Phrasing and Focus in Bengali

Sameer ud Dowla Khan
UC Los Angeles

Over 15 years after the prosody of Bengali was first analyzed using the Autosegmental-Metrical framework (Hayes & Lahiri 1991), linguists still disagree on many fundamental aspects of the system. Claims about focus, phrasing, and lexically-specified tone are constantly reanalyzed, and before any further details of the system are to be explored, some basic questions need to be addressed:

1. What is the inventory of tones in the language?
2. How many layers of prosodic structure are tonally marked?
3. What are the phonological restrictions on tonal sequences in the language?
4. How is focus realized with respect to prosody?

Part of the reason why these questions have not been adequately answered in the past is the lack of variety in the data collected and analyzed in previous studies. In order to set up an empirically-testable model of Bengali prosody within the framework of Intonational Phonology (Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988; see Ladd 1996 for review), the current study examines an extensive set of recordings of various types of sentences produced by a group of 25 adult native speakers of Standard Bangladeshi Bengali, collected in a recent experiment. Each speaker read 57 sentences aloud into a hand-held microphone connected via a USB cable to a laptop. The sentences were mostly made up of vowels and sonorant consonants. Their pragmatic, semantic, syntactic, and phonological properties were manipulated to test for: (a) different boundary tones, by using particular particles (*e.g.* focus clitics, interrogative particles) and punctuation (*e.g.* commas, question marks), and (b) the realization of focus domains, by using particular words of varying size and by using context sentences (*i.e.* preceding wh-questions eliciting focus on particular constituents). Recordings were made in WaveSurfer (Sjölander & Beskow 2005), and pitch events were analyzed in Praat (Boersma & Weenink 2005). Observations and measurements made using this new corpus of data reveal three fundamental facts that have been previously overlooked or refuted:

1. Both the pitch accent and boundary tone inventories include contour tones.
2. There are three layers of tonally-marked prosodic structure.
3. Tonal sequences are not constrained by the OCP.
4. Focused constituents bear rising pitch without the use of a boundary tone.

The following is a brief outline of B-ToBI (Bengali Tones and Break Indices),[1] a model of and transcription system for Bengali prosody, supplemented with images of representative pitch tracks and with acoustic measurements, to support the aforementioned claims of the current analysis. The B-ToBI model draws on elements of the general ToBI framework of prosodic transcription and analysis (Beckman &

[1] A preliminary version of the B-ToBI model was proposed in Khan (2007).

Hirschberg 1994); like other Intonational Phonology models of prosody, the proposed B-ToBI model recognizes two types of tones, characterized by their function and alignment: pitch accents mark the heads of prosodic units, and are aligned to phonologically stressed syllables, while boundary tones mark the edges of prosodic units, and are aligned with the right edges of phrases.

Prominence: Stress and pitch accent

In accordance with most previous work on Bengali, the current model assumes that native monomorphemic words consistently bear a phonetically weak primary stress on the initial syllable, and only stressed (*i.e.* initial) syllables can bear a head-marking tone (*i.e.* pitch accent). Three pitch accent types are found:

1. **Low (L*):** default pitch accent
2. **High (H*):** marks sarcastic or unexpected information
3. **Rising (L*+H):** marks focused constituents

Low pitch accents (L*) are realized as an F_0 minimum during a prominent syllable, while high pitch accents (H*) are realized as an F_0 peak. Both of these pitch accents have been described in previous studies, and will thus not be described in further detail. The rising pitch accent (L*+H), however, has only been described by one previous model of Bengali (Michaels & Nelson 2004, unpublished). Rising pitch accents (L*+H) are realized as an F_0 minimum during the prominent syllable, with a sharp rise in pitch in the following syllables. This analysis is in contrast with most previous studies, which have explicitly stated that Bengali does not bear bitonal pitch accents (Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999, among others). Pitch tracks of these three pitch accent types are provided below in Figures 1-3.

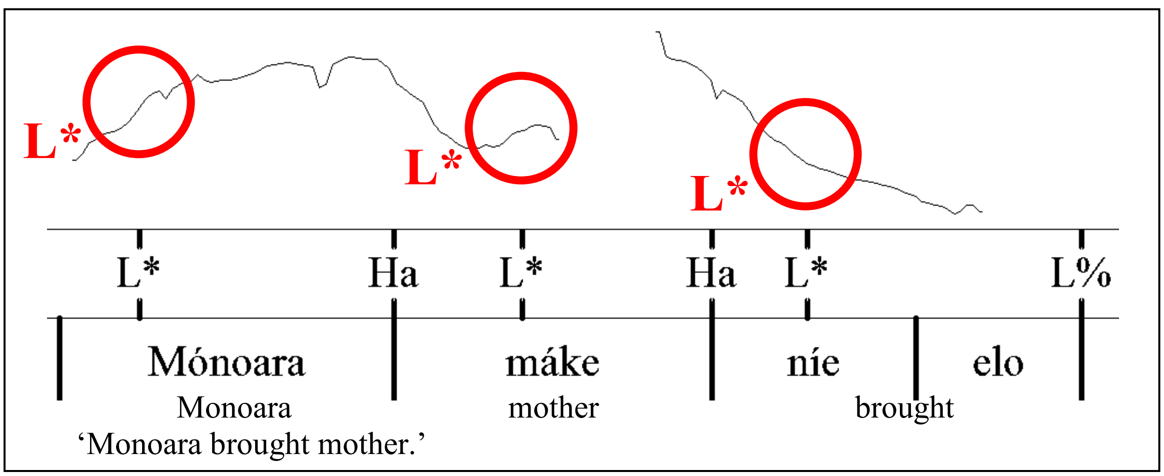


Figure 1. Three content words bear L*, realized as low pitch on the stressed syllable. [Ba19]

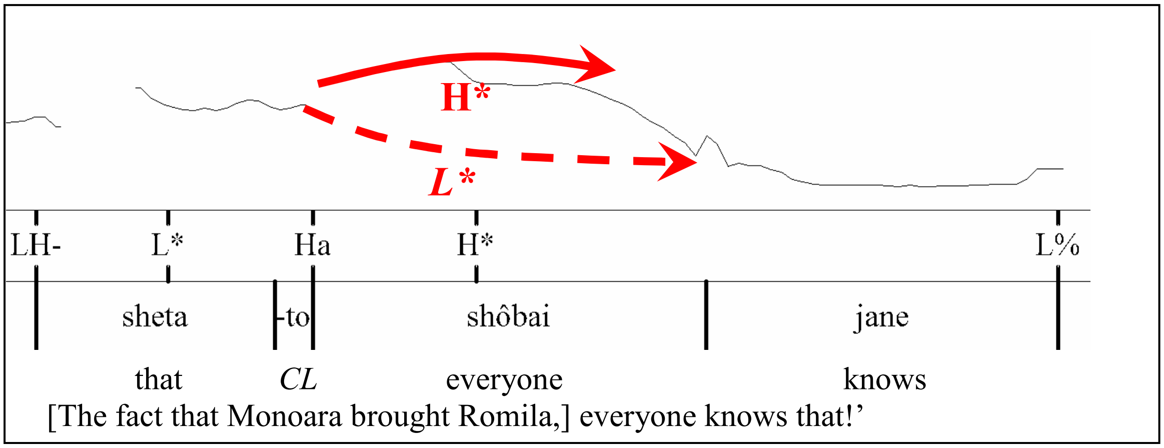


Figure 2. The subject *shôbai* 'everyone' bears H* instead of the default L* (predicted L* shown with dotted line; *cf.* L* on *shéta* 'that'). The H* is realized as a high pitch during the stressed syllable. [Re15]

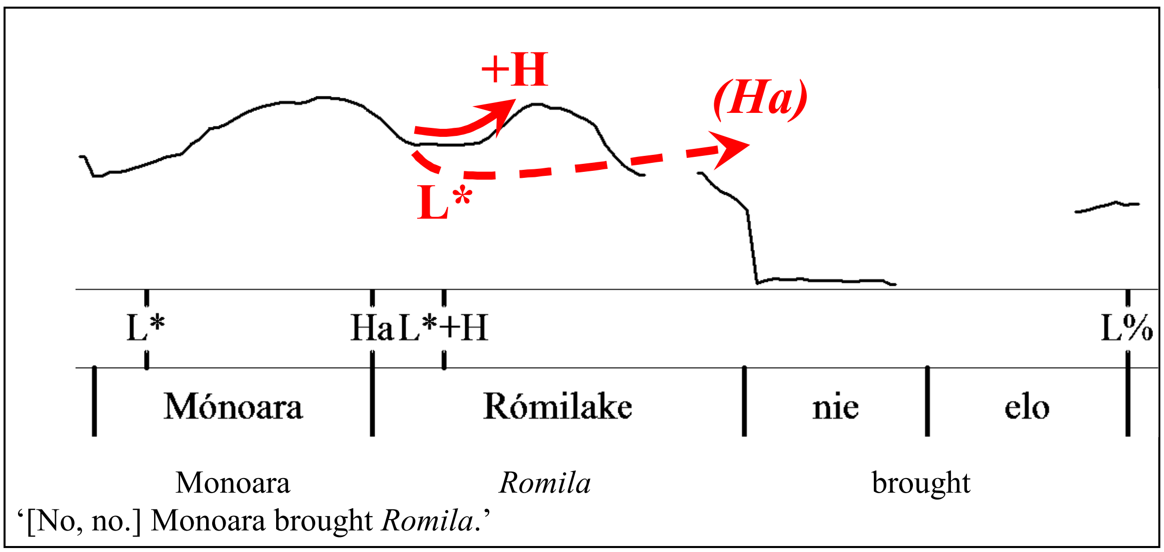


Figure 3. The focused object *Rómilake* 'Romila (ACC)' bears L*+H, realized as a low pitch on the stressed syllable followed by a sharp rise in pitch. Note the early rise and fall in F_0, in contrast to the default L*…Ha pattern expected in non-focused contexts (as shown by the dotted line). [Re23]

In depth: L vs. L*+H*

While all previous studies have agreed on the existence of bitonal boundary tones in Bengali, most studies point out a corresponding lack of bitonal pitch accents. The current analysis disagrees with this second statement; in addition to a contrast between level and bitonal boundary tones, the B-ToBI model includes a contrast between low (L*) and rising pitch accents (L*+H), a contrast that was once thought to be impossible. In the current analysis, the two pitch accent types are shown to vary in a number of ways; the most obvious differences are that L*+H (a) rises earlier than the rise for L*…Ha, (b) often includes falling pitch in longer words, and (c) reaches a higher pitch than surrounding Ha boundary tones do. See Figure 4 for a minimal pair of the two contours.

