UNIVERSITY OF CALIFORNIA

Los Angeles

Intonational Phonology and Focus Prosody of Bengali

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Linguistics

by

Sameer ud Dowla Khan

2008

The dissertation of Sameer ud Dowla Khan is approved.

Bruce P. Hayes

Patricia Keating

Jody Kreiman

Kie Ross Zuraw

Sun-Ah Jun, Committee Chair

University of California, Los Angeles

2008

TABLE OF CONTENTS

LIST OF FIGURES

the focused object [ɹomilake] ‘Romila-ACC’, and the high IP boundary tone (H%) assigned to confirmation questions. The tonal makeup of the focused word is ambiguous, and in fact could be composed of a focused smooth rise (L*…fHa), or a hybrid of the two sequences (L*+H…fHa). In any case, however, the H targets would be adjacent. [Re06] 303

LIST OF TABLES

ACKNOWLEDGMENTS

Completing this dissertation would not have been possible without the time and expertise of the many people who helped me at every step; I owe sincere gratitude to all of those who supported and guided me.

First, I would like to thank my primary adviser for this project, Professor Sun-Ah Jun, whose profound understanding of the phonology of intonational systems was an invaluable asset in helping to construct the prosodic model presented here. Indeed, it was in taking Sun-Ah's course on intonation that first inspired the project that led to this dissertation. Furthermore, her genuine interest in the subject served as a great source of encouragement for my progress. Sun-Ah's firm belief that a theory is only as strong as the data it can capture gave me the confidence to stand by my own analysis of my data instead of blindly defending a more fashionable theory. For her guidance and support, I cannot thank Sun-Ah enough.

I also owe a great deal to my committee members—Professors Bruce Hayes, Patricia Keating, Jody Kreiman, and Kie Ross Zuraw—for their constructive input and support. With his well-known expertise in the study of Bengali intonation, Bruce served as an important guide in carefully analyzing the patterns found in my data. Sharing his insight and even his data from his study of Kolkata Bengali, Bruce gave me a much deeper understanding of the intonational system of the language than I could have had

from reading his published work alone. As leaders in the fields of segmental phonetics and phonology, Pat and Kie exposed me to a wider range of perspectives of my study. Kie, in particular, was immensely helpful in shaping the presentation of my analysis in such a way that it could be appreciated both by experts and by readers less familiar with the phonological study of prosody. And beyond her help in regards to this dissertation, Kie is also responsible for forcing me to believe in myself as a linguist. Kie, you saw me at my most vulnerable time in graduate school, and I know that if it weren't for your wisdom and encouragement, I would not be where I am now. And although not directly involved with my dissertation project, I owe so much to Professors Susan Curtiss, Pamela Munro, and Colin Wilson, who have inspired me throughout my undergraduate and graduate career at UCLA.

I would also like to thank all of the subjects of my experiments, for their time and patience. ব্যস্ততার মধ্যেও আপনারা সবাই আমাকে সাহায্য করেছেন বলে আমি আন্তরিকভাবে কৃতজ্ঞ।

I don't know how many times I threatened to drop out of graduate school. If it weren't for my close friends in linguistics, I surely would have never finished my degree. I have so much respect and love for the senior students who guided me as mentors through my time at UCLA; Tim Arbisi-Kelm, Christina Esposito, Jeff Heinz, Kuniko Nielsen, Manola Salustri, and Shabnam Shademan—thank you for always inspiring me and instilling the confidence I needed to continue. To the colleagues in my cohort—Lawrence Cheung, Christina Kim, Tuan Le, Leonard Chacha Mwita, Reiko Okabe, Molly

Shilman, and Lauren Varner—I miss our study nights, when all we would do was find more and more clever ways to avoid working. Chris and Lauren, your friendship has been so precious to me, and even though we now live far apart, I feel so close to you both when we see each other. Aaron Lee and Molly Shilman, you have not only been there for me as two of my dearest friends, but you have also been willing to help guide me with my work, day or night. For all the advice you two have given me over the years, I feel your names belong alongside those of my committee members. And to all my other friends at UCLA's Linguistics Department, including but not limited to Asia Furmanska, Ananda Lima, Jennifer Michaels, J'aime Roemer, Katie Tang, and Kristine Yu, thanks for always being there for me.

Thank you Lisa Harrington, Mandana Hashemzadeh, Melanie Levin, and Kathryn Roberts, not only for doing everything that was needed to get me through the linguistics program, but also making it such a pleasure at the same time.

To my roommate and very dear friend Raymond Lo, thank you for putting up with my temper and all my quirks; I'm sure that I have never been the ideal roommate, and the added pressure of my dissertation could not have helped. Thank you for being so considerate and forgiving of me all these years.

Thank you to my friends who kept me alive and healthy by distracting from my work: Alex Abad-Santos, Matin Ebneshahrashoob, Asha Goud, Ashish Gupta, Jennifer

Paul, Michelle Sassounian, and many others. Eric Fiala, Nicole Gfroerer, Tatiana Libman, Ananda Lima, Robyn Orfitelli, Diego Rosso, Dan Shiber, Chad Vicenik, and Yulia Zabiyaka, thanks for hosting so many enjoyable dinners at your homes, simultaneously preventing me from starving to death and from becoming a complete recluse. Lotta Andonian, Dorothy Lê, and Amie McCampbell, thanks for making sure I got out the apartment several times a week to make it to the market for fresh produce. Justin Nuger, thank you for staying up with me virtually every night for the past five years, talking me through all of my difficult times and helping me make the decisions I was often too afraid to make on my own. Your opinion and your friendship both mean so much to me. Billy Kernkamp, thank you for listening to all my troubles even when I knew you were going through many of your own, and thank you for always offering to beat people up on my behalf. Matthew Boese, thanks for accompanying me on so many cycling and metro adventures and for having me accompany you on your journey across North America, forcing me away from all my stresses at the times that I most needed a break. And without Daisy Kwoh and Eric Fiala, my body would probably be at least ten pounds heavier and riddled with arterial blockages; thanks for keeping me running and cycling through graduate school.

To Nipa Auntie and all the teachers at the Bangla School of Los Angeles, thank you for not only helping me with my dissertation work, but also taking the risk of giving a second-generation Bangladeshi American the responsibility of teaching children about the Bengali language. My experiences at your school, as well as tutoring Nadia Khawaja,

J'aime Roemer, and Dallas Swendeman, have given me new insights into the language of my family. Thank you all for trusting me to teach you about a language I feel like I'm still trying to learn.

To John Chaplow, Jon Walden, James Hernandez, Chris Cheng, Simon Li, Alex Paris, Peter Sanderson, Michael Gorospe, Alexander Yoo, Raul Borja, and all the guys at GLU, thank you for making me part of your loving community, especially in such a busy year for me, when I thought I'd certainly never meet new people. It is an honor to be a part of something so beautiful.

To my father, Shafiq ud Dowla Khan, who left his home decades ago to give me and my sister Naveen an opportunity to pursue all of our academic and professional goals, and to my mother, Farida Amin Khan, who not only supported me as a parent but also guided me as my primary consultant through this and many other projects, I am forever indebted. As a family, we have gone through many challenges, and I must tell you I am so appreciative of how much you have done for me and Naveen, even when I know it could not have been easy. আমি যখন চিন্তা করি, তোমরা দুজন আমাদের জন্য কত কষ্ট করেছো, সত্যি আমার চোখে পানি আসে। মা গো, তুমি না থাকলে আমি কী করতাম? আমি যখনই তোমাকে ফোন করেছি, তুমি সাথে সাথে আমাকে ঘণ্টার পর ঘণ্টা সাহায্য করেছো। অন্য কোনও মা তার ছেলের জন্য এত করত বলে আমার মনে হয় না। আমি খুবই ভাগ্যবান তোমার মত মা পেয়েছি।

To all of my aunts, uncles, cousins, nieces, and nephews, thank you for always making me feel welcome and loved no matter how bad I've been about keeping in touch, especially in these last few years when I have been so distracted. It is rare to be able to travel almost anywhere in the world and feel so much at home.

And to my little sister Naveen, who I am so proud of, thank you for always being there for me and for always being so willing to accept me for who I am. তোমাকে যে কত ভালবাসি, নাভু!

VITA

June 17, 1982	Born, Newport Beach, California
2001-2003	Teaching Assistant Children's Center University of California, Irvine
2003	B.A., Linguistics University of California, Los Angeles Summa cum Laude
2005	Teaching Assistant Department of Linguistics University of California, Los Angeles
2006	M.A., Linguistics University of California, Los Angeles
2006-2007	Teaching Associate Department of Linguistics University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

Khan, S. D. (2004). 'Speaking of German...: exploring the dialects of everyone's favorite 'language''. Let's Go: Germany. New York: St. Martin's, p. 60.

Esposito, C. M., Khan, S. D., and Hurst, A. (2005). 'Breathiness in Indic Languages'. Poster presented at the 149th meeting of the Acoustical Society of America, Vancouver, British Columbia, 17 May.

Esposito, C. M., Khan, S. D., and Hurst, A. (2005). 'Breathy Nasals and /Nh/ Clusters in Bengali, Hindi, and Marathi'. Working Papers in Phonetics 104, pp. 82-106. Los Angeles: University of California.

Khan, S. D. (2006). 'Similarity Avoidance in Bengali Fixed-Segment Reduplication'. Poster presented at the 14th Manchester Phonology Meeting, Manchester, United Kingdom, 27 May.

Khan, S. D. (2006). 'Similarity Avoidance in East Bengali Fixed-Segment Reduplication'. Talk presented at the 34th Western Conference on Linguistics, Fresno, California, 27 October.

Khan, S. D. (2006). 'Bengali intonational structure: Evidence for three levels of structure'. Poster presented at the 4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii, 29 November.

Khan, S. D. (2007). 'The Distribution and Realization of H tones in Bengali'. Working Papers in Phonetics 106, pp. 34-40. Los Angeles: University of California.

Esposito, C. M., Khan, S. D., and Hurst, A. (2007). 'Breathy Nasals and /Nh/ Clusters in Bengali, Hindi, and Marathi'. In Indian Linguistics 68 (3-4). Pune: Linguistic Society of India.

Khan, S. D. (2007). 'Phrasing and Focus in Bengali'. Poster presented at the International Congress of Phonetic Sciences Satellite Workshop on Intonational Phonology: Understudied or Fieldwork Languages, Saarbrücken, Germany, 5 August.

Khan, S. D. (2007). ‘Phrasing and Focus in Bengali’. Workshop paper for the International Congress of Phonetic Sciences Satellite Workshop on Intonational Phonology: Understudied or Fieldwork Languages.

Khan, S. D. (2007). ‘Similarity Avoidance in East Bengali Fixed-Segment Reduplication’. In: Bainbridge, Erin & Brian Agbayani (eds.) Proceedings of the thirty-fourth Western Conference on Linguistics 2006, pp. 257-271. Fresno: California State University.

ABSTRACT OF THE DISSERTATION

Intonational Phonology and

Focus Prosody of Bengali

by

Sameer ud Dowla Khan

Doctor of Philosophy in Linguistics

University of California, Los Angeles, 2008

Professor Sun-Ah Jun, Chair

Using a framework based on the autosegmental-metrical model of intonational phonology and the transcription system known as Tones and Break Indices, or ToBI, the current study analyzes data from 20 speakers of Bangladeshi Standard Bengali to construct a comprehensive model of intonational phonology and a prosodic transcription system for Bengali (*i.e.* B-ToBI). The current model finds that Bangladeshi Standard Bengali has a rich tonal inventory, including pitch accents—associated to metrically-prominent positions of the sentence—and boundary tones—associated to the edges of three prosodic units: the accentual

phrase, the intermediate phrase, and the intonation phrase. As phonological units, these tones interact with one another and undergo various changes due to phonological constraints on the intonational contour. Furthermore, the current study finds that focus is prosodically marked through the use of a special high tone, which can surface in one of three realization patterns depending on the type of focus, as well as through distinctive phrasing patterns. The current model incorporates data collected from speakers with a wide range of dialect influences, and is thus also able to capture much of the patterns observed in previous studies of other dialects of the language, while also providing a large corpus of new data for future research in Bengali prosody.

CHAPTER ONE
INTRODUCTION

For nearly a century, the prosody—and more specifically, the intonational system—of the Bengali language has been described and analyzed in a variety of theoretical frameworks (Chatterji 1921; Ferguson & Chowdhury 1960; Ray, Hai, & Ray 1966, Hayes & Lahiri 1991; Fitzpatrick-Cole 1994; Lahiri & Fitzpatrick-Cole 1999; Truckenbrodt 2003; Michaels & Nelson 2004; Jun 2005; Selkirk 2006). Still, these previous studies have shied away from proposing a comprehensive phonological model of various structures and patterns using data collected from multiple speakers, and supported by F_0 and durational measurements. Instead, most studies attempted to transcribe intonational contours without performing an analysis of the patterns, analyzed data from a small number of speakers, reanalyzed and reclassified data introduced in other studies, or applied the known data to other subfields of linguistics. Due to the lack of a full-scale model of Bengali intonation and the lack of consensus among linguists regarding some of the most basic aspects of Bengali prosody, the research questions addressed in the current study range from basic phonetic measurements to matters of prosodic structure and phonological constraints. Each of the questions below will be addressed in the form of hypotheses tested against experimental data, and comparisons will be made with previous analyses where appropriate.

