differences, which is compatible with the idea that these two focus types are information-structurally distinct.

What about the perception side? Duration provides a reliable cue for focus types in both Bangla and Hindi on both subjects and objects, so it seems reasonable to expect listeners to be able to distinguish new-information focus from corrective focus on both subjects and objects in both languages. However, as we saw, whereas Hindi listeners are indeed able to do so, Bangla listeners are only able to distinguish focus types reliably when the object is in focus, not when the subject is in focus. (The data for the subject are going in the right direction, but do not reach significance. Thus, even if we were to say, optimistically, that Bangla listeners are somewhat able to distinguish focus types on the subject, they are *much* better at doing so when the focused element is the object.)

When we combine the production and perception data, we can identify some intriguing asymmetries between Hindi and Bangla. Although both languages encode focus type distinctions using duration, it appears that Hindi listeners rely on duration cues during perception more than Bangla listeners. Bangla listeners, in contrast, seem to rely more on *F0* cues (and are thus not able to reliably detect differences in focus type on the subject). Thus, our work provides the first crosslinguistic psycholinguistic

evidence that that Bangla and Hindi, two closely-related modern Indo-Aryan languages which share a lot of common syntactic properties, differ somewhat in the specifics of how focus types are encoded by speakers, but differ more strikingly in how sensitive listeners are to different kinds of prosodic cues. The deeper reasons for these differences are an intriguing direction for future work.

References

Boersma, Paul, and David Weenik. 2009. Praat: Doing phonetics by computer, www.praat.org

Bolinger, Dwight. 1961. Contrastive Accent and Contrastive stress. *Language* 37, 83–96.

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

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. *Subject and Topic*, ed. Charles. Li. London, New York: Press. 25-56.

Chen, Ying., Yi Xu, and Susan Guion-Anderson, S. 2014. Prosodic realization of focus in bilingual production of Southern Min and Mandarin. *Phonetica* 71, 249-270.

Choudhury, Arunima. 2015. Interaction between Prosody and Information structure: Experimental evidence from Hindi and Bangla, Doctoral Dissertation. USC

Choudhury, Arunima. 2010. A Study of Focus in Bangla, M.A. Paper, USC, Los Angeles

Dik, Simon. 1997. The Theory of Functional Grammar. *Part 1: The Structure of the Clause.* Kees Hengeveld (ed.) Berlin, New York: Mouton de Gruyter.

Féry, Caroline, and Manfred Krifka. 2008. Information structure: Notional distinctions, ways of expression. In *Unity and diversity of languages*, 123-136. P. van Sterkenburg (eds.) Amsterdam: John Benjamins.

Féry, Caroline. 2010. The intonation of Indian languages: an areal phenomenon. *Problematizing Language Studies*. (ed.) Imtiaz Hasnain and Shreesh Chaudhury. Festschrift for Rama Agnihotri. Delhi: Aakar Books, 288-312.

Gussenhoven, Carlos. 2007. Types of focus in English. *Topic and Focus*: *Cross-linguistic Perspectives on Meaning and Intonation* (eds.) Chungmin Lee, Matthew Gordon, and Daniel Büring, 83–101.

Gussenhoven, Carlos. 2008. Notions and subnotions in information structure. *Acta Linguistica Hungarica* 55, 381-395.

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

Hartmann, Katharina, and Matle Zimmermann. 2006. *Interdisciplinary Studies on Information Structure.* (eds.) Ishihara, S., M. Schmitz, and A. Schwarz. 5, 61-105

Halliday, Michael. 1967. *Intonation and grammar in British English*. The Hague: Mouton.

Hayes, Bruce, and Aditi Lahiri. 1991. Bengali Intonational Phonology. Natural Language and Linguistic Theory 9, 7-96

Ito, A., Nagasawa, T., Ogasawara, H., Suzuki, M., & Makino, S. (2006). Automatic Detection of English Mispronunciation Using Adaptation and Automatic Assessment of English Intonation and Speaker Rhythm. Educational Technology Research, 29:13-23. Japan Society of Educational Technology

Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Kanerva, Jonni M. 1990. *Focus and Phrasing in Chichewa Phonology*. NY: Garland.

Katz, Jonah, and Elisabeth Selkirk. 2011. Contrastive focus vs. discourse-new: evidence from prosodic prominence in English. *Language*, 87(4), 771-816.

Khan, Sameer. D. 2008. The Intonational Phonology and focus prosody of Bengali. Doctoral Dissertation. UCLA, Los Angeles.

Kidwai, Ayesha. 2000. *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu.* Oxford: Oxford University Press.

Kiss, Katalin É. 1998. Identificational Focus versus Information Focus. *Language* 74: 245–273.

Krifka, Manfred. 1993. Focus and presupposition in dynamic interpretation. *Journal of Semantics* 10, 269-300.

Molnár, Valéria. 2002. Information structure in a cross-linguistic perspective. *Language and computers.* H. Hasselgård, S. Johansson, B. Behrens, & C. Fabricius-Hansen (eds.), 39, 147-161. Amsterdam, The Netherlands: Rodopi.

Moore, Robert R. 1965. A study of Hindi intonation. Doctoral Dissertation, University of Michigan.

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

Pierrehumbert, Janet. 1980 The phonology and phonetics of English intonation. Doctoral Dissertation, MIT. Distributed 1988, Indiana University Linguistics Club.

Rochemont, M. S. (1986). Focus in generative grammar. Amsterdam, Philadelphia: John Benjamins.

Rooth, Mats. 1992. A Theory of Focus Interpretations. *Natural Language Semantics,* 1, 75–116.

Schwarzschild, Roger. 1999. Givenness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, 7, 144-177.

Vallduví, Enric and Maria Vilkuna. 1998. On Rheme and Kontrast. *In The Limits of Syntax,* (eds.) Peter Culicover & Louise McNally. New York: Press, 79–106.

Watson, Duane G., Michael K Tanenhaus, and Christine A Gunlogson. 2008. Interpreting Pitch Accents in Online Comprehension: H* vs. L+H*. *Cognitive Science* 32, 1232–1244

Wu, Wing Li and Yi Xu. 2010. Prosodic Focus in Hong Kong Cantonese without Post-focus Compression. In *Proceedings of Speech Prosody 2010*, Chicago

Xu, Yi. 2005-2012. ProsodyPro.praat: http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/

Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27. 55-105

Zimmermann, Malte and Edgar Onea. 2011. Focus Marking and Focus Interpretation *Lingua* 121: 1651-1670

Zubizarreta, Maria Louisa. 1998. *Prosody, Focus, and Word Order*, Linguistic Inquiry Monograph 33, MIT Press.

Arunima Choudhury & Elsi Kaiser
arunimac@usc.edu & emkaiser@usc.edu