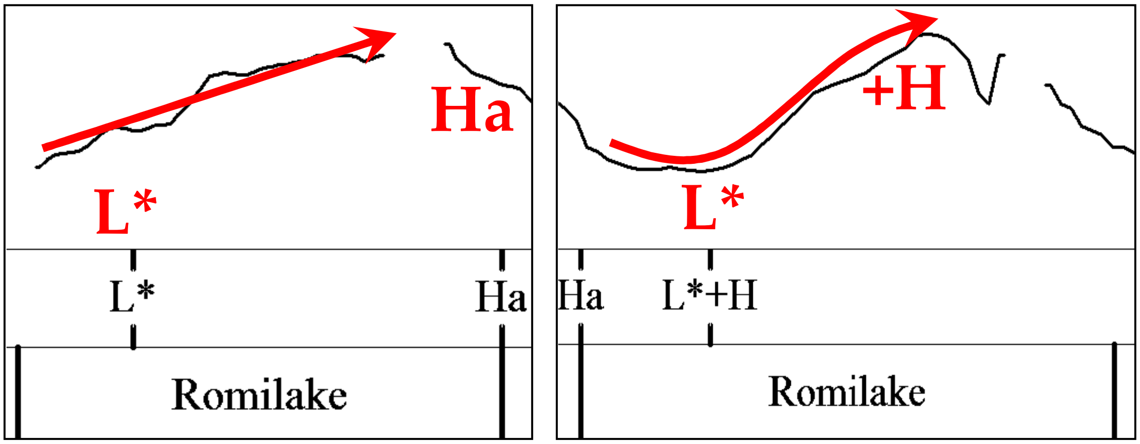


Figure 4. A minimal pair for the name *Romilake* 'Romila-ACC' produced by the same speaker in two different sentences, one in which the word was not focused (left), bearing a low pitch accent (L*) and high boundary tone (Ha), and one in which it was focused (right), bearing a rising pitch accent (L*+H). [Ro01, Ro23]

Measurements of the pitch differences between these tonal sequences are provided in the section on focus realization further below.

Phrasing and boundary tones

Unlike the two-phrase systems proposed in previous analyses, acoustic data from the current study reveal three layers of tonally-marked prosodic phrasing. The smallest phrase is roughly the size of a content word. This phrase has been called the Phonological Phrase (Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999), the Intermediate Phrase (Michaels & Nelson 2004), the Major Phrase (Selkirk 2006), and the Accentual Phrase (Jun 2005). The current analysis uses the latter term (Accentual Phrase, or AP), as the phrase roughly parallels structures with the same name in other languages including Farsi, Japanese, Kiche (K'iche'), Korean, and Northern Bizkaian Basque. As in Kiche (Nielsen 2005) and Farsi (Mahjani 2003, Esposito & Barjam 2007, Scarborough 2007), each pitch accent serves as the head of one and only one AP, and each AP contains one and only one pitch accent. As in Farsi, Japanese (Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988), Kiche, Korean (Jun 1993, 1998), and Northern Bizkaian Basque (Hualde 1988, 1999, Elordieta 1998), the AP is approximately the size of a content word, and often incorporates surrounding functional particles into the same AP. And as in Farsi, Japanese, Kiche, and Korean, the AP bears a boundary tone at its right edge; in Bengali, it is specifically a high tone:

1. **High (Ha)**: default

The default combination of the low pitch accent (L*) and a local high boundary tone (in the current model, Ha) has been described in all previous AM models of Bengali intonation as the most commonly-observed tonal pattern in the language. An alternative analysis might posit that this rising pitch pattern is not the result of interpolation between a low pitch accent and high boundary tone, but a bitonal, rising pitch accent as has been proposed for Greek (Arvaniti & Baltazani 2000, 2005). This alternative analysis, however, is not appropriate for Bengali, as the rising pitch pattern in Bengali non-focused constituents does not pattern like the rising pitch accent (L+H*) of Greek. Hypothesizing a rising pitch accent for this Bengali structure yields three obvious shortcomings:

(a) The highest pitch in the default contour is consistently found at the word's right edge, regardless of word length. In Figure 5 below, the direct objects *make* 'mother-ACC' and *Lina mamike* 'Aunt Lina-ACC' are produced without focus, and thus bear a low pitch accent (L*). Regardless of the length of the word (ranging from two syllables to five), the high tone is borne at the right edge of the constituent. This suggests that the high pitch at the end of a non-focused constituent is associated with the boundary and not the pitch accent, supporting the Ha boundary tone hypothesis.

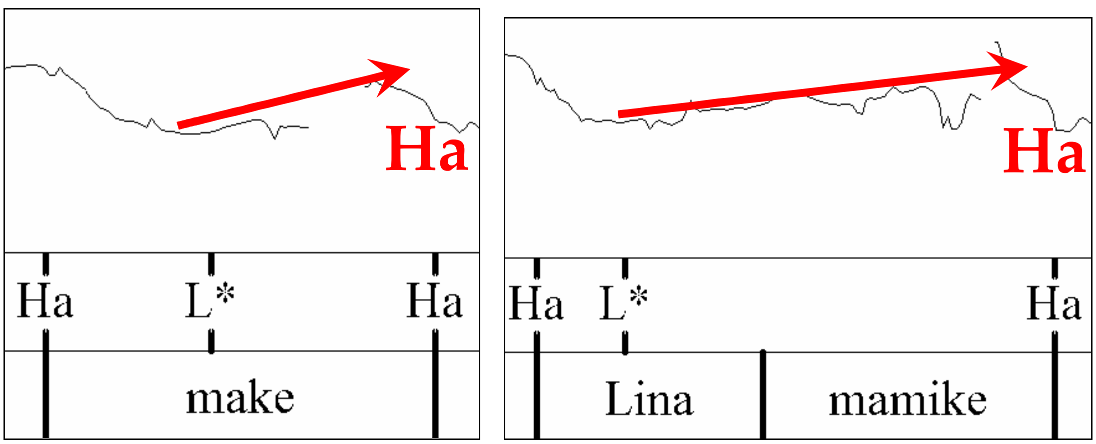
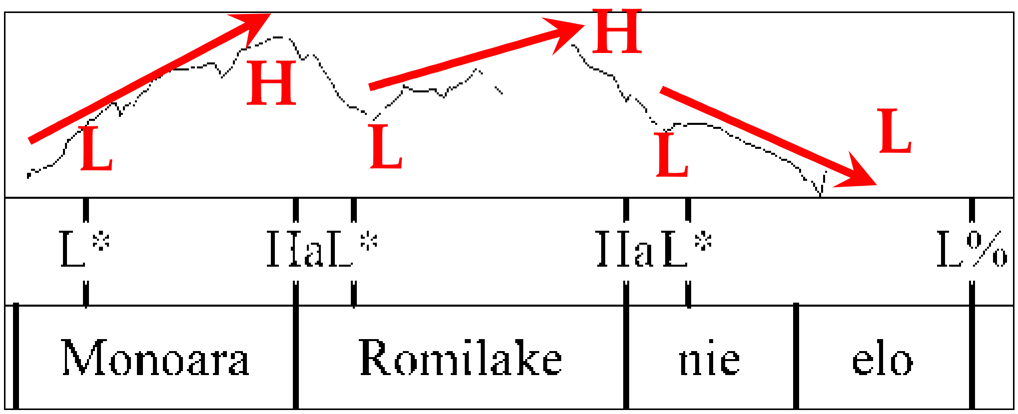


Figure 5. These two non-focused constituents (*make* 'mother-ACC' and *Lina mamike* 'Aunt Lina-ACC') illustrate how the high pitch is consistently realized at the right edge of the constituent, regardless of the constituent's length (measured here as the number of syllables), supporting analyses that label it a high boundary tone as opposed to a complex pitch accent. [To19], [To24]

(b) The highest pitch in the default contour is not realized when the word's right edge coincides with the boundary tone of a higher phrase, as shown in Figure 6 below. This analysis adopts the assertion that a boundary tone (Ha) is not realized when it coincides with the boundary tone of a higher phrase (T- or T%), indicating *boundary tone overriding*, which has also been shown in languages such as Korean (Jun 1999), rather than arguing that a bitonal pitch accent (L*+H) becomes a monotonal accent (L*) when it occurs before a boundary tone (T- or T%), which would suggest a constraint on pitch accent-boundary tone interaction.[2]

[2] One could propose that the final tonal contour of the sentence differs from the preceding contours due to its status as the nuclear pitch accent (as proposed in Hayes & Lahiri 1991 and Lahiri & Fitzpatrick-Cole 1999 – although these analyses still propose contours made up of low pitch accents and high boundary



Monoara Romila brought
'Monoara brought Romila.'