(1) What is the inventory of distinct pitch accents and boundary tones?

(2) How many levels of tonally-marked prosodic units are there?

(3) How much influence do phonological constraints have over intonation?

(4) What is the proper representation of focus realization, and of post-focal material?

Using a framework based on the autosegmental-metrical model of intonational phonology (Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988, Ladd 1996; see §2) and the transcription system known as Tones and Break Indices, or ToBI (Beckman & Ayers 1994, 1997; see §3), the current study offers a comprehensive model of intonational phonology and a prosodic transcription system for Bengali (*i.e.* B-ToBI), highlighting the unique and theoretically-relevant findings in order to better understand the typology of intonational systems and the interaction of prosody, syntax, and semantics. The specifics of the current model are based on data collected in a series of experiments introduced in Chapter Three.

1 INTRODUCTION TO THE STUDY OF INTONATION

In its most basic definition, the study of intonation involves postlexical, linguistically meaningful changes in pitch—the psycholinguistic reflex of the fundamental frequency of a speaker's voice—during the course of an utterance (Ladd 1996). While it is widely known that pitch can be part of a word's lexical entry in many, if not most, of the world's languages (*i.e.* tone languages like Mandarin or Xhosa), intonation is only concerned with the *postlexical* use of pitch, both in languages with lexically-defined pitch (*e.g.*

Xhosa, Japanese) and those without (*e.g.* English, Arabic). While a person's pitch can rise and fall due to nonlinguistic factors such as fatigue, emotion, or affect, intonation only covers *linguistic* uses of pitch, such as the marking of a yes/no question, or the marking of a focused constituent. In addition to variances in pitch, most current analyses of intonation also cover a wide variety of non-tonal prosodic phenomena, including metrics – the relative prominence of syllables in a word or words in a phrase – and phrasing – the grouping of words and phrases, and the durational phenomena related to a segment's position in a word or a phrase.

Much of the research on English intonation originates in the "American structuralist school" (*e.g.* Trager & Smith 1951) and the "British school" of intonational modeling (*e.g.* Halliday 1967). These schools based their models of pitch contours on impressionistic data, later shown to be phonetically inaccurate (Lieberman 1965). British models broke down pitch contours in an utterance into two groups of contours: heads and nuclei. Heads are simply prenuclear phrases, while nuclei are phrases that can bear sentence stress and one of many possible tonal contours, which are described as a whole. Some of the nuclear contours described for English include "falling-rising" and "low fall-rise". Often, each nuclear pitch contour is associated with one or more semantic uses, such as "calling" or "yes/no question" for the rising contour. In an attempt to describe the pitch contour in a very fastidious and language-non-specific framework, those in the American structuralist school and the British school of intonational description made an effort to describe and model precise details of the pitch contour, often leading to very

extensive and complicated pitch contour inventories and rules. This tradition was also carried over into early descriptions of Bengali intonation and phrasing, described in Chapter Two.

While narrow, language-non-specific transcriptions of intonation are still available, such as Hirst & Di Cristo's (1998) International Transcription System for Intonation (INTSINT), most linguists currently studying intonation have been developing and using broad transcriptions – those that only record the categorical aspects of an individual language's intonational system. As with a broad transcription of segments, this broad transcription of intonation is only concerned with the phonological/contrastive features needed to derive the phonetic output of an utterance in a given language. This implies that much of the observed phenomena, including variation due to differences in people's natural pitch range, emotional state, and other predictable factors, are not to be labeled or analyzed. The view of intonation as an interplay between contrastive phonological units and additional phonetic (*e.g.* realization of tones with respect to prosodic alignment), phonological (*e.g.* constraints on tonal interaction), and paralinguistic (*e.g.* emotion) factors breaks down the pitch contour of an utterance into discrete tonal targets and the trajectories connecting those targets – concepts that eventually came to be incorporated into the Autosegmental-Metrical model of intonation, or Intonational Phonology.

2 INTRODUCTION TO THE AUTOSEGMENTAL-METRICAL (AM) MODEL OF INTONATIONAL PHONOLOGY

The approach taken in this and other recent studies of Bengali intonation is the Autosegmental-Metrical (AM) model of intonation, also known as Intonational Phonology (Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988; see Ladd 1996 for review), which reduces the pitch contour to a string of pitch targets. While previous descriptions of English (*e.g.* Trager & Smith 1951) described four or more contrastive levels of pitch, transcribed with numbers representing their relative height, this inventory was reduced to only two opposing tonal targets in Pierrehumbert's (1980) analysis of English: High (H) and Low (L). The tonal targets are aligned to designated parts of the segmental string, and all intervening pitch is derived simply by phonetic interpolation between the targets. Complex contours are analyzed as combinations of these two tonal targets (*e.g.* Rising LH, Falling HL). All variation from these two targets is explained as effects of paralinguistic factors (*e.g.* surprise), phonological interactions of tones (*e.g.* identity avoidance effects), phonetic realization rules (*e.g.* tone undershoot), or semantic or pragmatic factors (*e.g.* emphasis).

Tonal targets are not static points; their positions do not correspond to fixed levels in the speaker's overall pitch range, but to *relative* levels of H and L, affected by many factors including but not limited to adjacent tones, prosodic alignment, nearby phrase boundaries, and prominence relations of words in a phrase, as well as the paralinguistic factors described earlier. With all these factors, it is not surprising if, in the same

utterance, a tone labeled L is actually realized with a higher pitch than a tone labeled H – in fact, this phenomenon is widely seen in the context known as downstep (Ladd 1983). And while tonal targets are placed in linguistically-defined locations throughout the utterance, the "spaces" between the targets are "filled in" by simply interpolating the pitch of the targets in language-specific ways. Thus, if a hypothetical phrase *tobásari nemo* bears a sequence of, say, a high (H) tone on the syllable *ba* and a low (L) tone on the syllable *mo*, its pitch contour will probably be realized as something like a gradual falling tone across the intervening three syllables *sa*, *ri*, and *ne*.

In the AM model of intonational phonology, a very important distinction is made between two different kinds of tones: head-marking tones (*i.e. pitch accents*) and edge-marking tones (*i.e. boundary tones*). Head-marking tones are those that are associated to metrically-prominent positions in the utterance, while edge-marking tones are associated to the edges of prosodic boundaries (described in further detail below). The pitch contour, or "tune", is made up of these pitch accents and boundary tones, represented on a tier separated from the segments, or "text". This separation in tiers is familiar from analyses of lexical tone, rooted in Goldsmith's (1976) work on autosegmental phonology. The "tune tier" is aligned to the "text tier" at metrically-prominent heads or phrasally-defined edges; at the post-lexical stage of derivation, the segments located at these prosodic positions and the tones associated to those positions are realized simultaneously (see Selkirk 1984 for a more detailed description of this pairing of tune and text).

The AM model of intonational phonology draws upon the metrical phonology representation of relative prominence among syllables, words, and phrases. As mentioned earlier, pitch accents (*i.e.* head-marking tones) can only align to prominent syllables. The exact type of prominence differs across languages – some languages mark their prominent syllables lexically for stress (*e.g.* English, Spanish), which can attract a variety of postlexical pitch accents, while other languages mark prominent syllables with a predetermined pitch contour (*e.g.* Japanese, Serbo-Croatian). For example, the first syllable of the English word *venerate* /ˈvɛnəɹeɪt/ is lexically specified to be the most prominent, and is thus eligible to bear one of the many postlexical pitch accents available in the English inventory. Similarly, the first mora of the Japanese word /sáɴkakɯ/ 'triangle' is lexically-specified to bear a particular pitch accent, although the mora may not bear the phonetic or phonological attributes of stress seen in stress-accent languages (Venditti 1997, Beckman 1986). Some languages use a combination of stress and lexically-specified pitch contour to mark syllable prominence (*e.g.* Swedish, Norwegian). For example, the Norwegian words for 'the seal' and 'the suspender' are segmentally homophonous (*i.e. selen*), but the former is specified to bear a low tone on the stressed syllable, and the latter is specified to bear a falling tone on the stressed syllable (Kristofferson 2003). All three of these prominent syllable types (*i.e.* stress accent, lexical pitch accent, stressed lexical pitch accent) can be considered lexically-specified locations. The tones that are attracted to these prominent, lexically-specified locations are traditionally labeled with an asterisk (*) next to the tonal target (*e.g.* L* for a low pitch accent). Complex tones can include "leading" or "trailing" components, where the

prominent syllable bears the tone marked by the asterisk (*) and the preceding or following syllable bears the unmarked tone, separated by a plus sign (+) from the tone bearing the asterisk – for example, if the hypothetical word *tobásari*, with stress on *bá*, were to be assigned a L*+H pitch accent, the stressed syllable *bá* would be produced with the tone bearing the asterisk, which in this case is L*. Thus the stressed syllable *bá* would bear a low tone, while the following syllable *sa* would bear the "trailing" high tone (H). Similarly, if the same word *tobásari* were to bear a L+H* pitch accent, the stressed syllable *bá* would bear the high portion of the pitch accent (H*) while the previous syllable *to* would bear the "leading" low portion (L). These two possible rising tone types are schematized in Figure 1 below.

Figure 1. The hypothetical word *tobásari* bearing two different types of rising pitch accent: L*+H (left) and L+H* (right).

However, depending on the language and on certain prosodic factors (*e.g.* word size), the exact realization of leading and trailing tones can vary; for example, trailing H tones can be realized two or three syllables after the main stress in the model of Bengali presented in Chapter Three. However, when the word is shorter than three syllables, as in [make] 'mother-ACC' or [nun] 'salt', the trailing H tone can be realized on the second syllable or even the second half of the first syllable.

In addition to prominent syllables, tones can also be associated with the edges of phrases. A language may have one or more prosodic phrase types, grouping together words and/or other smaller phrases, whose edges can be marked by a tone, final lengthening, segmental realizations, or a combination of these phenomena. While prosodic grouping can extend below the word level (*e.g.* feet, syllables, morae), tonally-marked prosodic phrases (*i.e.* those at or above the word level) are categorized into no more than three types. In the analyses that derive prosodic phrasing from syntax (Nespor & Vogel 1986, Selkirk & Tateishi 1989, Hayes 1989, Hayes & Lahiri 1991, among others), the phrases above the word are typically called the Intonation Phrase (IP), which roughly corresponds to the clause, and the Phonological Phrase. In the analyses of prosodic phrasing based on suprasegmental phenomena such as intonation and final lengthening (Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988, Jun 1993, Beckman 1996, Jun 2005, among others), the phrases above the word include the IP and two smaller phrases: the Intermediate Phrase (ip), which is a smaller unit than the IP, and the Accentual Phrase (AP), which is roughly equivalent to a single content word, often including adjacent clitics and function words (see Jun 2005 for an overview). Whether they can be derived from syntactic phrasing or defined by suprasegmental phenomena, these prosodic phrases are often assumed to be structured following the Strict Layer Hypothesis (Selkirk 1984, 1986; Nespor & Vogel 1986), which states that every prosodic constituent is dominated by a constituent of the immediately higher level in the hierarchy, and conversely, every prosodic constituent also contains at least one member of the

immediately lower level in the hierarchy.[1] Thus, every AP is dominated by an ip, and every IP contains at least one ip. Following a strict definition of the Strict Layer Hypothesis, this also means that the edge of every IP aligns with the edge of an ip, and the edge of every ip aligns with the edge of an AP. Not every language has been claimed to have all three prosodic phrase types, however. While some languages tonally mark all three of these phrase types (*e.g.* Georgian: Jun *et al.* 2007), most languages only mark one (*e.g.* Portuguese: Viana & Frota 2007) or two (*e.g.* English: Beckman & Pierrehumbert 1986).

3 INTRODUCTION TO TOBI TRANSCRIPTION SYSTEMS

One of the most widely-used transcription systems based on the AM model of intonation is Tones and Break Indices (ToBI), put together by a team of linguists, speech scientists, and engineers, and laid out in Silverman *et al.* (1992) and Beckman & Hirschberg (1994) as a model of English prosody. Since then, linguists have applied the principles used in what is now called MAE_ToBI (Mainstream American English ToBI) to describe the intonational systems of languages as diverse as Japanese (Beckman & Pierrehumbert 1986, summarized in Venditti 1995), Greek (Arvaniti & Baltazani 2000), and Serbo-Croatian (Godjevac 2001). See Jun (2005) for a collection of ToBI-based transcription systems for twelve languages.

[1] See Inkelas (1989), Itô & Mester (1992), and Selkirk (1996) for a discussion of the decomposition of the strongest form of the Strict Layer Hypothesis into individual constraints.

There are six parts to a ToBI transcription: an audio recording of the utterance, a record of the F_0 contour, usually superimposed on a spectrogram or waveform, and four transcription tiers (*i.e.* words, tones, break indices, and miscellaneous). The word tier includes the orthographic (or transliterated, for non-Roman scripts) representation of the segments in the utterance. The tone tier includes the distinctive tonal events, including pitch accents and boundary tones. The break index tier includes integer numbers corresponding to the perceived juncture size between words. Finally, the miscellaneous tier may include any additional information about the utterance (*e.g.* paralinguistic information such as stuttering or laughing), or other information such as the transcriber's notes to colleagues regarding a troublesome contour.

The model of Bengali prosody proposed in the current study uses a ToBI-style transcription system (*i.e.* B-ToBI) to annotate experimental data, formulated within the Autosegmental-Metrical framework of Intonational Phonology.

4 INTRODUCTION TO BENGALI AND ITS DIALECTS

With over 170 million speakers, Bengali is one of the world's most widely-spoken languages (Gordon 2005). It is an Indo-European language spoken in eastern South Asia; the approximate boundaries of the Bengali-speaking region—which, in addition to Bangladesh and Indian West Bengal, includes parts of the Indian states of Assam, Bihar, Jharkhand, Mizoram, and Tripura, as well as northwestern Burma (Myanmar)—are

outlined in Figure 2 below. The dotted lines to the north and west of the Bengali-speaking region represent the undetermined boundary between dialects of Bengali and those of neighboring Indic languages.

Figure 2. Map of the approximate boundaries of the Bengali-speaking region. Dotted portions of the boundary represent areas where dialects of Bengali overlap with dialects of neighboring Indic languages.

The Bengali-speaking region is home to a vast dialect continuum, and many dialects are largely unintelligible to those only familiar with the standard. Often, the dialects at the extremes of this continuum (*e.g.* Chakma, Chittagonian, Hajong, Kharia Thar, Mal Paharia, Sylheti, Tangchangya) are considered separate languages (Gordon 2005), although the border between these putative languages and Bengali is unclear.

Grierson's (1928) well-known survey of Bengali dialects is to this day used as the basic classification of the language's many variants (Gordon 2005). Grierson divides the Bengali language into an Eastern Branch and a Western Branch, whose dividing line does not follow any particular national or geographical boundary. I summarize here the distribution of the various dialect groupings according to this classification.

The Western Branch of Bengali straddles the border between Bangladesh and Indian West Bengal, and includes the Central, Northern, Western, and Southwestern dialects. Reaching into the states of Bihar, Jharkhand, and Orissa, the Western and Southwestern dialects can be thought of as parts of the continuum of Bengali into the Sadri and Oriya languages, respectively. Northern Bengali is spoken in the foothills of the Himalaya, on both the Indian and Bangladeshi sides of the border, and is part of the continuum from Bengali to Assamese. Central Bengali is the grouping that contains the two most widely-used forms of the language: Nadia dialect, which is the basis of Standard Bengali, and Kolkata dialect, which also has great influence over the standard. Standard Bengali is officially known as [tɕolt̪i bʰaʃa] 'current language', and also informally and perhaps misleadingly as [ʃud̪ʰːo baŋla] 'pure Bengali' or [bojeɹ bʰaʃa] 'book language'.

The Eastern Branch includes the Eastern, East-Central, and Southeastern dialects. Eastern Bengali dialects are spoken in a large area encompassing the Dhaka, Sylhet, and Barisal Districts of Bangladesh, as well as in parts of Tripura and Assam States in India.

East-Central dialects are spoken in the Khulna District of Bangladesh, and are considered midway between Eastern and Central Bengali. Southeastern dialects are found in the Chittagong District of Bangladesh, as well as parts of Mizoram State in India and Rakhine State in Myanmar, and are at the eastern extreme of the Bengali dialect continuum, practically surrounded by Tibeto-Burman languages. All dialects in the Eastern Branch are considered non-standard, although the Eastern Bengali dialects around the Bangladeshi capital of Dhaka presumably have some influence on the standard form of the language spoken in that country.

Grierson's (1928) classification of Bengali dialect groupings is illustrated below in Figure 3, and also arranged in Outline 1, with major dialect names provided under each grouping.

Figure 3. Map of the eight major dialect regions of the Bengali language according to Grierson (1928). Northern, Central, Western, and Southwestern dialects fall under the Western Branch, and Eastern, East-Central, and Southeastern dialects fall under the Eastern Branch. Rajbanshi was originally classified within the Eastern Branch in Grierson (1928), but was renamed Rangpuri and reclassified within the Western Branch in Shahidullah (2000).

I. **Western Branch**

a. Central Bengali

i. In Indian West Bengal: Nadia (Standard Bengali), Kolkata, Haora, Tamluk, Medinipur, Murshidabad, Barddhaman

ii. In Bangladesh: Kushtia

b. Northern Bengali

i. In Indian West Bengal: East Malda, Koch Bihar

ii. In Bangladesh: Rajshahi, Dinajpur, Bogra, Pabna

c. Western Bengali

i. In Indian West Bengal: Kharia Thar, Mal Paharia, Manbhum

ii. In Indian Bihar: Saraki

d. Southwestern Bengali

II. **Eastern Branch**

a. Eastern Bengali

i. In Bangladesh: Dhaka, southeastern Faridpur, Mymensingh, Comilla, Bakerganj, Sylhet, Hajong, Sandwip Island

ii. In Indian Assam: Cachar

b. East-Central Bengali

i. In Bangladesh: Jessore, Khulna, Faridpur

c. Southeastern Bengali

i. In Bangladesh: Noakhali, Chittagong, Chakma, Tangchangya

ii. In Myanmar: Sittwe

d. Rajbanshi[2]

i. In Bangladesh: Rangpur

ii. In Indian West Bengal: Siripuria, Jalpaiguri, Bahe

iii. In Indian Assam: Goalpara

Outline 1. Grierson's (1928) classification of Bengali dialects.

[2] Shahidullah (2000) suggests that the Rajbanshi dialect grouping be renamed Rangpuri, to avoid confusion with the Rajbanshi tribe. He also finds that the dialect grouping belongs under the Northern Bengali group, part of the Western Branch of Bengali, instead of forming its own group as part of the Eastern Branch.

Dialects differ widely in all aspects of the grammar, with the greatest variation found in the phoneme inventory, allophony, and inflectional morphology. Therefore, although speakers of all dialects are familiar to some degree with the standardized form of the language, regional dialect influence on an individual's production of Standard Bengali can be significant. For this reason, it may be relevant to use labels such as "Kolkata Standard Bengali" or "Bangladeshi Standard Bengali". In addition to morphology and vocabulary choice, the tense *vs.* lax vowel distinction, oral *vs.* nasal vowel distinction, modal *vs.* breathy voicing distinction, /ɹ/ *vs.* /ɽ/ distinction, /s/ *vs.* /ʃ/ distinction, vowel rounding harmony, voicing harmony, /ɹ/-assimilation, and degree of spirantization are some of the possible areas in which a Bengali speaker's native dialect(s) may influence his or her production of Standard Bengali (Chatterji 1921, Grierson 1928). And as the production of Standard Bengali is known to vary due to speakers' familiarity with nonstandard dialects, there is every reason to believe that speakers' prosodic patterns may also vary considerably. Therefore, it is of great importance both to identify the possible dialect influences on Bengali speakers (even when speaking some form of Standard Bengali) when analyzing their prosody, and to be cautious in making direct comparisons between models of Bengali prosody, as the models may be describing very different systems. Therefore, I identify the dialect of Bengali described in each of the previous studies (Chapter Two), and what dialects may serve as influences on each individual's production of Standard Bengali in the current study (Chapter Three).

5 INTRODUCTION TO THE DISSERTATION

Describing the prosodic patterns of Bengali is hardly a new or pioneering research endeavor; as mentioned earlier, and as more thoroughly discussed in Chapter Two, many linguists have analyzed and modeled the stress, pitch, and phrasing of various dialects of Bengali in a variety of theoretical frameworks, including the AM theory of intonational phonology. What is lacking, however, is a comprehensive model, based on data collected from a large number of speakers from a wide range of dialect backgrounds, producing scripted sentences controlled for word length, sentence type, focus domain, focus type, and other variables, as well as producing naturalistic speech in both (semi-)standard and nonstandard varieties of Bengali. Every component of the model proposed in the current study is reflected in the speech of multiple speakers, and where appropriate, numerical measurements are made to support each finding. The goals of the current study include the collection of a large corpus of scripted and naturalistic data, the proposal of a model of Bengali prosody that takes into account influences from nonstandard dialects, and comparison of the proposed model to previous models of different varieties of Bengali prosody as well as to those of other languages. I hope that the current study not only serves as a reliable model of Standard Bengali prosody but also can be applied towards establishing prosodic models for nonstandard varieties of Bengali as well as related (and severely understudied) languages such as Assamese, Bishnupriya Manipuri, Oriya, and the Bihari languages (*e.g.* Maithili, Magahi, Sadri).

The remainder of this dissertation is divided into four chapters. Chapter Two summarizes previous literature on the prosody of different dialects of Bengali, including studies of stress, tone, and phrasing. Chapter Three describes in detail the experiments carried out for the current study as well as the resulting intonational phonological model and accompanying transcription system (*i.e.* B-ToBI), including the entire inventory of tone and phrase types in Standard Bengali as spoken by speakers from various parts of Bangladesh and West Bengal, and issues of tonal interaction. Chapter Four discusses the prosodic realization of focus, including its tonal and phrasal characteristics. Both Chapters Three and Four include comparisons of the current model with previous descriptions of various Bengali dialects, proposing additional tones and prosodic phrasing patterns not previously described, as well as an abstract tone for focus realization. Finally, Chapter Five summarizes the findings of the current study, discusses their similarities to related languages, and ends with concluding remarks.

CHAPTER TWO
PREVIOUS STUDIES

In Chapter One, I briefly discussed the study of intonation, the theoretical framework known as Intonational Phonology, and the prosodic transcription system known as ToBI. In this chapter, I focus on the study of intonation, both within and outside the Intonational Phonology framework, as it pertains to various dialects of Bengali. As intonation is closely tied with other aspects of prosody, including stress and phrasing, I summarize previous descriptions and models of all aspects of the prosody. I begin in §6 with a review of studies dealing solely with the stress pattern of the language, with additional phonological evidence supporting the designation of Bengali as a stress-accent language. In continue in §7 with a review of pre-Intonational Phonology studies of the prosodic system. Finally, I discuss previous studies of the prosodic system within an Intonational Phonology framework in §8.

6 STRESS

Intonation is inextricably tied to other aspects of prosody, including stress and phrasing. The manner in which tones align with segments is sensitive to the distribution of prominent syllables and phrase breaks. As mentioned in Chapter One, the Autosegmental-Metrical theory of Intonational Phonology posits two tone types: pitch accents (*i.e.* head-marking tones) and boundary tones (*i.e.* edge-marking tones). Pitch accents are tones that are characterized by their association with prominent syllables, where "prominence" refers to metrical structure. Thus, in languages where the syllables

of a word are ranked relative to one another in a metrical hierarchy, the most prominent of those syllables will be eligible to bear a pitch accent. As this ranking within the metrical hierarchy is the most basic definition of stress, any discussion of a language's pitch accent system must be accompanied by a clear understanding the assignment of stress. In this section, I include a short review of previous studies that deal with the stress pattern of Bengali. I begin in §6.1 with descriptions of word-level stress, and discuss phrase-level stress in §6.2. The phonetic characteristics of stress are described in §6.3. Phonological properties of stress are listed in §6.4, and lastly, pitch accent attraction is discussed in §6.5.

6.1 Word stress

Stress-accent languages are those in which words are assumed to bear stress (*i.e.* strong metrical prominence) on (at least) one syllable, and in which pitch accents are attracted to these metrically prominent syllables. Stress assignment (*i.e.* the patterns of assigning metrical prominence to the syllables of a word) varies considerably cross-linguistically. In many stress-accent languages, the location of the stressed syllable is not entirely predictable, and can in fact be a contrastive feature between words (*e.g.* Spanish [ˈxuɣo] 'juice' *vs.* [xuˈɣo] 'played').[3] In other stress-accent languages, stress assignment is predictable from either syllable weight (*e.g.* heavy-syllable stress in variants of Modern Standard Arabic [ixtiˈlaːf] 'difference', [muˈxaːlafa] 'contrast'[4]) or the position of a syllable in the word (*e.g.* word-initial stress in Finnish [ˈlyhytnækøinen] 'short-sighted',

[3] Spanish examples and analysis taken from Gutiérrez-Palma & Palma-Reyes (2007).

[4] Arabic examples taken from Cowan (1993); analysis taken from Birkeland (1954).

[ˈkeŋːænːɑuhɑ] ‘shoelace’;[5] word-final stress in Thai [ʔaˈrɔ̀j] ‘delicious’, [sǐːnáːmˈtāːn] ‘brown’[6]; penultimate stress in Swahili [tɛʤɛˈmɛa] ‘depend’, [kiˈtaⁿda] ‘bed’[7]). Of course, in languages where stress assignment is predictable from syllable weight or position, stress cannot serve as a contrastive feature, and no minimal pairs can exist.

While all studies of Bengali prosody (Chatterji 1921; Goswami 1944; Ferguson & Chowdhury 1960; Anderson 1962; Ray, Hai, & Ray 1966; Bykova 1981; Shaw 1984; Kawasaki & Shattuck-Hufnagel 1988; Hayes & Lahiri 1991; Lahiri & Fitzpatrick-Cole 1999; Das 2001; Michaels & Nelson 2004; Selkirk 2006) would classify Bengali as a stress-accent language, they also add that word stress is not a contrastive feature in the language. Thus no two words can be differentiated solely on the basis of stress, unlike in languages such as Spanish.