Figure 6. The high pitch found at the ends of non-focused constituents (in the current analysis, Ha) is not found when the constituent's right edge coincides with the boundary of a larger phrase (which, in this case, is the low IP boundary tone L%). [Az01]

(c) The current analysis already posits a separate rising pitch accent (L*+H) for focused constituents. As mentioned before, the rising pitch accent in a focused constituent does not behave like the rising contour in a non-focused constituent in that L*+H rises earlier than the rise for L*…Ha, includes falling pitch in longer words, and reaches a higher pitch than surrounding Ha boundary tones do. Figure 4 is repeated below in Figure 7 as a visual comparison of the two contours.

tones rather than rising pitch accents). In the current data set, however, there is no evidence to suggest that the last AP in an IP (or ip) should be considered "nuclear", as it is no more prominent than preceding APs.

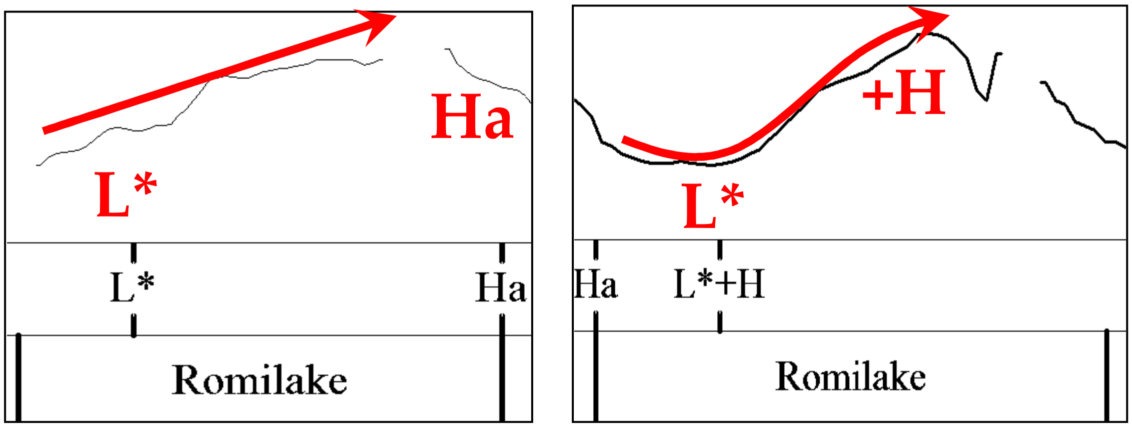


Figure 7. A comparison of the same word *Romilake* 'Romila-ACC' produced by the same speaker but in two different sentences, one in which the word was not focused (left), bearing a low pitch accent (L*) and high boundary tone (Ha), and one in which it was focused (right), bearing a rising pitch accent (L*+H). [Ro01, Ro23]

Also described in all previous studies of Bengali intonation is the Intonation Phrase (IP), which is the largest tonally-marked phrase. An IP can bear one of four boundary tones, primarily realized on the last syllable of the phrase. This also largely reflects previous models' analyses:

1. **Low (L%):** marks declarative sentences and focused wh-questions
2. **High (H%)**: marks echo wh-questions and various other interrogatives
3. **Rising (LH%):** marks default wh-questions
4. **Falling (HL%):** marks yes-no questions

Between the AP and IP levels of phrasing is the Intermediate Phrase (ip), which has been proposed for several languages, although not for Bengali (except for Michaels & Nelson 2004, which uses the label "intermediate phrase" for what the current study calls an accentual phrase). The ip can bear one of three boundary tones:

1. **Low (L-)** and **Rising (LH-):** both mark larger units, such as because-clauses, if-clauses, and relative clauses.
2. **High (H-):** marks smaller units, such as locative phrases and preposed topics

As all phrase types (AP, ip, IP) have the option of bearing a high tone at their right edge, a three-way comparison of Ha, H-, and H% is provided below.

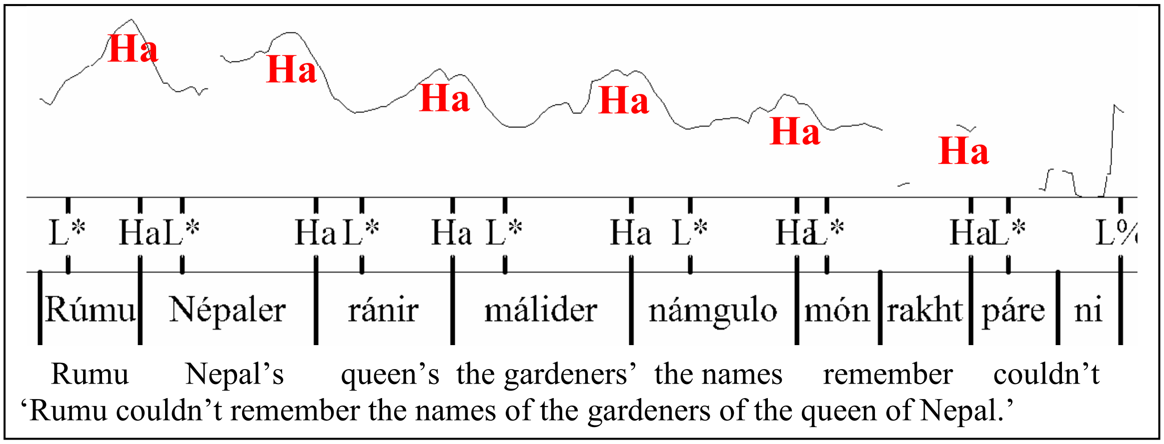


Figure 8. Each non-final content word bears Ha. [Fa50]

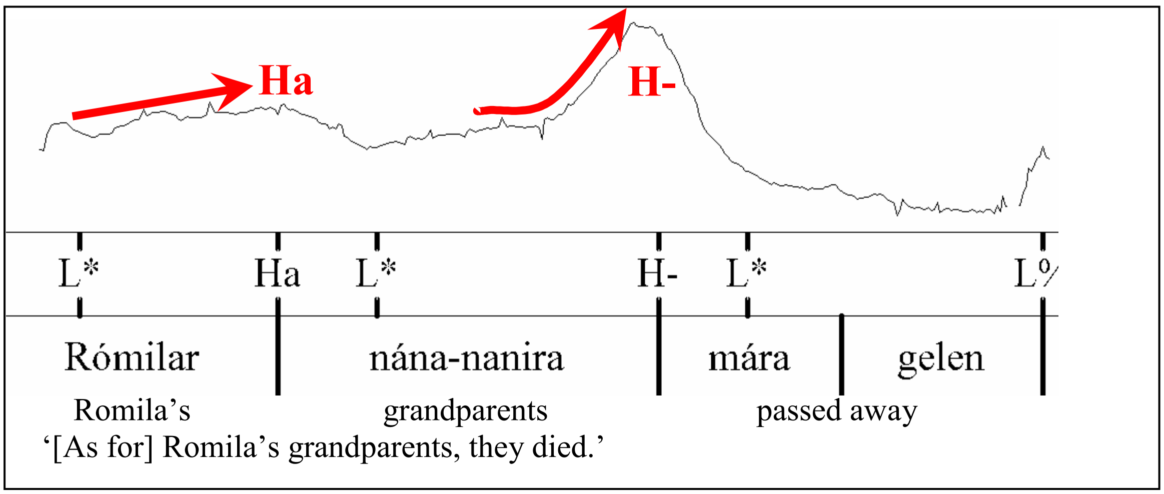


Figure 9. To indicate that the multi-word subject of the sentence (the ip *Rómilar nána-nanira*) is a topic, translated into English as something like 'As for Romila's grandparents', the phrase is marked on its right edge by H-, realized as a sharp F_0 rise on the ip-final syllable. (This differs from LH-, which includes a dip in F_0 before the final rise). [Na49]

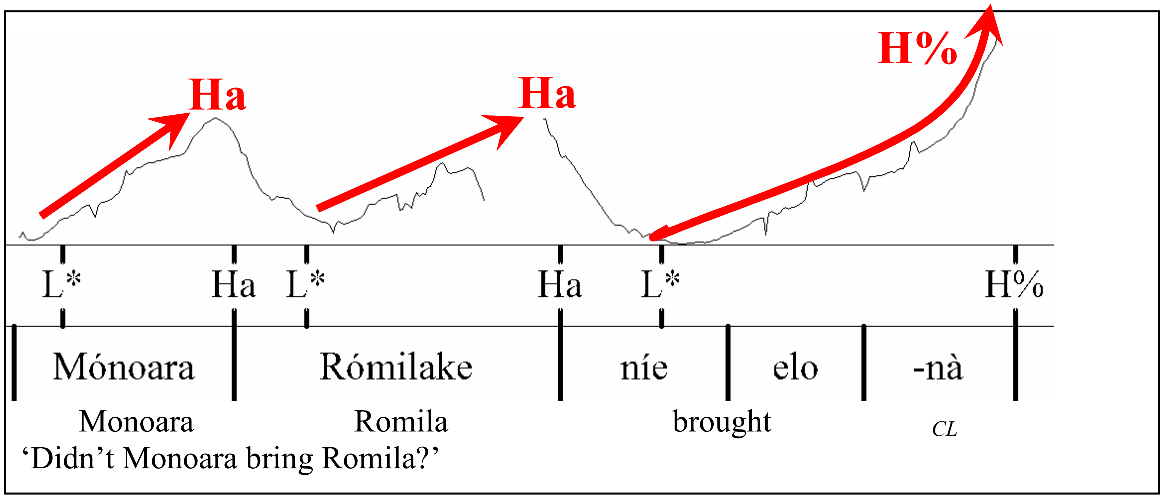


Figure 10. This question is marked with H%, realized as a rise in pitch from the final pitch accent rightwards, and an extreme rise on the final syllable. [Fa06]

These three prosodic phrases not only differ in their semantic/syntactic properties, but also in their acoustic properties, both durational and tonal:

1. **Duration**: the duration of the final syllable of an IP is longer than that of an ip, which is longer than that of an AP, which is longer than that of a word.
2. **Tone interpolation**: IP and AP boundary tones cause pitch interpolation starting just after the pitch accented syllable, while ip tones only cause pitch interpolation in the final syllable. In the current analysis, this is referred to the ip boundary tone locality constraint.[3]
3. **Tone height**: the pitch of an IP boundary tone is more extreme (lower L, higher H) than that of the corresponding ip tone, which is more extreme than that of the corresponding AP tone. For a three-way comparison, the maximum pitch of H% is higher than that of H-, which is higher than that of Ha.