It is claimed in nearly all studies (Chatterji 1921, Goswami 1944, Ferguson & Chowdhury 1960, Anderson 1962, Bykova 1981, Kawasaki & Shattuck-Hufnagel 1988, Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999, Michaels & Nelson 2004; Selkirk 2006) that words are consistently stressed on the initial syllable (*e.g.* [ˈʃudʑi] ‘farina’, [ˈmakʰietɕʰilen] ‘you-HON had mashed’). Only two studies, to my knowledge, claim that stress can occur on non-initial syllables. Shaw (1984) claims that stress is assigned to the

[5] Finnish examples taken from Wuolle (2002); analysis taken from Suomi & Ylitalo (2002).

[6] Thai examples taken from Pittayaporn (2007); analysis taken from Potisuk *et al.* (1996).

[7] Swahili examples and analysis taken from Brandon (1975).

second syllable (*e.g.* [maˈli] ‘gardener’, [aˈkaʃ] ‘sky’, [koˈbi̪ta] ‘poem’), unless the first syllable is heavy (*i.e.* closed), in which case the initial/heavy syllable is stressed (*e.g.* initial stress in [ˈʃɔŋʃaɹ] ‘family’, [ˈɔntɕol] ‘region’). Das (2001), which focuses on the prosody of Tripuri Bengali, also claims that the stress pattern of both Tripuri Bengali and Standard Bengali is quantity-sensitive. Proposing detailed rules of metrical foot composition, Das claims that initial stress can only be found on a light syllable in words where the initial syllable is parsed into a binary foot (*e.g.* [ˈalɔsɔna] ‘discussion’, [ˈɔnonuˌgɹɔhoniˌjɔ̪ta] ‘unacceptability’).[8] In his detailed analysis of Tripuri Bengali and Tripuri English, Das also discusses the distribution of secondary stress. As neither of these two non-initial stress studies was focused on the phonetic correlates of stress, no acoustic measurements were made.[9]

It can be presumed from comparing these previous studies that Bengali words indeed bear stress, and that stress assignment is either consistently word-initial, as described in the majority of studies, or dependent on syllable weight, as described in Shaw (1984) and Das (2001). Phonological and prosodic evidence presented in Chapter Three sheds further light on the question of stress assignment, and suggests that initial syllables indeed bear a phonologically-significant but phonetically-weak stress.

[8] In descriptions of Standard Bengali that propose fixed initial stress, these words would be transcribed [ˈalotɕɔna], [ˈɔnonugːɹohonioo̪ta].

[9] It is unclear what phonetic cues previous studies had used as correlates of stress. Presumably, the occurrence of pitch maxima and minima (indicating pitch accent attraction), high amplitude, and vowel length could serve as potential cues to metrical prominence.

6.2 Phrasal stress

Most studies (Chatterji 1921, Goswami 1944, Ferguson & Chowdhury 1960, Ray *et al.* 1966, Bykova 1981, Shaw 1984, Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999) also agree that word stress is largely subservient to phrasal stress; while one syllable (*i.e.* the initial syllable in most studies) of each word is considered stressed, one word in each phrase will be considered the most prominent. In a typical sentence, the most prominent word can be either the leftmost (Chatterji 1921, Goswami 1944, Ray *et al.* 1966, Bykova 1981, Shaw 1984) or the rightmost (Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999, Selkirk 2006) in the phrase. The latter three studies also find that phenomena such as focus can change the default pattern of phrasal stress assignment.

6.3 Phonetic correlates of stress

While numerous studies describe the existence of phonetically weak stress for native words, only one study (*i.e.* Kawasaki & Shattuck-Hufnagel 1988) has explicitly investigated the existence of phonetic correlates to stress in the language. Kawasaki & Shattuck-Hufnagel (1988) finds that high F_0 and intensity do not serve as reliable cues to stress in Bengali.[10] Two additional studies (Chatterji 1921, Anderson 1962) make a claim about phonetic correlates of Bengali stress, both stating that stress is often phonetically realized in conjunction with vowel lengthening, although durational measurements were not provided in either study. Thus there is currently no clear evidence about the phonetic nature of stress in Bengali.

[10] The lack of correlation between a high F_0 and stress may, of course, be due to the fact that Bengali words are typically produced with a low pitch accent on the initial syllable; see Chapter Three.

6.4 Phonological correlates of stress

Although stress in Bengali appears to be phonetically weak, it is nevertheless phonologically salient. Bengali clearly does assign a special status to particular syllables, whose properties are roughly equivalent to the non-phonetic correlates of stress in English. This special status is always a property of the initial syllable of native monomorphemic words, and normally also on borrowings and complex words.[11] There is often a similar status on all odd-numbered syllables after the initial syllable.[12] I will refer to this special status as (primary- and secondary-) phonological stress, to distinguish it from the phonetically-measurable stress such as that of English. I summarize below some phonological observations about phonologically stressed syllables, including the ability to host lax vowels (§6.4.1), the ability to host nasal vowels (§6.4.2), adaptation of loanwords (§6.4.3), the avoidance of syncope (§6.4.4), attraction of metathesized vowels (§6.4.5), and the restriction against diphthongizing with vowel-initial enclitics (§6.4.6).

6.4.1 Tense-lax vowel distinction

Bengali has seven vowels in its inventory /i, e, ɛ, a, ɔ, o, u/ (plus seven nasal vowels for dialects that have contrastive nasalization), but phonologically unstressed syllables rarely host lax vowels /ɛ, ɔ/ (Dasgupta 2003); the appearance of lax vowels appears to be

[11] Primary stress is nearly always word-initial; however, some words borrowed from Perso-Arabic, English, or Sanskrit can bear non-initial primary stress in some speakers' pronunciations. For example, the Sanskritic re-borrowing /niɹ-batɕ-ɔn/ 'election' is produced in news broadcasts with primary stress either on the word-initial syllable, giving [ˈniɹbatɕon], or on the root-initial syllable, giving [niɹˈbatɕon].

[12] Secondary stress is extremely variable, as it is highly sensitive to (speakers' intuitions of) word-internal morpheme boundaries. See Khan (2006) for an introductory discussion of phonological words in Bengali.

sensitive to phonological stress. For example, the monomorpheme [ɛk] ‘one’ is produced with a lax vowel in isolation, but is produced with a tense vowel when attached to the negative prefix /ɔn/-, giving /ɔn-ɛk/ [ɔnek] ‘multiple’. This raising of the lax vowel /ɛ/ to [e] appears to be conditioned by the shift of stress from the root to the word-initial syllable. Similarly, the /ɔ/ of /pɔɹad̪ʑ-i̪to/ [pɔɹad̪ʑi̪to] ‘defeated’ is raised to [o] when the negative prefix [ɔ]- is added, giving /ɔ-pɔrad̪ʑ-i̪to/ [ɔpoɹad̪ʑi̪to] ‘undefeated’.[13]

6.4.2 Oral-nasal vowel distinction

In dialects of Bengali that have contrastive vowel nasalization,[14] phonologically unstressed syllables rarely host nasal vowels /ĩ, ẽ, ɛ̃, ã, ɔ̃, õ, ũ/ (Dasgupta 2003). While words like [kãta] ‘thorn’ and [t̪ẽt̪ul] ‘tamarind’ are common, there are few monomorphemic words of the type *[katã] or *[t̪et̪ũl]. Secondarily-stressed syllables can also be nasalized, as in the French borrowing [ɹest̪oɹã] ‘restaurant’. Rare exceptions to this rule include Sanskritic reborrowings such as [at̪ːã] ‘soul’ and [at̪ːĩo] ‘relative’, which Dasgupta (2003) describes as being actually pronounced [ãt̪ːa] and [ãt̪ːio] in all contexts other than stage performances. A handful of foreign words such as [oɹãɔ̃] ‘Oraon’ (an ethnic group) make up the only true exceptions. In words made of a repeated monosyllable (*e.g.* [tuptup] ‘dripping’, [mama] ‘mother’s brother’), a nasal vowel in the initial syllable may optionally be produced as an oral vowel in the phonologically

[13] These transcriptions come from recordings of various speakers. The variant [ɔpɔɹadʑi̪to] ‘undefeated’ is listed as the standard pronunciation in Biswas (2007), a source that tends to preserve morpheme-initial lax vowels regardless of the stress pattern.

[14] Shahidullah (2000) recognizes Eastern Bengali as the only dialect group to not contrast oral and nasal vowels. This feature is often carried over into the Standard Bengali as spoken in Eastern dialect areas.

unstressed reduplicant (*e.g.* [dʑʱĩdʑʱĩ]~[dʑʱĩdʑʱi] ‘cricket’, [kʰãkʰã]~[kʰãkʰa] ‘desolate’). Thus, the appearance of nasal vowels appears to be sensitive to phonological stress.[15]

6.4.3 Loanword adaptation

Loanwords are often pronounced with the initial syllable phonologically stressed, regardless of the stress pattern of the donor language; this often leads to reduction or total deletion of a non-initial syllable, even if it is the stressed syllable in the donor language (*e.g.* [amerika]~[ɛmeɹika] ‘America’, colloquial [ɛmɹika], presumably from American English [əˈmɛɹəkə]). This suggests that the stress pattern of the original word can often be ignored in when borrowed into Bengali. However, a few foreign words with non-initial stress have been borrowed into the language with adjustments made to fit the initial-stress system of the language; in words that have primary stress on the second syllable in the donor language, the unstressed initial syllable is dropped completely in the Bengali version to conform to the language’s stress pattern (*e.g.* [maɹkin] ‘American’, presumably from American English [əˈmɛɹəkən], and the colloquial pronunciations of the Muslim greeting [slamalikum]~[slamlikum] from Arabic [asːaˈlaːmu ʕaˈlajkum] ‘peace be upon you’).

[15] This is also similar to the optional loss of aspiration in the second iteration of a repeated monosyllable in words like [t̪ʰut̪ʰu]~[t̪ʰut̪u] ‘spit’.

6.4.4 Syncope

Many disyllabic roots lose their unstressed second syllable when suffixes are added. Suffixed forms without syncope would be considered ungrammatical. This alternation is not productive. Examples are provided in Table 1 below.

Base UR[16]	Base SR	Suffixed UR	Suffixed SR
/atok/	[atok] 'obstruction'	/atok-ano/	[atkano] 'obstructing'
/pagol/	[pagol] 'crazy'	/pagol-ami/	[paglami] 'crazy behavior'
/d̪uʃtu/	[d̪uʃtu] 'naughty'	/d̪uʃtu-ami/	[d̪uʃtami] 'naughty behavior'
/aɹob/	[aɹob] 'Arab'	/aɹob-i/	[aɹbi] 'Arabic'

Table 1. Examples of unstressed vowel syncope in suffixed forms.

Some trisyllabic monomorphemes freely alternate between trisyllabic forms without any syncope and disyllabic forms with syncope of the unstressed syllable. Thus, while the stressed initial syllable is protected from deletion, the unstressed second syllable is subject to this effect. This alternation is not productive. Examples are provided in Table 2 below.

UR	SR
/dʑanala/	[dʑanala]~[dʑanla] 'window'
/is̪tiɹi/	[is̪tiɹi]~[is̪tɾi] 'ironing'
/t̪aɾa-t̪aɾi/	[t̪aɾat̪aɾi]~[t̪aɾt̪aɾi] 'quickly'
/abo-ha͡ɔa/	[abohaɔ͡a]~[abhaɔ͡a] 'weather'

Table 2. Examples of unstressed vowel syncope in free variation with forms without syncope.

[16] UR = underlying representation (*i.e.* abstract form); SR = surface representation (*i.e.* phonetic form)

The fact that syncope affects only a word's second syllable, which is presumed to bear no stress, suggests that Bengali words in fact do bear stress, the lack of which leaves vowels vulnerable to deletion.

6.4.5 High vowel metathesis

In dialects that have high vowel metathesis, underlying high vowels /i, u/ in non-final phonologically unstressed syllables shift to phonologically stressed syllables, surfacing as the off-glide [j]. This evacuation of the second syllable is presumably related to its unstressed status, suggesting that stress is phonologically salient in these dialects of Bengali. Examples are provided in Table 3 below.

Base UR	Base SR	Suffixed UR	Suffixed SR
/manuʃ/	[manuʃ] 'person'	/manuʃ-e/	[majnʃe] 'people'
/pal-ok/	[palok] 'fosterer'	/pal-ika/	[pajlka] 'adopted'
/balu/	[balu] 'sand'	/balu-a/	[bajlːa] 'sandy'
/fazil/	[fazil] 'precocious'	/fazil-ami/	[fajzlami] 'precocity'

Table 3. Examples of unstressed vowel metathesis in suffixed forms.

Metathesis can also cause underlying high vowels in non-final phonologically unstressed syllables to shift to phonologically stressed syllables in free variation with syncopated versions, as shown in Table 4 below.

UR	SR
/naɹikel/	[najɹkel]~[naɹkel] ‘mature coconut’
/ʃoɹiʃa/	[ʃojɹʃa]~[ʃoɹʃa] ‘mustard’

Table 4. Examples of unstressed vowel metathesis in free variation with forms without syncope.

6.4.6 Enclitic alternations

When attached to a vowel-final stem, the enclitics -/i/ ‘only’, -/o/ ‘also’, -/e/ LOC, and -/eɹ/ GEN are produced without the full vowel [j, ə̯, ɛ̯, ɹ]. Switching to non-vocalic forms is blocked, however, if the vowel preceding the enclitic bears primary phonological stress; hiatus is maintained in such situations.[17] This alternation is productive. Examples are provided below in Table 5.