Representative examples supporting all three properties described above are provided below, and in the case of the duration and tone height properties, acoustic measurements are also provided.

AP vs. ip vs. IP: Durational differences

One important difference between APs, ips, and IPs is final syllable duration. The final syllable of a word found at the end of an IP is far longer than the final syllable of the same word found at the end of an AP. The final syllable of a word found ip-finally shows moderate lengthening. One clear way to illustrate this difference is to measure the

[3] Due to the effects of this constraint, it is often ambiguous whether a given ip tone is high (H-) or rising (LH-). The clearest examples of each tone type indicate that the high ip boundary tone (H-) causes pitch to rise sharply in the final syllable, while the rising ip boundary tone (LH-) causes pitch to first dip in the penultimate syllable and then rise in the final syllable.

duration of the final syllable of a word in different prosodic positions. Factors such as speech rate can be controlled by measuring the duration of the final syllable as a percentage of the whole word's duration. Because the phrasing of a sentence into APs and ips is highly variable, and vulnerable to factors including speech rate and pragmatics, AP and ip boundaries are often used in the same point in a sentence spoken by two speakers, or the same speaker on different occasions. By measuring the percent duration of the final syllable of the same word spoken by the same speaker in a minimal or near-minimal set of sentences, it can be seen that ip-final syllables are statistically longer than AP-final syllables [paired t(8) = 3.05, p < .05], as shown in Figure 11 below.

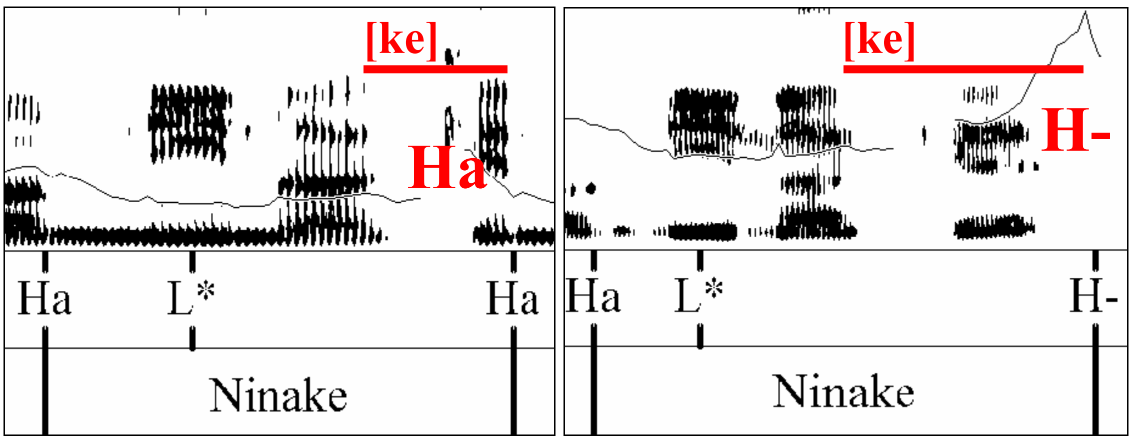


Figure 11. A comparison of the same word *Ninake* 'Nina-ACC' produced by the same speaker in two different sentences, one where the word occurred at an AP break and one where it occurred at an ip break. In the first example, the duration of the word *Ninake* is 403ms, while the final syllable duration is 120ms; in the second example, the duration of the word *Ninake* is 540ms and the final syllable duration is 238ms. A spectrogram has been included in the background to clarify the location of the segment boundaries, and red lines indicate the duration of the final syllable [ke]. [Pi22, Pi21]

Durational measurements of the final syllables of APs and ips consisting of the same text (measured as a percentage of the total word duration) are found in Table 1 below, and are illustrated in the bar graph in Figure 12 below. The boundary tone was always a high tone (Ha or H-, depending on the phrase measured); contour tones are not included because they tend to lengthen the final syllable more extremely and would lead to an unfair comparison.

Speaker	AP-final	ip- final	Word measured	Examples used
Ba	33.1%	35.3%	Monoara	01 & 23
Ba	36.1%	34.1%	Monoara	29 & 28
BM	34.1%	43.5%	Romilake	02 & 01
Do	20.2%	22.6%	Monoara	01 & 01
Jh	21.8%	31.6%	Monoara	01 & 08
Jh	29.7%	44.0%	Monoara	02 & 04
Jh	33.1%	35.9%	Monoara	01 & 10
Pi	33.1%	35.3%	Ninake	22 & 21
Sh	26.1%	31.0%	Monoara	02 & 01
Sh	34.1%	43.5%	Monoara	21 & 22

Table 1. Durational differences between the final syllables of ten APs and of identical ips, measured as a percentage of the total word duration, measured within speaker.

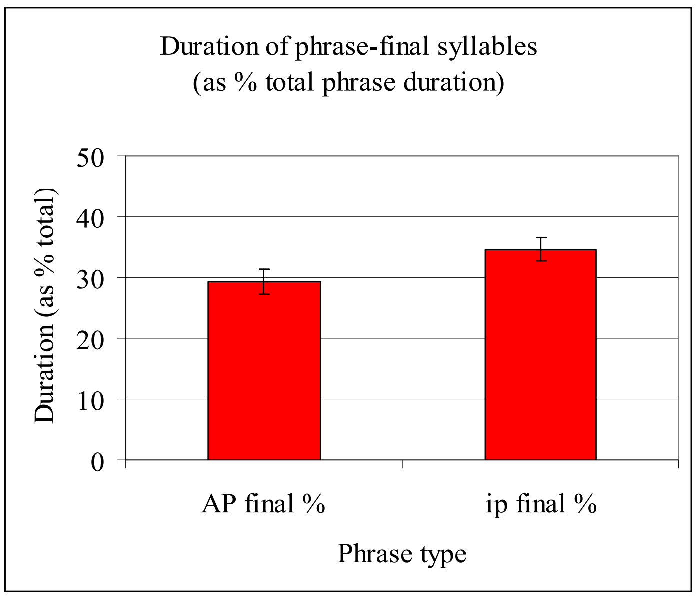
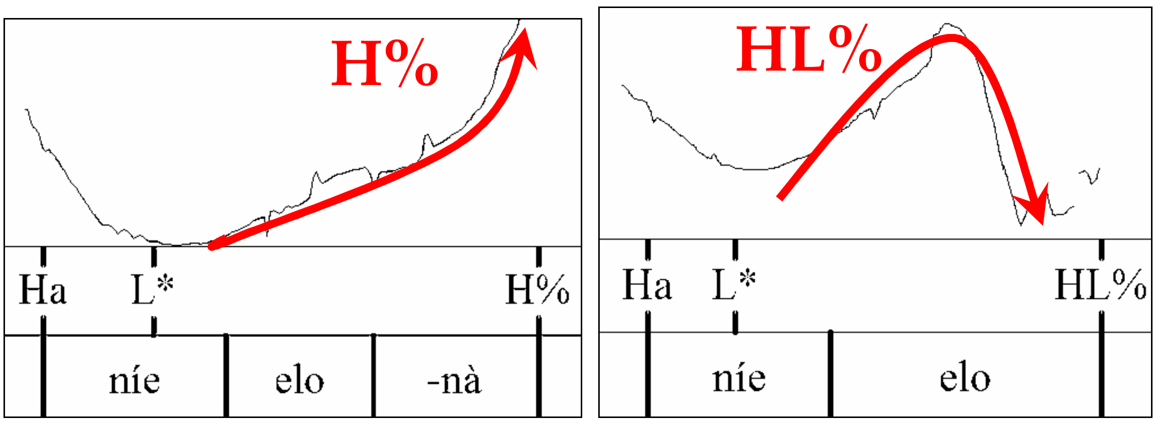


Figure 12. Durational differences between the final syllables of ten APs and of identical ips, measured as a percentage of the total word duration, measured within speaker. Error bars indicate standard error.

Unfortunately, the same comparison cannot be made for the AP-IP or ip-IP contrasts, as IP boundaries most often occur in different syntactic positions than AP or ip boundaries, thus preventing any direct comparison of truly minimal sets in the current study's corpus of data. The only near-minimal ip-IP sets involved a contour ip boundary tone (LH-) and a low IP boundary tone (L%), where the ip-final syllables are consistently longer than the IP-final syllables simply due to the lengthening of syllables bearing contour tones.