[17] This alternation is specific to the aforementioned enclitics; the otherwise homophonous suffixes -/i/ FIRST.PERSON, -/o/ SECOND.PERSON, -/e/ THIRD.PERSON do not have any such restriction against losing vocalic status.

Base UR	Base SR	Encliticized UR	Encliticized SR
/kʰõpa/	[kʰõpa] ‘hairbun’	/kʰõpa-i/	[kʰõpaj] ‘only (a) bun’ *[kʰõpa.i]
		/kʰõpa-o/	[kʰõpa͡ɔ] ‘also (a) bun’ *[kʰõpa.o]
		/kʰõpa-e/	[kʰõpa͡ɛ] ‘in (a) bun’ *[kʰõpa.e], *[kʰõpa͡ɛ.e]
		/kʰõpa-eɹ/	[kʰõpaɹ] ‘bun’s’ *[kʰõpaeɹ], *[kʰõpa͡ɛeɹ]
/pa/	[pa] ‘(a) leg’	/pa-i/	[pa.i] ‘only (a) leg’ #[paj] (*cf.* [paj] ‘I find’)
		/pa-o/	[pa.o] ‘also (a) leg’ #[pa͡ɔ] (*cf.* [pa͡ɔ] ‘you find’)
		/pa-e/	[pa.e]~[pa͡ɛ.e] ‘on (a) leg’ #[pa͡ɛ] (*cf.* [pa͡ɛ] ‘finds’)
		/pa-eɹ/	[pa.eɹ]~[pa͡ɛ.eɹ] ‘leg’s’ #[paɹ] (*cf.* [paɹ] ‘edge’

Table 5. Examples of different suffix/enclitic alternations adjacent to stressed and unstressed vowels.

The fact that stressed vowels avoid diphthongizing with vowel-initial enclitics while unstressed vowels show no such restriction suggests that stress is a salient feature of Bengali phonology.[18]

6.5 Pitch accent attraction

In terms of theoretical motivation, another compelling reason to classify Bengali as a stress-accent language is related to the association of pitch accents. As mentioned in Chapter One, the autosegmental-metrical model of intonational phonology assumes that post-lexical tones can associate either to prosodic heads or to prosodic edges, where

[18] Another interpretation of this fact involves positing a violable constraint on minimal word size. See Fitzpatrick-Cole (1994) and Lahiri & Fitzpatrick-Cole (1999) for a detailed discussion.

“prosodic edge” refers to the boundary of a prosodic phrase, and “head” refers to the most metrically-prominent syllable within that phrase. Thus, for a pitch accent to align to a word, it must find the most metrically-prominent syllable therein, *i.e.* the primary-stressed syllable, as it does in languages such as English, Spanish, and German. Given this assumption about tonal alignment, we can take pitch accent assignment to be an indicator of stress assignment. More generally, we can take tonal assignment to be an indicator of stress assignment in the absence of prosodic edges.

As seen in §1.1, it is almost universally agreed that Bengali stress is consistently word-initial, and thus a hypothetical Bengali word such as /ˈɡɔnomi̪ta/ could only bear a tone on the first syllable /ˈɡɔ/ and the last syllable /̪ta/. While the tone of the final syllable /̪ta/ could be understood to be a boundary tone aligned to the right edge of the phrase, the tone on the syllable /ˈɡɔ/ would be ambiguous in its underlying association; is the tone attracted to the syllable’s stressed status (*i.e.* indicative of a pitch accent) or its position as leftmost in the phrase containing the word (*i.e.* indicative of a boundary tone)? Disambiguating evidence is found, however, in a number of studies (Chatterji 1921, Anderson 1962, Shaw 1984, Hayes & Lahiri 1991), where it is pointed out that function words and particles either do not bear stress or at least do not bear the strongest stress in the phrase even when they are in phrase-initial position. Hayes & Lahiri (1991) includes data of this kind, showing pitch accent placement on the initial syllables of non-phrase initial words, when the phrase-initial word is a function word. In other words, adding a hypothetical function word to non-phrase-final content word, giving /ˈena ˈɡɔnomi̪ta/, can

disambiguate whether the tone of the first syllable of /ˈgɔnomi̪ta/ is a pitch accent or boundary tone. The findings of Hayes & Lahiri (1991) suggest that a phrase like /ˈena ˈgɔnomi̪ta/ can bear its pitch accent on the initial syllable of the content word, /ˈgɔ/, instead of the initial syllable of the entire phrase, /ˈe/. This could not be explained under an analysis that posits that the tone is attracted to the left edge of a phrase (*i.e.* a boundary tone), but it could be explained in an analysis that posits that the tone is attracted to the metrically strongest syllable in the phrase (*i.e.* pitch accent), which could very well be the stressed syllable of a content word instead of that of a preceding function word. If this is the case, then it is clear that Bengali stress is not only relevant for segments and their alternations, but also for postlexical tone association.

6.6 Summary

While previous studies disagree on the pattern of stress assignment in different varieties of Bengali, the majority opinion is that stress is consistently assigned to word-initial syllables. While phonetic evidence for word stress is inconclusive, phonological evidence supports the word-initial stress analysis, as word-initial syllables have a special status in their ability to host a greater range of phonemic distinctions, and in their ability to resist alternations undergone by unstressed syllables (*e.g.* syncope, metathesis, diphthongization). In addition to this phonological evidence, the power of stressed syllables to attract postlexical tones away from prosodic edges confirms that Bengali

words bear phonological stress, and the tones associated to phonologically stressed syllables are indeed pitch accents.

7 PRE-INTONATIONAL PHONOLOGY STUDIES

In this section, I provide reviews of the three major pre-Intonational Phonology studies of Bengali prosody: Chatterji (1921), Ferguson & Chowdhury (1960), and Ray, Hai, & Ray (1966). These analyses include descriptions of stress, phrasing, juncture, tones, contours, and focus. I also include a comparison of the various phrasing hierarchies proposed by these three studies in §7.4.

7.1 Chatterji (1921)

7.1.1 One-word utterances

While he acknowledges that "[i]ntonation or pitch of voice is not a significant[19] element of speech in Bengali" (p. 20), Chatterji (1921) described the distinction in pitch contours on utterances controlled for segmental content, thus serving as the first scientific investigation of intonation in Bengali. The dialect described is his own, which he identifies as "the standard colloquial [dialect] of Calcutta", "habitually spoken by the educated classes of Calcutta and of West Bengal generally" (p. 2). This is taken to be a Central Bengali dialect, presumably drawing on features of the Kolkata and Standard Nadia dialects.

[19] Although he does not specify what is meant by the word "significant" here, Chatterji may have been referring to the fact that (at least most dialects of) Bengali are non-tonal, and thus words cannot be distinguished solely based on their pitch.

By using the words হাঁ *hã* [hɛ̃] ‘yes’ and হুঁ *m* [m̩m̩]~[n̩n̩][20] (as in English, a syllabic consonantal interjection normally signifying agreement) as his utterances, Chatterji distinguished high rising, mid rising, low falling, mid falling-rising, and high falling contours, shown in Table 6 in conjunction with the interjection *m*.

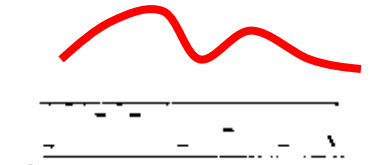
Tone	**Chatterji’s transcription (using *m*)**	**Meaning (using *m*)**
High rising	ˊm	Query: “Yes?”
Mid rising	ˏm	Annoyance or disgust
Low falling	ˎm	‘I see’
Mid fall-rise	ˬm	‘Yes, it may be so, but –’
High falling	˩m	Threat: ‘Very well, I shall see’

Table 6. Tone inventory described by Chatterji (1921), with contrasts illustrated on the interjection *m*.

7.1.2 Multiword utterances

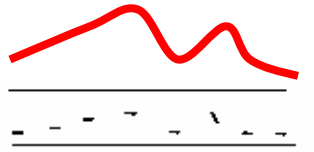
Chatterji also provides ten example sentences with hand-drawn representations of pitch contours above the text. Although he does not point out this fact explicitly, all ten of Chatterji’s transcribed sentences include what appear to be low tones on the initial syllable of all content words, as shown in (1) through (4) below, with solid lines representing pitch tracks added for clarity: [Chatterji §63, p. 21]

[20] The IPA transcriptions here and in later examples, as with the morpheme-by-morpheme glosses, are my own addition.



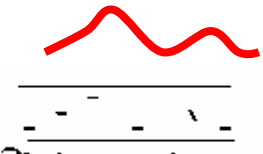
ˈbæbʃaṭa ʃiggir ˈʄombe

(1) [ˈbɛbʃata ʃigːɪɹ ˈdʑombe]
the business soon will flourish
‘The business will flourish quickly.’



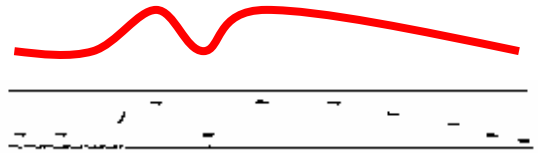
ˈgoṭakɔtɔkˈṭaka(ĕ)eno

(3) [ˈgotakɔ̰tok ˈtaka (ḛ)eno]
some rupee you’ll bring
‘You will bring some rupees.’



ʃheleṭi ma·bolle

(2) [tɕʰeleti ma: bolːe]
the boy mother called
‘The little boy called “Mother!”’



ˈʃɔlo, ˈaːʄi brinˈdaːbonʄattra kori

(4) [ˈtɕɔlo ˈaːdʑi bɹinˈd̪aːbondʑat̪ːɾa koɹi]
come today-CL Brindaban-journey we-do
‘Come, let’s start for Brindaban even to-day.’

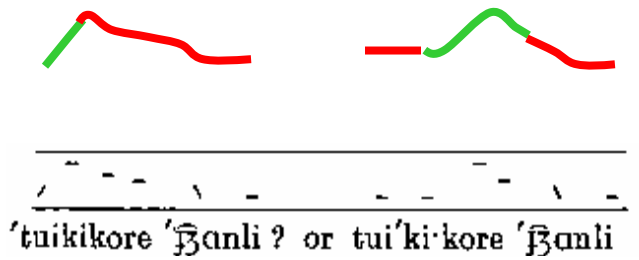
Many of these initial syllables also correspond to stressed syllables, transcribed with a preceding acute accent mark (´) in Chatterji's transcription, thus suggesting that there is some regular association of low pitch on initial (and stressed) syllables – a pattern described in later studies as the result of a low pitch accent (L*) on the stressed syllable and a high boundary tone at the word's right edge.[21] The only exception to this association of low tone and initial/stressed syllables is found in example (4), where the compound *brin´da:bon-ʒʒattra* [bɹinˈd̪a:bon-dʑat̪:ɾa] 'Brindaban-journey' is transcribed with stress on the second syllable, also marked for length (possibly due to the stress, as vowel length is not contrastive in Bengali). In most studies of Bengali stress, this word would be claimed to have initial stress, allowing it to conform to the generalization that initial syllables of content words in Bengali are stressed and bear a low pitch.[22] Another generalization seen across the above examples (along with the other six examples provided in Chatterji's text) is that words that do not immediately precede a major prosodic break (marked by a comma or question mark) or utterance boundary display a rising contour from the stressed syllable's low tone to the word edge. That is to say, words typically show a rising pitch unless they precede some sort of pause, an observation that is noted in many subsequent studies of Bengali intonation.

7.1.3 Focus

[21] Depending on the analysis, this boundary tone could be transcribed as H_P (*i.e.* a high tone borne at the right edge of a phonological phrase) or Ha (*i.e.* a high tone borne on the right edge of an accentual phrase).

[22] It is not clear if this distinction in stress is indicative of a larger dialect difference between Chatterji's (1921) data and the data described in later studies.

Lastly, although Chatterji did not explicitly discuss the realization of focus, he does provide a nice example of the contrast between narrow focus on a subject and normal wh-word focus, reproduced in (5) below, with green lines added to represent the portion of the pitch track associated with the focused constituents: [Chatterji §63, p. 21]



(5) [ˈt̪uj ki koɹe ˈdʑanli] [t̪uj ˈkiː koɹe ˈdʑanli]
you how knew you how knew
'How did *you* know?'[23] 'How did you know?'