AP vs. ip vs. IP: Tone interpolation differences

Boundary tones in Bengali are, by definition, associated to edges of phrases; still, their effect is normally felt well before and after the associated boundary. The AP boundary tone (Ha) causes pitch to interpolate smoothly from the preceding pitch accent (normally low L*), generating a relatively constant pitch slope across the AP. Thus we can interpret the AP boundary tone (Ha) as having some sort of influence over the entire AP, and not just the right edge. Similarly, the IP boundary tones (L%, H%, LH%, and HL%) cause pitch to interpolate from the last pitch accent all the way to the right edge of the IP. In the final syllable, the most extreme pitch excursions take place. Examples of gradually rising pitch before the high IP boundary tone (H%) and falling IP boundary tone (HL%) are shown below in Figure 13.



brought CL brought

Figure 13. Near-minimal pairs of the high (H%) and falling (HL%) IP boundary tones, produced by the same speaker. The sentences (*Monoara Romilake nie elo nà* 'Didn't Monoara bring Romila?' and *Monoara Romilake nie elo?* 'Did Monoara bring Romila?') could not form a perfect minimal pair as different sentence-level clitics were needed to elicit the boundary tones. [Fa06, Fa03]

In the case of the high IP boundary tone (H%), the pitch rises more sharply in the final syllable, while in the case of the falling IP boundary tone (HL%), pitch ceases to rise and begins to sharply fall in the final syllable. Similarly, pitch gradually falls before a low IP boundary tone (L%) or rising IP boundary tone (LH%). In the case of the low IP boundary tone (L%), the pitch drops more sharply in the final syllable, while in the case of the rising IP boundary tone (LH%), pitch ceases to fall and begins to sharply rise in the final syllable. Thus, we can conclude that IP boundary tones, like the AP boundary tone, have some influence over a large domain, and can cause pitch interpolation from the preceding tone. This is in sharp contrast to ip boundary tones, which appear to cause no pitch interpolation outside the final one or two syllables of the ip. A good comparison is the rising IP boundary tone (LH%), which would be expected to cause pitch interpolation from the preceding tone, versus the rising ip boundary tone (LH-), which would only cause pitch interpolation in the final one or two syllables of the ip. This contrast is best brought out when the final pitch accent of an ip or IP is several syllables away from the right boundary, as shown in Figure 14 and Figure 15 below.

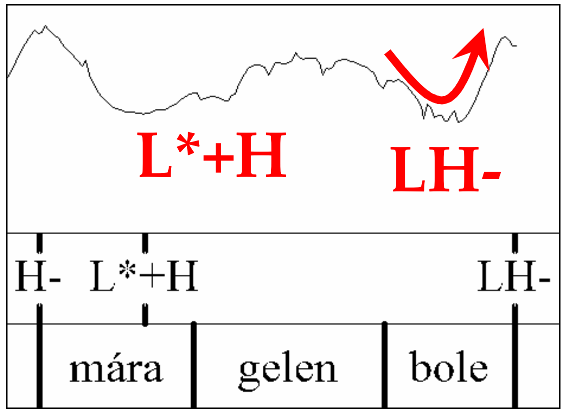
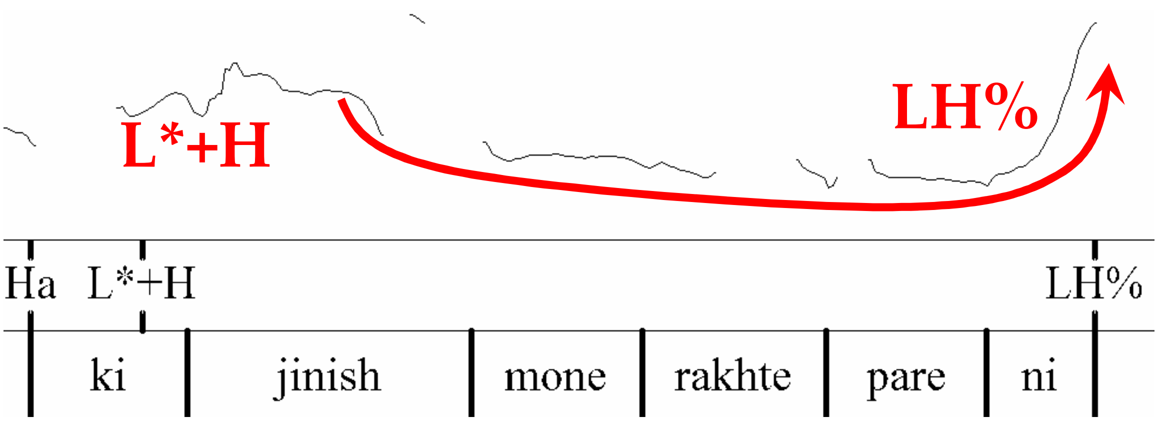


Figure 14. The rising ip boundary tone (LH-) only influences the pitch of the final one or two syllables. Note how the rising/high component of the bitonal pitch accent (L*+H) is realized across several syllables ([ɹa.ge.len]) before the bitonal boundary tone (LH-) takes effect on the penultimate syllable [bo].



what thing remember could not

'What thing could [the queen of Nepal] not remember?'

Figure 15. The rising IP boundary tone (LH%). Note how the low component of the bitonal boundary tone causes the pitch to fall almost immediately after the last pitch accent, several syllables away from the IP-edge. Pitch begins to fall right after the first syllable ([dʒi]) after the pitch-accented syllable ([ki]). [Jh 47]

This suggests that unlike AP and IP boundary tones, ip boundary tones only have the power to cause pitch interpolation within a restricted domain. This property of ip boundary tones in Bengali is referred to in the current model as the *ip boundary tone locality constraint*, as mentioned previously, and helps support the theory that there are three distinct levels of prosodic structure in Bengali: AP, ip, and IP.

AP vs. ip vs. IP: Tone height differences

As shown in Table 2, the pitch after a low pitch accent (L*) rises more to a more extreme height at a high IP boundary tone (H%) than at a high ip boundary tone (H-) [paired $t(4) = 4.57$, $p < 0.05$], which reaches a more extreme height at a high AP boundary tone (Ha) [paired $t(4) = 3.76$, $p < 0.05$]. Because of the optionality in phrasing, it can be difficult to reliably elicit an ip break in a given sentence. Thus, perfectly minimal pairs between Ha (AP break) and H- (ip break) from the same speaker could not be recorded for all speakers. Thus, this measure focuses on six speakers' productions of the three high boundary tone types across the same or similar words. All words measured for the AP- *vs.* ip-final distinction were controlled for the following properties: the words measured were all three syllables in length and each subject's triplet included words found in the same sentential position.[4] The words in the ip- *vs.* IP-final comparison were not controlled for sentential position due to syntactic constraints on their occurrence.

Speaker	**L*…Ha**	**L*…H-**	**L*…H%**
BM	36.3	109.3	147.0
Sh	41.5	55.6	97.7
Do	70.9	108.9	120.3
Ba	8.5	54.9	72.8
Jh	63.4	155.0	184.8
Pi	39.7	86.0	106.3

Table 2. Pitch differences (in Hz) between low pitch accents (L*) and following high boundary tones of different phrase levels (Ha, H-, and H%), measured on identical or similar words across six speakers.

Figure 16 below illustrates the difference in tone height between the high AP boundary tone (Ha) and the high ip boundary tone (H-). Note that the pitch scales are constant across the two examples (170-350Hz).

[4] The name *Monoara* could be analyzed as consisting of four syllables underlyingly /mo.no.a.ɹa/, but in running speech, unstressed non-low vowels tend to become glides in hiatus, yielding [ˈmon.o̯a.ɹa]. Also, *nie elo nà* 'took-*CL*' underlyingly includes five syllables /ni.e.e.lo.na/ but is pronounced with four [ˈni.e.lo.na] in careful speech and three [ˈnie̯.lo.na] in running speech due to hiatus resolution.

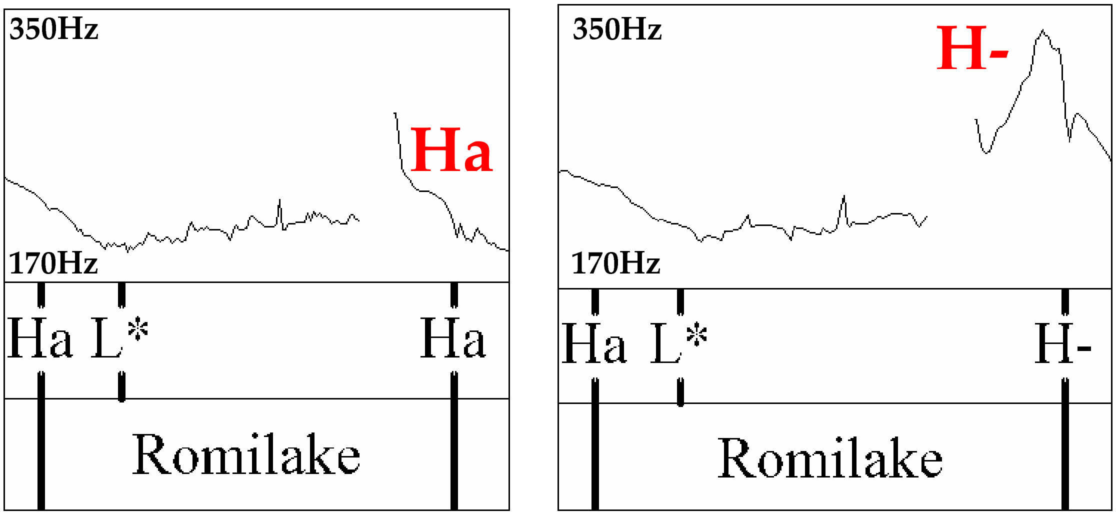


Figure 16. The name *Romilake* ‘Romila-ACC’ produced by the same speaker, once with the high AP boundary tone (Ha) and once with the high ip boundary tone (H-). [BM03, BM01]

Because ip breaks and IP breaks tend to occur in very different syntactic locations, perfect minimal pairs between a high ip boundary tone (H-) and a high IP boundary tone (H%) were not found in the data set. Instead, Figure 17 below shows the contrast between H- and H% within the same speaker but across different words in different sentences. The pitch scale is held constant across the two examples (100-500Hz).