The second sentence in (5) appears to have neutral focus on the wh-phrase [ki koɹe] 'how' (*lit.* 'having done what'), transcribed as primary stress and phonetic lengthening on the initial syllable [ki] 'what', giving [ˈkiː]. The corresponding pitch is low, and followed by a very high pitch on the next syllable to the right ([ko]), which falls slightly to mid level by the third and last syllable of the wh-phrase ([ɹe]). The initial syllables of the other words in the sentence (*i.e.* [t̪uj] and [ˈdʑan]) are also realized with low or low-falling pitch. The first sentence in (5), in contrast, has primary stress on the first word [t̪uj] 'you', giving [ˈt̪uj], and not on the wh-phrase [ki koɹe] 'how'. I assume this represents narrow focus on [t̪uj] 'you', which I would write in English orthography as 'How did *you* know?'. In this first version, the stressed syllable [ˈt̪uj] bears a similar pitch contour as the wh-phrase in the second (*i.e.* default wh-focus) version. It is realized with

[23] The italics to show emphasis are my addition.

low/rising pitch, and is followed by a very high pitch on the next syllable [ki], which is followed by steadily falling pitch on [ko] and [ɹe]. Although the segments, word boundaries, and syntactic constituency are different across the focused elements of the two variants of the sentence (*i.e.* [ˈt̪uj] and [ˈkiː koɹe]), the pitch pattern is the same. The stressed syllable of the presumably focused word takes a low pitch, and is followed by a high pitch and steadily falling pitch on the subsequent syllables, regardless of word boundaries and whether or not the affected words form a constituent ([ki koɹe] 'how' is a constituent, but [t̪uj ki koɹe] 'you how' most likely is not). This suggests that focus is realized as some sort of pitch accent (as it is anchored on a stressed syllable) that extends at least one or two syllables to the right (regardless of the existence of any word boundaries). Such examples are also found in the experiments conducted for the current study on Bengali speakers from dialect areas far removed from that of Kolkata, suggesting that it is not specific to Chatterji's Central Bengali dialect. Although Hayes & Lahiri (1991) describes the same dialect as Chatterji (1921), *i.e.* Standard Bengali as spoken in Kolkata, they do not describe data resembling this pitch contour, and in fact their analysis anchors the high pitch on focused words not to the pitch accent but to a phrasal edge (see §8.1).

7.2 Ferguson & Chowdhury (1960)

Ferguson & Chowdhury (1960) were the first to label three prosodic juncture types in the Central Bengali dialect identified as the "Standard Colloquial Bengali of the educated people of Calcutta" (p. 22): terminal juncture (sentence-level, labeled /||/),

phrase juncture (above the word-level, /|/), and microjuncture (below the word-level, /-/). Phrase-final lengthening, minimal word lengthening, hiatus and its resolution, pauses, and segmental alternations are mentioned as cues to these boundaries. As was generally true at that time, prosodic junctures are not related in the analysis to intonational phenomena; intonation is described as whole tonal contours that are not explicitly broken down into constituent parts, regardless of the number of words or word types in the utterance.

7.3 Ray, Hai, & Ray (1966)

7.3.1 Phrasing and interphrasal disjuncture

In Ray, Hai, & Ray's (1966) grammar of Standard Central Bengali, intonation is described using notation for both phrasing and tonal contours. They lay out three "demarcation" (*i.e.* phrasing) symbols (shown below in Table 7) to mark three different interphrasal disjuncture types: /#/ for a "final pause, a stretch of silence with relaxation", /=/ for a "tentative pause, a cessation of movement without relaxation, with a stretch of either silence or nonfunctional voicing", and /+/ for "a momentary dip in loudness coincident with a momentary spurt in speed, [marking] a stress group boundary insofar as that does not coincide with a pause group boundary". These three interphrasal disjuncture types presumably do not correspond directly to the three disjuncture types seen in Ferguson & Chowdhury (1960), as the latter study identified two disjuncture types above the word level, and one within words (*i.e.* the microjuncture). Furthermore, Ray *et al*. state that /#/ and /=/ both demarcate "pause groups" (which they further describe as

“grammatical clauses”), while /+/ demarcates “stress groups”, thus implying the existence of only two phrase/group types, despite the three disjuncture types.

Symbol	Description	Use
#	Final pause, silence	Grammatical clause R-edge
=	Tentative pause, optional silence	Grammatical clause R-edge
+	Momentary dip in loudness	Stress group R-edge

Table 7. Interphrasal disjuncture types identified by Ray *et al.* (1966).

Another three “demarcation” symbols marking other non-tonal suprasegmental phenomena in Ray *et al.*’s model are taken from the International Phonetic Alphabet (IPA): /'/ for “loud stress”, /:/ for the phonetic lengthening of vowels, and /./ for syllable breaks. Stress is not described phonetically or in terms of location within a word; it is described as a relatively strong cue to the beginning of a clause nucleus. It may occur once, twice, or not at all within a “stress group”. The contrastive properties of stress are described only in terms of the phrase level, with the example /uni tay 'khelen/[24] ‘He therefore ate’ contrasted with /uni 'tay khelen/ ‘He ate just that’.

[24] /uni/ = 3RD PERSON HONORIFIC; /tay/ = [t̪aj] ‘(for) that’-FOCUS; /khelen/ = [kʰelen] ‘ate’-HONORIFIC

7.3.2 Pitch phonemes

Ray *et al.* go on to describe the possible pitch contours of Bengali; a pitch contour is defined as "what begins everytime with a pause or a juncture or a stress". It is composed of one to four distinct pitch phonemes strung together across a stress group, and is demarcated by one of the three disjuncture symbols shown above in Table 7. They then list the nine distinct pitch phonemes that can make up a contour, with each pitch phoneme transcribed using numbers and two diacritics (i.e. underline and overline), shown in Table 8 below, along with their description of the phoneme's use:

Pitch phoneme	Description	Use
$\overline{3}$	High rising	Challenges
$\underline{3}$	High falling	Questions[25]
3	High even	Calls, or in storytelling
$\overline{2}$	Mid rising	Frames of reference as tentative queries
$\underline{2}$	Mid falling	Frames of reference as tentative assumptions
2	Mid even	Limiting conditions to an expected statement
$\overline{1}$	Low rising	Requests
$\underline{1}$	Low falling	Final statements
1	Low even	Non-final statements

Table 8. Pitch phonemes identified by Ray *et al.* (1966).

[25] Ray *et al.* note in a later chapter that questions are marked with the "terminal high-to-low glide" (*i.e.* $\underline{32}$) in the Dacca (Dhaka) dialect. This may be relevant to my study, as the subjects in my experiment are all from Bangladesh and are presumably familiar with the Dhaka dialect.

7.3.3 Example sentences transcribed

I reproduce here some of their example sentences [Ray *et al.* §2.2.2 pp. 8-9] with pitch phonemes and disjunctures transcribed along with pitch tracks added by me:[26]

2 3 2 $\overline{3}$
(6) /tumi na + 'jabe =/
[t̪umi na ˈdʑabe]
you PT will.go
'Aren't you supposed to go?'

2 2 $\overline{2}$ 2 3 $\underline{2}$ 2 2 2 1 $\underline{1}$
(7) /ey kOlomTa = jar dam Ek So Taka = ami 'upohar peechi #/
[ej kɔlomta dʑaɹ d̪am ɛkʃo taka ami ˈupohaɹ peetɕʰi]
this pen whose price 100 rupees I gift have.received
'This pen, the price of which is Rs.100, I have received as a gift.'

2 2 2 33$\underline{2}$
(8) /ka:l to + 'skul ache: =/
[ka:l t̪o ˈskul atɕʰe:]
tomorrow PT school exists
'There is school tomorrow (exclamation-cum-explanation).'

Ray *et al*. recognize that the descriptions of the uses of these nine pitch phonemes are only relevant when immediately preceding a pause. Even though most non-final words are transcribed as having rising pitch (usually starting at 2 and ending at 3), there is no discussion of pitch contour patterns that are not immediately prepausal. The authors do describe that, when not immediately prepausal, low pitches ($\overline{1}$, $\underline{1}$, and 1) often signify

[26] Again, the IPA transcription and morpheme-by-morpheme glosses are my own addition.

restraint, while high pitches ($\overline{3}$, $\underline{3}$, and 3) tend to indicate excitement. “Gliding” (*i.e.* contour) tones ($\overline{3}$, $\overline{2}$, $\overline{1}$, $\underline{3}$, $\underline{2}$, and $\underline{1}$) often signify emphasis (presumably “focus”), as shown in the following example [Ray *et al.* §2.2.2 p. 9]:

2 $\overline{3}$ $\overline{3}$

(9) /ˈpritthi:n =/
[ˈpɹit̪ʰ:i:n]
Prithvin
‘Is it really Prithvin?’

By positing three disjuncture sizes and nine pitch phonemes (basically tonal targets) instead of whole contours as the basic units of intonation, Ray *et al.* lay out the basic properties of what would become an autosegmental representation of Bengali intonation.

7.4 Summary

I summarize in Table 9 the prosodic units above the word and inter-unit disjuncture types (greater than a word boundary) described in Chatterji (1921), Ferguson & Chowdhury (1960), and Ray *et al*. (1966). Each cell represents a unit/disjuncture, with the name of the prosodic unit provided on the first line of each cell, and the name of and the symbol for the disjuncture provided below it. When the author does not provide a name for the unit, a dash (–) is given.

<table>
<tr><th>Chatterji (1921)</th><th>Ferguson & Chowdhury (1960)</th><th>Ray et al. (1966)</th></tr>
<tr><td>(no label for unit)
slight pause /||/</td><td>(no label for unit)
terminal juncture /||/</td><td>pause group
final pause /#/</td></tr>
<tr><td rowspan="2">sense group
/|/</td><td rowspan="2">(no label for unit)
phrase juncture /|/</td><td>pause group
tentative pause /=/</td></tr>
<tr><td>stress group
/+/</td></tr>
</table>

Table 9. Summary of prosodic units above the word and disjuncture types described in Chatterji (1921), Ferguson & Chowdhury (1960), and Ray *et al.* (1966).

8 INTONATIONAL PHONOLOGY STUDIES

Building on the Autosegmental-Metrical model of intonation (Pierrehumbert 1980; Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988), Hayes & Lahiri (1991), Lahiri & Fitzpatrick-Cole (1999), Michaels & Nelson (2004), Jun (2005), and Selkirk (2006) all propose Intonational Phonological models of Bengali prosody. These can be set apart from previous models in that they recognize only two underlying tonal targets (*i.e.* L and H) and distinguish between pitch accents and boundary tones, claiming that pitch accents are associated to stressed syllables while boundary tones are associated to phrase edges. In this section, I provide a review of these six studies, followed by a comparison of their claims of phrasing in §8.6.[27]

[27] Truckenbrodt (2003) adopts the model of Bengali intonation put forth in Hayes & Lahiri (1991), and uses data provided in that study as well as data from Fitzpatrick-Cole (1994, 1996) to support his constraints on variable phonological phrasing. As the model I propose in Chapter Three does not make reference to the derivation of variable phrasing, I will not discuss Truckenbrodt's analysis in further detail.

8.1 Hayes & Lahiri (1991)

8.1.1 Overview

The first intonational phonology analysis of Bengali (Kolkata dialect of Central Bengali) is Hayes & Lahiri (1991), which incorporates Beckman & Pierrehumbert's (1986) fundamental distinction between two types of tone: pitch accents and boundary tones. While pitch accents align to prominent syllables, boundary tones associate to the edges of prosodic, and by extension, phonological domains. Hayes & Lahiri's (1991) view of prosodic phrasing draws upon the theory of the Prosodic Hierarchy (Selkirk 1980, Nespor & Vogel 1986, Hayes 1989), which distinguishes several levels of prosodic structure both within individual words and groups of words. In adopting the Prosodic Hierarchy model of phrasing, Hayes & Lahiri (1991) reduce the number of interphrasal disjuncture types from Ray *et al.*'s three (*i.e.* /#/, /=/, and /+/) to two, but adopt Ray *et al*.'s two-phrase (*i.e.* stress group and pause group) model. These two prosodic units in the Hayes & Lahiri (1991) model are labeled the phonological phrase (P-phrase) and the intonation phrase (I-phrase).

The six basic principles of the Hayes & Lahiri (1991) model of Bengali prosody are (1) that Bengali stress attracts pitch accent assignment, (2) that the intonational structure of Bengali draws directly from the Prosodic Hierarchy, (3) that "phrase accents" such as those described in Bruce (1977) can be reanalyzed as boundary tones (as per Beckman & Pierrehumbert 1986), (4) that phrasal stress can be assigned by rule in the absence of

semantic factors that would otherwise assign focus, (5) that tones are subject to phonological rules much in the same way that segments are, and (6) that the tonal inventory of Bengali is constrained by the Obligatory Contour Principle (OCP; Leben 1973, McCarthy 1986), much in the same way it has been shown to constrain surface realizations of tone in languages with lexical tone.

8.1.2 Phrase types

In the Hayes & Lahiri model, words are grouped into P-phrases, which serve as both the domain of segmental alternations and the basic unit of intonational phrasing. A P-phrase may or may not bear a pitch accent – a P-phrase can have zero, one, or multiple pitch accents.[28] P-phrases can occur in either the intonational head or nucleus, a concept associated with the British school of intonation research;[29] the nucleus is the rightmost tonally-marked phrase in an I-phrase, and the head is the collection of all prenuclear phrases. While the nucleus can be made up of more than one P-phrase; however, only the first P-phrase in the nucleus bears a pitch accent and P-phrase-level boundary tone. However, while some previous studies (*e.g.* Chatterji 1921, Ferguson & Chowdhury 1960) analyzed the pitch contours of nuclear phrases as whole units that could not be further broken down, Hayes & Lahiri discovered regular patterns in pitch contours, analyzing them as strings of pitch accents and boundary tones.