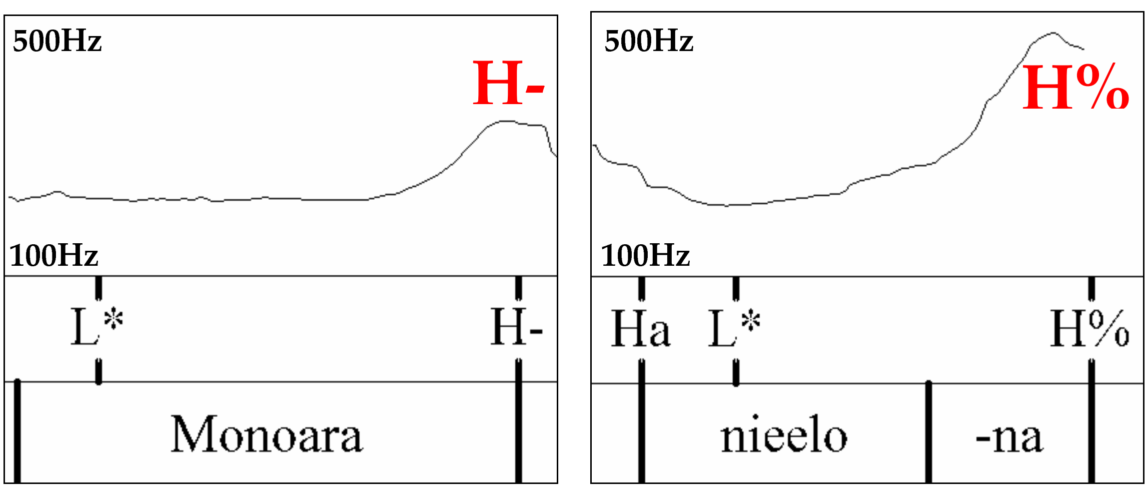


Figure 17. A contrast between the high ip boundary tone (H-) in the name *Monoara* and the high IP boundary tone (H%) on the verb *nie elo nà* ‘brought-*CL*’ produced by the same speaker in two different sentences. [Do01, Do06]

Tone sequences: role of the OCP

Some previous studies (Hayes & Lahiri 1991, Fitzpatrick 2000, Selkirk 2006, among others) explicitly assert that tonal sequences in Bengali are directly affected by the Obligatory Contour Principle, or OCP (Leben 1973, McCarthy 1986). Sequences of identical tones are transformed so that one of the tones is deleted. For example, if two high boundary tones H_I and H_P (the labels used in Hayes & Lahiri 1991 and Lahiri & Fitzpatrick-Cole 1999) occur in sequence, the boundary tone belonging to the smaller phrase (in this case, the H_P) deletes, leaving the sequence indistinguishable from an underlying instance of only one boundary tone. These studies also note the lack of sequences of high pitch accents (H*) followed by high (H_I) or falling (H_IL_I) boundary tones, and the lack of sequences of low pitch accents (L*) followed by low (L_I) or rising (L_IH_I) boundary tones, and attribute these apparent paradigm gaps to the influence of the OCP. This would suggest that any sequence of pitch accents and/or boundary tones would have to alternate between low and high targets.

Data from the current study, however, indicates that this is not the case. The low and rising pitch accents (L* and L*+H, respectively) seem to freely occur even when the following boundary tone (described in more depth in the subsequent section) includes a high element, such as the high or falling IP boundary tones (H% and HL%, respectively). The sequence L*…H% contrasts with the sequence L*+H…H%, and the sequence L*…HL% contrasts with the sequence L*+H…HL%, counter the claims of previous analyses. Similarly, the sequence L*+H…L% contrasts with the sequence L*…L%, counter previous analyses where the latter structure would violate the OCP. This contrast was found in the current study and is illustrated below in Figure 18. Analyses that make use of the OCP would not be able to predict the steadily falling pitch between the two low tones in the example on the left.

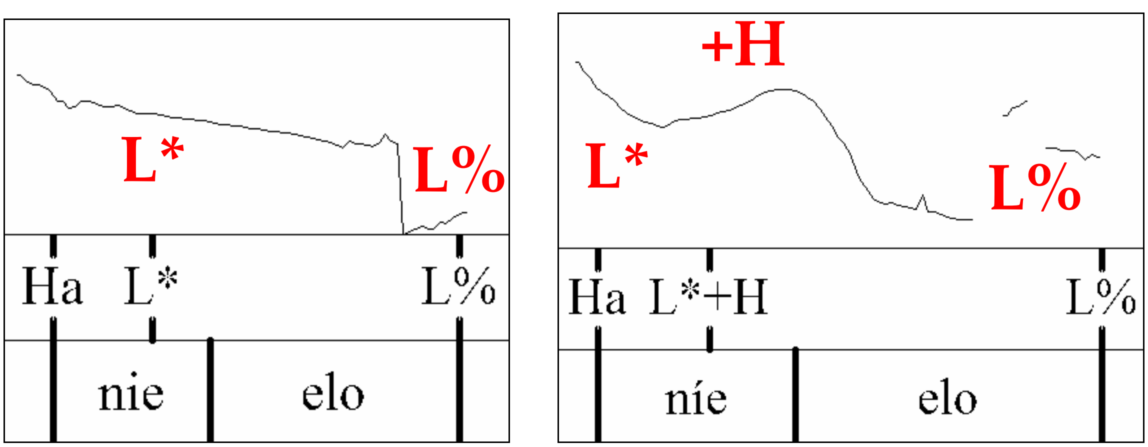


Figure 18. A minimal pair of the sequences L*…L% and L*+H…L% on the string *nie elo* 'brought', extracted from the same sentence *Monoara Romilake nie elo* 'Monoara brought Romila', produced by two different speakers. [Tu09, Fa09]

Similarly, the sequence L*…LH- (the boundary tone will be described in more detail in the subsequent section) contrasts with L*+H LH-, as shown below in Figure 19.

Analyses that make use of the OCP would not be able to predict the flat low pitch in the example on the left.

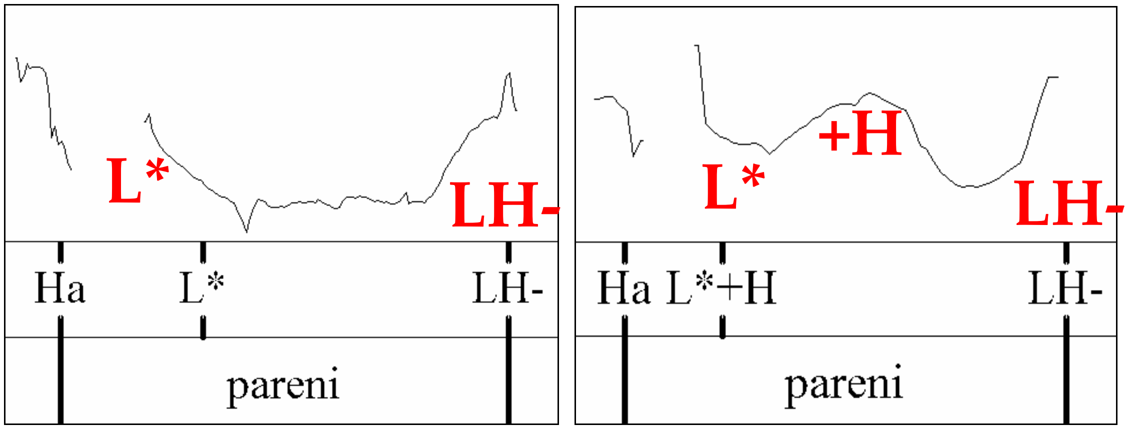


Figure 19. A minimal pair of the sequences L*…LH- and L*+H…LH- in the string *pare ni* 'could not', extracted from the phrase *Rumu jei namgulo mone rakhte pare ni* 'The names which Rumu could not remember…' produced by two different speakers. [Tu52, Re52]

The largely unrestrained inventory of tonal sequences observed in the current study's dataset indicates the lack of an obvious constraint on tonal interaction, suggesting (counter most previous analyses) that the OCP plays no major role in Bengali prosody.

Focus realization

Lack of focus-R-edge boundary tone

Previous studies agree that focus is realized in Bengali with a rising pitch, although the nature of the high target of this pitch rise is controversial. While Hayes & Lahiri (1991) posit a low pitch accent followed by a high boundary tone, Lahiri & Fitzpatrick-Cole (1999) further propose a lexically-specified high tone (H*) on certain focused phrases. Selkirk (2006) describes a floating [H] tone associated with the right edge of focused elements. Whether it is labeled as a boundary tone, a boundary-seeking floating high tone, or a lexically-specified high tone on a word-edge clitic, all of these studies agree that the high tone portion of the focus constituent's rising pitch associates with the right edge of the entire focus constituent, or at least the right edge of some word in the constituent. The boundary-seeking high tone analysis of focus realization has its share of theoretical challenges, including the violation of the Strict Layer Hypothesis (Selkirk 1984, 1986; Nespor & Vogel 1986), the positing of underlying tones in a non-lexical tone language (Lahiri & Fitzpatrick-Cole 1999), and the projection of a right-edge prosodic boundary by a pitch accent associated with left-edge prominence (further explained in Selkirk 2006).[5] The boundary-seeking high tone analysis makes two important claims:

1. The highest pitch in a focused word will occur at the word's right edge.
2. As the focused word increases in length, its pitch maximum will occur farther from the pitch accented syllable.

Thus, examples of focus where the highest pitch is realized well before the right edge of a word are either ignored in the analysis, or described as marginal exceptions. The data from the current study, however, include numerous such examples, like those shown below in Figure 20, which compares shorter and longer focused constituents produced by the same speaker in the same carrier sentence *Monoara* ____ *nie elo* 'Monoara brought ____'.