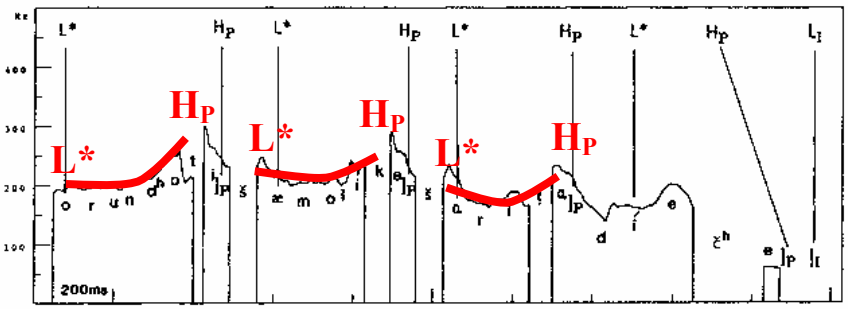
---

[28] Hayes & Lahiri state "[e]ach Phonological Phrase may contain one or more pitch accents (T*), and may end in a boundary tone (T_P)" (§2.5 p. 55). They do not explicitly state that a P-phrase can be devoid of pitch accents, although in their example sentences, this often occurs post-focally (*e.g.* sentence 16, §4.1, p.57).

[29] This is the extent of Hayes & Lahiri's "British school"-isms. They even point out that the head-nucleus distinction is not crucial to their analysis, and that they employ the terms simply for their descriptive convenience.

8.1.3 The head

According to the Hayes & Lahiri (1991) model, the head is the group of all prenuclear P-phrases.[30] An I-phrase might have no head P-phrase, if it is made up of only a nucleus. Each P-phrase in the head bears one low pitch accent (L*), assigned to the most prominently-stressed syllable(s), and a high boundary tone (H_P) marking the right edge of the P-phrase. (Stress, according to Hayes & Lahiri, is consistently word-initial, but phonetically very weak and thus difficult to perceive.) An example utterance including a head consisting of three P-phrases (*i.e.* [ˈɔɹund̪ʱo̪ti ˈʃɛmolike ˈʃaɹita] 'Arundhati…the sari to Shamoli') is provided in (10).



(10)	orundʰoti	šæmolike	šariṭa	díečʰe
	[ˈɔɹund̪ʱo̪ti	ˈʃɛmolike	ˈʃaɹita	ˈd̪ietɕʰe][31]
	Arundhati	to Shamoli	the sari	*gave*

'Arundhati *gave* the sari to Shamoli.'

[30] It is not explicitly stated in Hayes & Lahiri (1991) that the head is the group of all prenuclear P-phrases (as opposed to each individual prenuclear P-phrase), but this is inferred in sentences such as "Where the head contains several P-phrases, each one receives its own rise" (§8.1, p. 77). This is unlike Lahiri & Fitzpatrick-Cole's (1996) definition of a head (each individual prenuclear P-phrase).

[31] The IPA transcription is my own addition. Pitch accented syllables are marked with primary stress (ˈ). The vowel transcribed variously as /æ, E, ê, ɛ/ in previous studies is labeled /ɛ/ in this dissertation in the interest of consistency; I make no claims as to the exact phonetic quality of the vowel across speakers. See the Appendix for a comparison of transcription schemes for Bengali.

In (10), each word constitutes its own P-phrase; the first three P-phrases (*i.e.* [ˈoɹund̪ʱo̪ti] ‘Arundhati’, [ˈʃɛmolike] ‘to Shamoli’, and [ˈʃaɹita] ‘the sari’) make up the head, and the last P-phrase ([ˈd̪ietɕʰe] ‘gave’) is the nucleus. Each P-phrase in the head receives the default head contour, consisting of a low pitch accent (L*) and a high boundary tone (H_P).

8.1.4 The nucleus and the I-phrase

The tonally-minimal utterance in Bengali is an I-phrase made up of a nucleus, bearing only one pitch accent (T*) and one I-phrase boundary tone (T_I), giving the bitonal $[[T^*]_{PP}T_I]_{IP}$. The nuclear pitch accent (once again, aligned to the stressed syllable) can be either low (L*) for interrogatives and focused constituents, or high (H*) for declaratives. P-phrases in the nucleus do not bear the high P-phrase boundary tone (H_P) mentioned earlier, unless the nucleus is under focus; in these cases, the high P-phrase boundary tone (H_P) marks the focused nucleus’s right edge as in **(10)**. This will be described further in §8.1.5.

P-phrases are grouped into Intonation Phrases (I-phrases). The last tonally-marked P-phrase of the I-phrase is always the nucleus, which, in addition to the pitch accents and boundary tone mentioned earlier, can also take one of four I-phrase-level boundary tones or “suffixes” (T_I). The four boundary tones found include a low tone (L_I) for declaratives, a low rising contour (L_IH_I) for continuation, a high tone (H_I) for offerings, and a high falling contour (H_IL_I) for yes/no questions. These four IP-boundary tones (T_I) interact

with the nucleus's pitch accent and possible H_P boundary tone to create a large number of possible sentence-final pitch contours, as shown in Table 10.

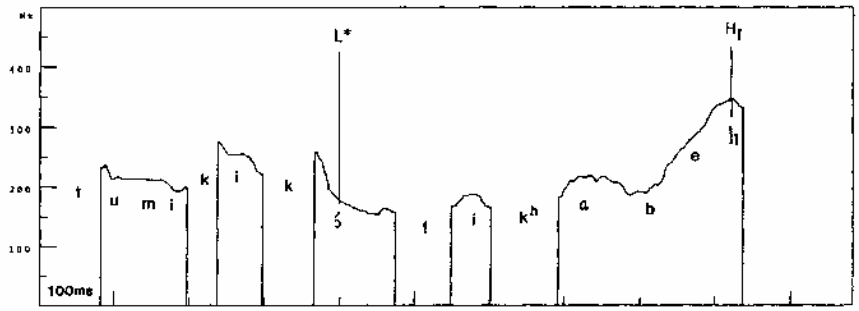
Pitch accent	P-phrase boundary tone	I-phrase boundary tone	Meaning
L*	none	H_I	Offering
L*	none	H_IL_I	Yes/no question
L*	H_P	L_I	Focus
L*	H_P	L_IH_I	Focus with continuation rise
H*	none	L_I	Declarative
H*	none	L_IH_I	Declarative with continuation rise
L+H*	none	L_I	Finality (downstep)
L+H*	none	L_IH_I	Finality (downstep) with continuation rise

Table 10. Nuclear tonal inventory in Hayes & Lahiri's (1991) model.

I include here figures of some of Hayes & Lahiri's examples to illustrate selected nucleus types from Table 10 above. The question in (11) ends in what is called the "offering" nucleus, marked by a low pitch accent (L*) and a high I-phrase boundary tone (H_I).[32] Notice how the pitch rises at a shallow slope until the final word [k^habe], where it rises more sharply.[33]

[32] Note that the object and verb in this sentence are parsed together into a single P-Phrase. See Truckenbrodt (2003) for an account of the optionality in phrasing.

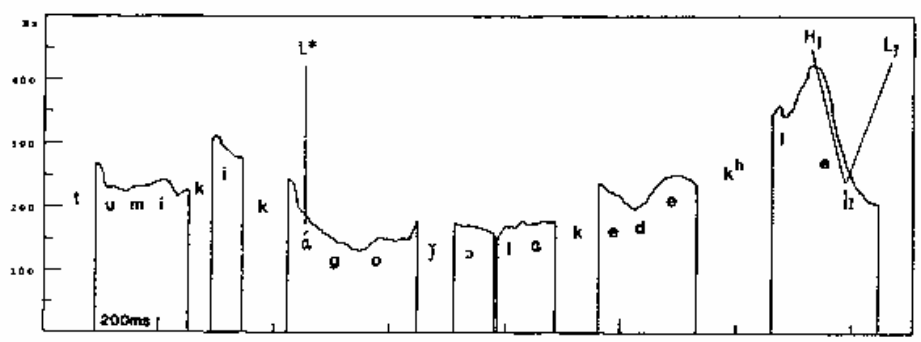
[33] NB: the voiced stop [b] in [k^habe] is a depressor consonant, triggering local pitch lowering.



(11)	tumi	ki	kɔ́fi	kʰabe
	[t̪umi	ki	ˈkɔfi	kʰabe]
	you	Q	coffee	drink

‘Would you like some coffee?’ (offering nucleus; H&L 28, §5.1, p. 66)

Yes/no questions like that in (12) are marked with a low pitch accent (L*) on the nucleus (in this case, *kágoǰɔlake* [kagodʑɔlake] ‘the newspaperman’, as it is under narrow focus, as will be described further on) and a falling I-phrase boundary tone (H_IL_I), realized in (12) on the last syllable of *dekʰle* [d̪ekʰle] ‘saw’.



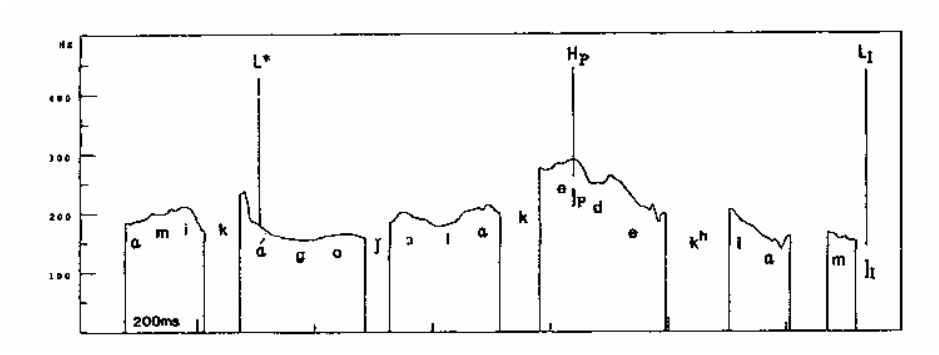
(12)	tumi	ki	kágoǰɔlake	dekʰle
	[t̪umi	ki	ˈkagodʑɔlake	d̪ekʰle]
	you	Q	*newspaperman*	saw

'Did you see the *newspaperman*?' (yes/no nucleus, H&L 16, §4.1, p. 57)[34]

8.1.5 Focus

Focus realization is described in the Hayes & Lahiri (1991) model as a sequence of a low pitch accent (L*) on the main stressed syllable of the P-phrase under focus, followed by a high P-phrase boundary tone (H_P) marking the right edge of the phrase under focus, and a low I-phrase boundary tone (L_I). Thus, a P-phrase under focus (which causes it to be a nuclear P-phrase) looks exactly like a P-phrase in the head in terms of the underlying tonal targets and their alignment; the main difference is the lack of additional tones between the focused P-phrase and the I-phrase boundary tone. As the nucleus bears phrasal stress, the realization of nuclear tones is more exaggerated (lower L*, higher H_P).

[34] The high P-phrase tone (H_P) tone is deleted in this example due to the existence of a following H tone (*i.e.* H_IL_I). See §8.1.6 for a description of H_P deletion.

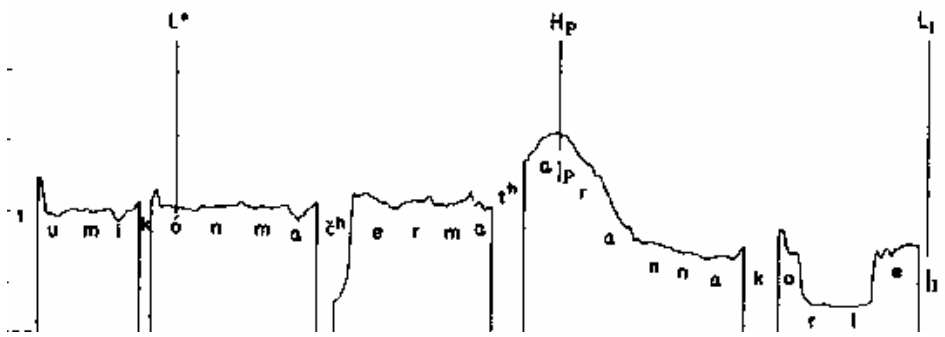


(13) ami kágoǰɔlake dekʰlam
[ami ˈkagodʑɔlake d̪ekʰlam]
I *newspaperman* saw
'I saw the *newspaperman*.' (focus nucleus, H&L 18, §4.2, p. 59)

The sentence in (13) illustrates an example of focus realization in a declarative sentence. The constituent under narrow focus, *kágoǰɔlake* [ˈkagodʑɔlake] 'newspaperman-ACC', bears the declarative focus nuclear tone pattern: the pitch starts at a low pitch accent (L*) on the main stressed syllable [ˈka], steadily rising to a high boundary tone marking the ridge edge of the focused P-phrase (*i.e.* at the right edge of the last syllable [ke]), and then drops smoothly across the tonally-unmarked P-phrase *dek^{h}lam* [d̪ekʰlam] 'I saw' towards the low I-phrase boundary tone (L_I).

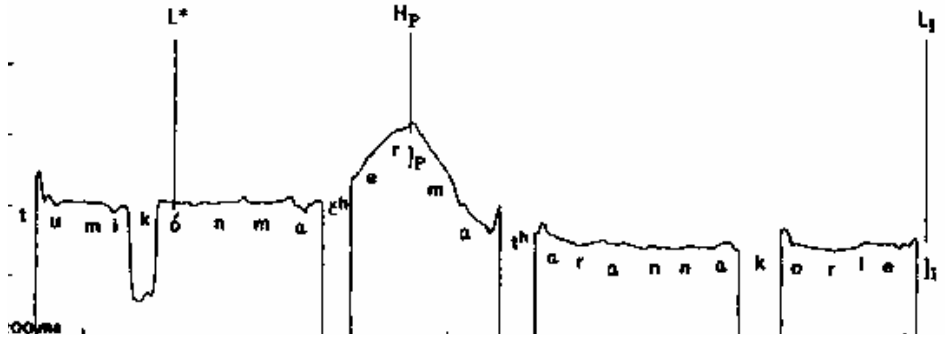
Focus domains spanning more than one word are treated as if they were made of only one very long word; since the focus nuclear tone pattern involves one low pitch accent (L*) on the domain's strongest (*i.e.* leftmost nonclitic) word, one high boundary tone (H_P) at the domain's right edge, and a low boundary tone (L_I) at the end of the utterance, the same frame can be stretched to fit any focus domain, regardless of size. Compare the tones projected on the larger focus domain [kon matɕʰeɹ mat̪ʰa] 'which fish-

head' in (14) to the tones in the smaller domain [kon matɕʰeɹ] 'which fish's' of (15) below.