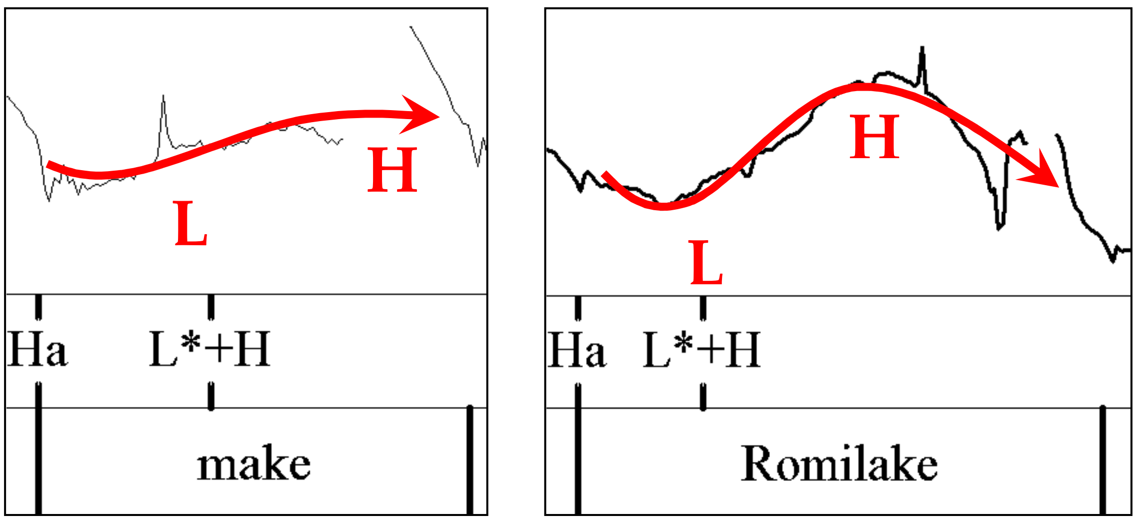


Figure 20. A comparison of shorter (*make* 'mother-ACC') and longer (*Romilake* 'Romila-ACC') focused words, produced by the same speaker in the same carrier sentence *Monoara* ____ *nie elo* 'Monoara brought ____'. [BM20, BM23]

Figure 20 illustrates how very short words such as *make* 'mother-ACC' can appear as though they bear the default combination of a low pitch accent (L*) and high boundary tone (Ha). However, when the constituent is longer, as in *Romilake* 'Romila-ACC', there is a clear fall in pitch between the pitch maximum between the second and third syllables [mi.la] and the right edge of the focused constituent, suggesting the lack of a high boundary-seeking tone. Thus the current study adopts the analysis in Michaels & Nelson (2004), which attributes the rise in pitch to a rising pitch accent (L*+H) on the focused word. In such an analysis, the rise in pitch is not associated to a boundary at all. This association of the focus rising pitch to a pitch accent as opposed to a boundary-seeking tone carries the following implications:

[5] In the interest of space, I will not expand on the theoretical problems associated with the boundary-seeking high tone analysis of Bengali focus realization. For a more detailed explanation, see Selkirk (2006) and Khan (forthcoming).

1. The highest pitch in a focused word will occur at a relatively fixed distance from the pitch accented syllable.

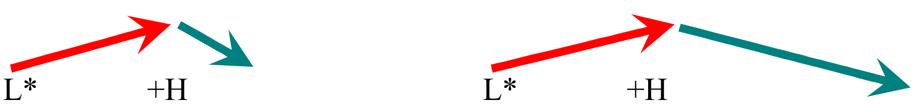
2. As the focused word increases in length, its pitch maximum will occur farther from the word's right edge.

The two predictions listed above are repeated below in four schematic representations of the rise in pitch in shorter and longer focused constituents predicted by the boundary-seeking H tone analysis (Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999, Selkirk 2006) and by the bitonal pitch accent analysis (Michaels & Nelson 2004, current study). Below each schematic representation is a summary of the predictions each analysis makes with respect to whether the duration of rise or fall in pitch will be shorter or longer in smaller or larger focused constituents, and whether the percentage of the duration of rise or fall in pitch will be greater or smaller in smaller or larger focused constituents.

Predictions of the boundary-seeking H tone analysis

	L* H	L* H
	[Short constituent]	[Significantly longer constituent]
Pitch rise:	**shorter**	**longer**
Pitch fall:	**same (negligible)**	**same (negligible)**
% Pitch rise:	**same (~100%)**	**same (~100%)**
% Pitch fall:	**same (negligible)**	**same (negligible)**

Predictions of the bitonal pitch accent analysis



	L* +H	L* +H
	[Short constituent]	[Significantly longer constituent]
Pitch rise:	**same**	**same**
Pitch fall:	**shorter**	**longer**
% Pitch rise:	**larger %**	**smaller %**
% Pitch fall:	**smaller %**	**larger %**

Thus, the bitonal pitch accent analysis predicts that the rising pitch towards the +H will constitute a larger percentage of the total duration of a shorter constituent than of a longer constituent, and that the falling pitch from the +H to the word's right edge will constitute a smaller percentage of a shorter constituent than of a longer constituent. To test these predictions, 25 subjects were asked to produce sentences eliciting corrective focus on constituents of various lengths, ranging from two syllables (*make* 'mother-ACC') to five (*Linamamike* 'Aunt Lina-ACC'). Corrective focus was elicited by having the subject read a mini-dialog in which hypothetical Speaker A makes an incorrect statement such as *Monoara Rumuke nie elo, tai nà?* 'Monoara brought Romila, right?', and Speaker B responds with *Na, na. Monoara* ____ *nie elo* 'No, no. Monoara brought ____'. The blank left in the last sentence in the minidialog (*Monoara* ____ *nie elo* 'Monoara brought ____') was filled in with the focused object of controlled length. Then a native speaker consultant listened to all 100 recordings and labeled some as infelicitous if she felt the subject did not properly produce the expected corrective focus pattern. Due to the awkward nature of eliciting focus in a reading task, many subjects had one or two examples judged infelicitous, leaving only eight subjects whose data were considered perfect. Two durational measures were taken for each of the 32 focused words examined:

(a) **% Duration of pitch rise:** the duration of rising pitch from the pitch minimum (interpreted as the locus of pitch accent assignment in either analysis of focus realization) to the pitch maximum, as a percentage of the total duration of the word, and
(b) **% Duration of pitch fall:** the duration of falling pitch from the word's pitch maximum to the right edge, as a percentage of the total duration of the word.

Mean rise and fall durational measurements are shown below in Table 3, shown as a percentage of the duration of the entire word:

Number of syllables	**Word**	**Mean duration (as % of total duration)**	
		% Pitch rise L* → H	**% Pitch fall H → R-edge**
2	*make* 'mother-ACC'	58.9%	12.8%
3	*Ninake* 'Nina-ACC'	45.2%	34.0%
4	*Romilake* 'Romila-ACC'	43.8%	41.0%
5	*Lina mamike* 'Aunt Lina-ACC'	35.0%	54.2%

Table 3. Means of the duration of rising and falling pitch in focused words of varying length, shown as a percentage of the duration of the entire word. Note that the percentages will not add up to 100% as these durations do not include any initial fall in pitch towards the pitch minimum (*i.e.* the L* portion of L*+H).

The values in Table 3 above are plotted as a line graph in Figure 21 below.

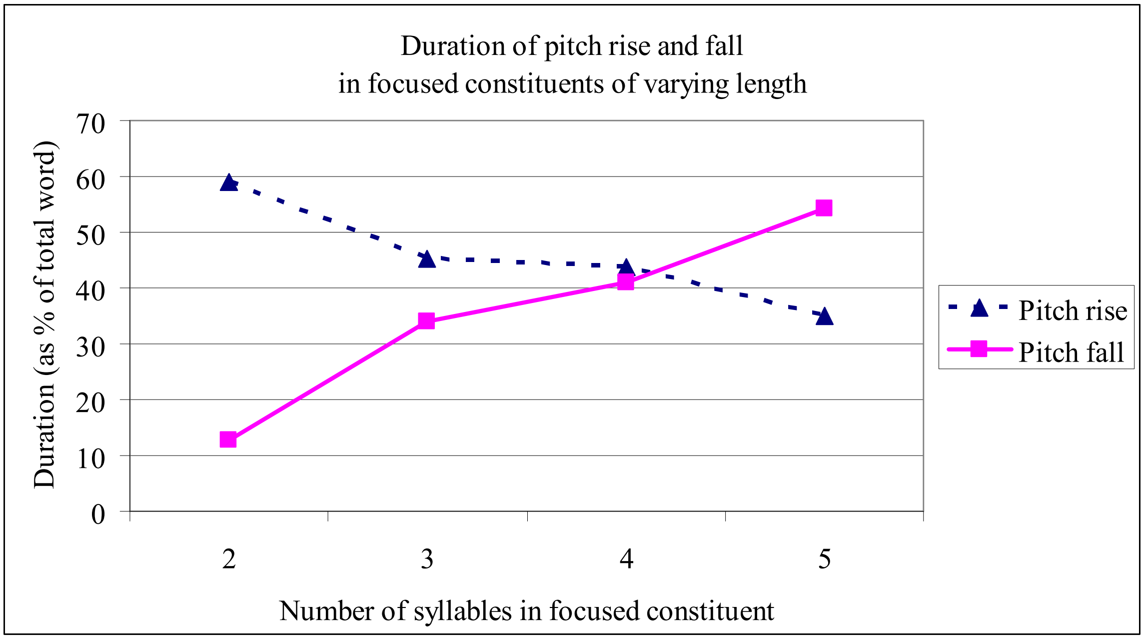
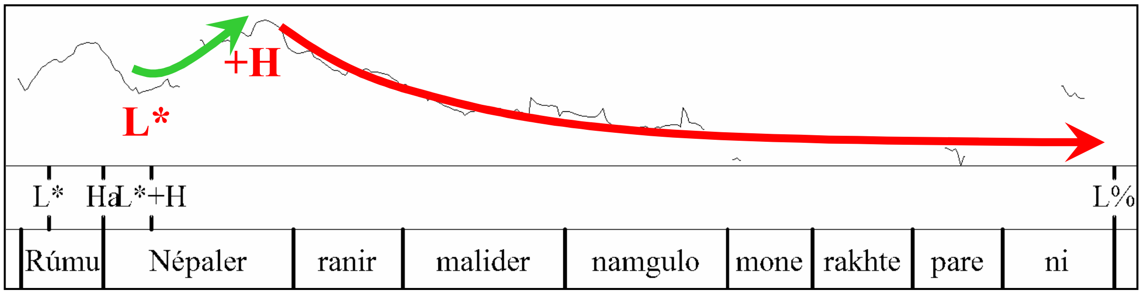


Figure 21. Duration of pitch rise and pitch fall in focused constituents of varying length (2, 3, 4, or 5 syllables), shown as a percentage of the total constituent duration.

From the measurements summarized in Table 3 and illustrated in Figure 21, it is clear that the duration of pitch rise as a percentage of the total duration of the focused constituent *decreases* as the constituent's size increases, while the duration of pitch fall as a percentage of the total duration of the focused constituent *increases* as the constituent's size increases. Thus, the high tone of the focused word is somehow anchored to the pitch accent rather than the word's right edge, supporting the analysis that the rising pitch seen on focused constituents is due to rising pitch accent (L*+H) and not to a boundary tone such as Hayes & Lahiri's (1991) H_P or a boundary-seeking tone such as Lahiri & Fitzpatrick Cole's (1999) H* or Selkirk's (2006) $[H]_{FOC}$.

Deaccenting/dephrasing

By positing that focused constituents bear a bitonal pitch accent (L*+H) and do not project a local high AP boundary tone (Ha), post-focal deaccenting/dephrasing is also explained. Post-focal words bear no tones, and are instead realized with smooth pitch interpolation towards the final boundary tone (deaccenting/dephrasing), as shown in Figure 22.



Rumu *Nepal's* queen's the gardeners' the names remember could not

'Rumu couldn't remember the names of the gardeners of the queen of *Nepal*.'

Figure 22. The focused word *Népaler* 'Nepal's' bears L*+H. Post-focal words bear no pitch accents or boundary tones; their pitch is determined by phonetic interpolation of adjacent tones. [Fa38]

If there were a high boundary tone (Ha) immediately following the focused constituent *Nepaler* 'of Nepal' in Figure 22, all the unaccented material between *Nepaler* 'of Nepal' and the end of the sentence would be prosodically unparsed material, which would lead to a violation of the Strict Layer Hypothesis. Instead, the current analysis assumes that the right edge of the focused constituent's AP is actually at the end of the sentence, thus coinciding with the low IP boundary tone (L%), which overrides it (as described earlier). Thus, the final AP in this sentence is *Nepaler ranir malider namgulo mone rakhte pare ni* 'couldn't remember the names of the gardeners of the queen of *Nepal'*. This concurs with the theory that each AP contains one and only one pitch accent, and that each pitch accent projects one and only one AP.

Conclusions

In using varied data from numerous speakers, the current study aims to serve as the foundation of an Intonational Phonology model of Bengali prosody (*i.e.* B-ToBI). This newly-acquired acoustic data suggest that Bengali is among the few languages, including Kiche (Nielsen 2005), Basque (Hualde 1988, 1999; Jun 1999, 2005), and Farsi (Jun 2005, Esposito & Barjam 2007, Scarborough 2007), described as tonally marking three prosodic phrases (IP, ip, and AP). The current study also reveals a larger tonal inventory than described in the literature, including bitonal pitch accents, previously thought to be nonexistent or even impossible in the language. Finally, acoustic measurements of the current study's observed data confirm the implications of the bitonal pitch accent analysis proposed in Michaels & Nelson (2004) and the current study, which posit a bitonal pitch accent and no local boundary tone on focused constituents, thus avoiding the problematic assertions made in previous analyses of Bengali focus. Further investigation, including perception data testing the psychological reality of the structures posited in this analysis, should reveal more details about the prosodic system of Bengali.

References

Arvaniti, Amalia & Mary Baltazani (2000). 'GREEK ToBI: a system for the annotation of Greek speech corpora.' In Maria Gavrilidou, George Carayannis, Stella Markantonatou, Stelios Piperidis, Gregory Stainhaouer (eds.), *Proceedings of Second International Conference on Language Resources and Evaluation (LREC 2000)*: **II**: 555-562. European Language Resources Association: Paris.

Arvaniti, Amalia & Mary Baltazani (2005). 'Intonational analysis and prosodic annotation of Greek spoken corpora.' In Sun-Ah Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press: 84-117.

Beckman, Mary & Julia Hirschberg (1994) 'The ToBI Annotation Conventions'. Manuscript, Ohio State University.

Beckman, Mary & Janet Pierrehumbert (1986). 'Intonational structure in English and Japanese'. *Phonology Yearbook* **3**, 255-310.

Boersma, Paul & D. Weenink (2005). Praat: doing phonetics by computer, v.4.3. Summer Institute of Linguistics. Software freely available.

Elordieta, Gorka (1998). 'Accent, tone and intonation in Lekeitio Basque.' In F. Martínez-Gil and A. Morales-Front (eds.) *Issues in the phonology and morphology of the major Iberian languages.* Washington: Georgetown University Press: 3-78.

Esposito, Christina M. & Patrick Barjam (2007). 'The Intonation of Questions in Farsi – Wh-Questions, Yes-No Questions and Echo Questions'. *UCLA Working Papers in Phonetics* **105**.

Fitzpatrick, Jennifer (2000). 'On intonational typology.' In Peter Siemund (ed.) *Methodological Issues in Language Typology: Sprachtypologie und Universalienforschung* **53**: 88-96.

Hayes, Bruce & Aditi Lahiri (1991). 'Bengali intonational phonology.' *Natural Language and Linguistic Theory* **9**, 47-96.

Hualde, José Ignacio (1988). 'A theory of pitch-accent, with particular attention to Basque.' *Euskara Biltzarra/Congreso de la Lengua Vasca* **1**: 53-60.

Hualde, José Ignacio (1999). 'Basque accentuation.' In H. van der Hulst, (ed.) *Word prosodic systems in the languages of Europe*. Berlin: Mouton de Gruyter: 947-993.

Jun, Sun-Ah (1993). *The Phonetics and Phonology of Korean Prosody*. Ph.D. Ohio State University.

Jun, Sun-Ah (1998). 'The Accentual Phrase in the Korean prosodic hierarchy.' *Phonology* **15**(2), 189-226.

Jun, Sun-Ah (1999). 'K-ToBI (Korean ToBI) Labeling Conventions.' In *Speech Sciences* **7**: 143-170.

Jun, Sun-Ah (ed.) (2005). *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press.

Khan, Sameer ud Dowla (2006). 'Bengali intonational structure: Evidence for three levels of phrasing'. Poster presented at the 4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu.

Khan, Sameer ud Dowla (2007). 'A model of Bengali prosody: the foundations of B-ToBI.' UCLA Ms.

Khan, Sameer ud Dowla (forthcoming). 'The distribution and realization of H tones in Bengali.' *UCLA Working Papers in Phonetics* **106**.

Ladd, D. Robert (1996). Intonational Phonology. New York: Cambridge University Press.

Lahiri, Aditi & Jennifer Fitzpatrick-Cole (1999). 'Emphatic Clitics and Focus Intonation in Bengali.' In *Phrasal Phonology*, eds. René Kager & Wim Zonneveld, 119-144. Nijmegen: University of Nijmegen Press.

Leben, William Ronald (1973). 'Suprasegmental Phonology'. Doctoral dissertation, MIT.

Mahjani, Behzad (2003). *An instrumental study of prosodic features and intonation in Modern Farsi (Persian)*. Masters thesis, Department of Theoretical and Applied Linguistics, University of Edinburgh.

McCarthy, John J. (1986). 'OCP Effects: Gemination and antigemination'. *Linguistic Inquiry* **17**: 207-263.

Michaels, Jennifer Marie & Catherine E. Nelson (2004). 'A Preliminary Investigation of Intonation in East Bengali'. UCLA Ms.

Nespor, Marina & Irene Vogel (1986). *Prosodic Phonology*. Foris: Dordrecht.

Nielsen, Kuniko (2005). 'Kiche Intonation.' *UCLA Working Papers in Phonetics* **104**: 45-60.

Pierrehumbert, Janet (1980). 'The phonology and phonetics of English intonation'. Ph.D. thesis, MIT, published 1988 by IULC.

Pierrehumbert, Janet & Mary Beckman (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.

Scarborough, Rebecca (2007). 'The Intonation of Focus in Farsi'. *UCLA Working Papers in Phonetics* **105**.

Selkirk, Elisabeth O. (1984). 'On the Major Class Features and Syllable Theory'. M. Aronoff & R. Oehrle (eds.) *Language Sound Structure*, Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Selkirk, Elisabeth O. (1986). 'On derived domains in sentence phonology'. *Phonology Yearbook* **3**: 371–405.

Selkirk, Elisabeth O. (2006). 'Bengali intonation revisited: An optimality theoretic analysis in which FOCUS stress prominence drives FOCUS phrasing.' In *Topic and Focus: Intonation and Meaning*, eds. Chung-Min Lee, Matthew Gordon, & Daniel Büring, 215-244. The Netherlands.

Sjölander, K. & J. Beskow (2005). WaveSurfer 1.8.5. Kungliga Tekniska högskolan: Software freely available.