(14) tumi	kón	mačʰer	mat̪ʰa	ranna	korle
[t̪umi	kon	matɕʰeɹ	mat̪ʰa	ɹan:a	koɹle]
you	which	fish's	head	cook	did

'*Which fish-head* did you cook?' (H&L 20, §4.2, p. 60)[35]



(15) tumi	kón	mačʰer	mat̪ʰa	ranna	korle
[t̪umi	kon	matɕʰeɹ	mat̪ʰa	ɹan:a	koɹle]
you	which	fish's	head	cook	did

'*Which fish*-head did you cook?' (H&L 21, §4.2, p. 61)

[35] NB: The sharp drop and rise in the pitch contour during the word [koɹle] 'did' is likely due to a tracking error and not to an actual change in the F_0 of the speaker's voice.

The tone pattern that frames the focused constituent (L*…H_P) in Hayes & Lahiri (1991) makes a prediction that all focused elements should end in a high boundary tone (H_P), regardless of the size of the constituent. I come back to this point when I discuss the data from the current study in Chapters Three and Four.

8.1.6 OCP effects

Hayes & Lahiri note that there are no tonal patterns in Bengali that include sequences of the same tone (*e.g.* two high tones in a row); they claim that this is not an accidental gap, but that the tones that make up a grammatical utterance in Bengali must obey the Obligatory Contour Principle (OCP) (Leben 1973, McCarthy 1986). This principle monitors the pitch contour, and forces the deletion of certain tones in order to avoid sequences of underlying identical tones. The examples shown all include sequences of high (H) tones repaired to include only one high (H) tone, although there is no reason to believe that sequences of low (L) tones are not also subject to the OCP.

In (11) above, the offering nucleus is described as a low pitch accent (L*) on the initial syllable of the object, rising to a high I-phrase boundary tone (H_I) at the end of the verb. The lack of a P-phrase boundary tone for the object *kɔ́fi* [ˈkɔfi] ‘coffee’ is explained as an effect of the OCP. Since the P-phrase boundary tone of Bengali is a high tone (H_P), and the next tone to the right of the P-phrase *kɔ́fi* is the high I-phrase boundary tone (H_I), there is an underlying sequence of two high (H) tones. Thus, the OCP keeps this sequence from being produced by triggering what is called H_P deletion, which, not

surprisingly, deletes high P-phrase boundary tones (H_P) when adjacent to another high (H) tone. This is schematized in (16) below.

(16) L* H_P H_I

H_P deletion ↓

L* Ø H_I

tumi-ki kɔ́fi k^habe
[ṭumi ki ˈkɔfi k^habe]
you Q coffee drink

'Would you like some coffee?' (H&L 28, §5.1, p. 66)

Had H_P deletion not occurred in (16), we would have expected a high plateau between the high P-phrase boundary tone (H_P) and the high I-phrase boundary tone (H_I). Hayes & Lahiri explain that long plateaus of high pitch are never found in their data; for example, a focus nucleus (L* H_P) followed by a tonally unmarked verb and a yes/no question boundary tone (H_IL_I) is not realized as a pitch rise on the focused constituent, a plateau towards the I-phrase edge, and a final fall, as would be expected given those tones. Instead, the OCP triggers H_P deletion, thus leaving us with a low pitch accent (L*) on the focused element, rising smoothly through the focused word and the following verb, reaching a peak on the last syllable of the sentence and sharply dropping thereafter. This means that the statement shown in (12) above is ambiguous: the speaker could be asking either 'Did you see the newspaperman?' (with no focus) or 'Did you see *the newspaperman?*' (with narrow focus on the object). This is schematized in (17) below.

(17) L* H_P H_IL_I

H_P deletion ↓

L* Ø H_IL_I

tumi-ki [*kágoǰɔlake*]FOC dek^{h}le?
[t̪umi ki ˈkagodʑɔlake d̪ek^{h}le]
you Q *newspaperman* saw

'Did you see the *newspaperman*?' (yes/no nucleus, H&L 16, §4.1, p. 57)

Claims of Hayes & Lahiri's (1991) study of Kolkata Bengali that are not held in the current model's study of Bangladeshi Standard Bengali include (1) the restriction on sequences of adjacent H tones (resolved by the deletion of the P-phrase boundary tone H_P), (2) the restriction against high plateaus of pitch, (3) the use of a tonal "frame" demarcating the focus domain, (4) the properties of the P-phrase with respect to how many pitch accents it can contain and what boundary tones are projected, (5) the distribution of the high pitch accent (H*), (6) the lack of contour pitch accents,[36] (7) and the distinction between focused declaratives and interrogatives.

8.1.7 Other issues

Hayes & Lahiri do not explicitly discuss the phenomenon of tonal deletion that appears to occur post-focally in their data. In (17), for example, the verb *dek^{h}le* bears no

[36] Hayes & Lahiri (1991) do, however, posit a pitch accent that is formally composed of two tones (L+H*). Although it uses two tonal targets in its label, this pitch accent does not represent a contour pitch accent, but a downstepped high pitch accent, and is thus not considered an exception to Hayes & Lahiri's claim that Bengali lacks contour pitch accents. This diacritic use of bitonality is also used for downstep in previous studies of English (Pierrehumbert 1980).

pitch accent, although it is parsed as its own P-phrase (as is explicitly shown in the diagram on p. 58 of H&L 1991). Hayes & Lahiri state that a focused constituent should receive the greatest metrical prominence in its I-phrase, and that the greatest metrical prominence in an I-phrase should be the rightmost prominence. In this way, the lack of a pitch accent on the verb *dek^{h}le* does not have to be explained in further detail.

The authors also discuss phonological phrasing and its dual role as a domain of intonational contours and a domain of segmental rule application. As the segmental rules they cite (/r/-assimilation and voicing assimilation) are not regularly applied in the speech of the subjects participating in the current study, I am unfortunately unable to discuss these phenomena for the current model.[37]

8.2 Lahiri & Fitzpatrick-Cole (1999)

Lahiri & Fitzpatrick-Cole (1999) extend the initial model set in Hayes & Lahiri (1991) by examining aspects of focus realization and P-word boundaries in Kolkata Bengali. They posit that emphatic clitics *–i* and *–o*, which attract focus to the words onto which they attach, are lexically specified for a high pitch accent (H*). These lexical tones

[37] This may be due to dialect differences; although the subjects in the current study and those in Hayes & Lahiri (1991) identify as speakers of Standard Bengali, many of the current study's subjects are from various parts of East Bengal (*i.e.* Bangladesh), primarily from areas where Northern and Eastern Bengali dialects are spoken. Thus, the subjects of the current study presumably speak (Northern or Eastern) Bangladeshi Standard Bengali, while Hayes & Lahiri (1991) describe the speech of speakers from Kolkata, in India's West Bengal, whose dialect could be labeled Central, Indian, or Kolkata Standard Bengali.

interact with the postlexical pitch accents and boundary tones of the language, and create otherwise impossible contours when they occur p-phrase-medially.[38]

The inventory of pitch accents and boundary tones in Lahiri & Fitzpatrick-Cole (1999) is not drastically different from that of Hayes & Lahiri (1991). Apart from some subtle differences in pitch accent use, Lahiri & Fitzpatrick-Cole also avoid explicitly building nuclear tonal patterns as whole units, instead opting to leave pitch accents and i-phrase boundary tones separate in the model. For example, while Hayes & Lahiri describe the contour for yes/no question nuclei as a whole unit (L* H_IL_I), Lahiri & Fitzpatrick-Cole break down this nuclear contour into a "neutral question accent" (L*) and "yes/no question boundary tone" (H_IL_I). The high p-phrase boundary tone (H_P) used in the head and in focused constituents, however, is still considered part of the "stem" or accent of the focused phrase, and is not decomposed explicitly. The full set of nuclear accents in Lahiri & Fitzpatrick-Cole (1999) is shown below in Table 11, followed by the set of boundary tones in Table 12 (L&F-C 4, §2, p. 121).

Accents ("stems")	Use
L*	Neutral question accent
H*	Neutral declarative accent
L* H_P	Focus accent

Table 11. Nuclear pitch accent inventory in Lahiri & Fitzpatrick-Cole's (1999) model.

[38] Unlike Hayes & Lahiri, Lahiri & Fitzpatrick-Cole do not capitalize the terms "p-phrase" and "i-phrase".

Boundary tones ("affixes")	Use
L_I	declaratives
L_IH_I	continuation rises
H_I	offerings
H_IL_I	yes/no questions

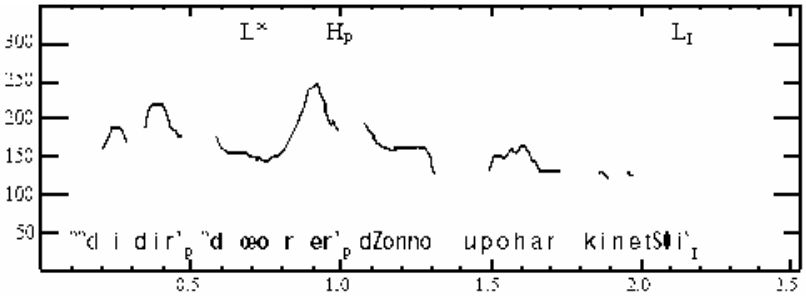
Table 12. Nuclear boundary tone inventory in Lahiri & Fitzpatrick-Cole's (1999) model.

Like the Hayes & Lahiri (1991) model, the Lahiri & Fitzpatrick-Cole (1999) model distinguishes p-phrases in the head from those in the nucleus (with a subtle difference – Lahiri & Fitzpatrick-Cole label each individual prenuclear p-phrase as a separate "head", instead of grouping them into a set of phrases called a "head"); while nuclear p-phrases can take any of the tonal patterns shown in Table 11 above, head p-phrases (which are, by definition, not focused) can only take the L* H_P contour described in Hayes & Lahiri (1991). The two models also largely agree on the role of the OCP, H_P deletion, the lack of plateaus in pitch, the lack of contour tones, and a number of other phenomena.

Lahiri & Fitzpatrick-Cole begin their discussion of emphatic clitics *–i* and *–o* (which add the meaning 'only' and 'also', respectively, to the words they attach to, *e.g.* [ɹɔt̪on-**i** dʐabe] '*Ratan* will go' or '**Only** *Ratan* will go'; [ɹɔt̪on-**o** dʐabe] '*Ratan* will go, **too'** or '**Even** *Ratan* will go') and their interaction with focus with a description of how emphatic clitics attach to full p-phrases (as opposed to bound stems, *etc.*). They proceed to show how to test whether or not a stem is a p-phrase, using evidence from reduplication facts, /r/-assimilation, voicing assimilation, and prosodic lengthening. As I mentioned earlier, my speakers do not regularly apply voicing assimilation or /r/-assimilation as described in

Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole (1999), and thus I will simply adopt the latter authors' claim that emphatic clitics attach to full p-phrases.

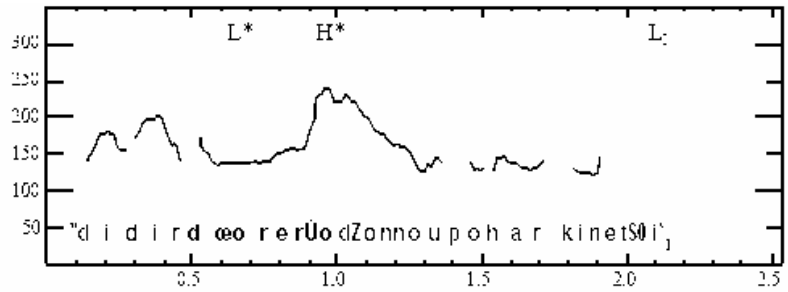
To describe how emphatic clitics trigger focus on their respective arguments, the authors show examples of plain narrow focus (without clitics) as described in Hayes & Lahiri (1991). They show that the tonal contour (*i.e.* L*…H_P) that frames the focused constituent in answers to *wh*-questions, as shown in (18) below is identical to the contour framing a constituent hosting an emphatic clitic, as shown in (19).



(18) didir [dæorer]$_{FOC}$ dZonno upohar kinetS^{h}i
[d̪id̪iɹ d̪ɛ̯ɹɐɹ dʑonːo upohaɹ kinetɕ^{h}i]
sister's brother-in-law's for gift I bought
'I bought a present for sister's *brother-in-law*.' (L&F-C 27b, §4, p. 133)
Answer to: 'Which relative of sister's did you buy a present for?'

Note how the focused constituent *dæorer* [d̪ɛ̯ɹɐɹ] 'brother-in-law-GEN' in (18) above begins with a low pitch accent (L*) and rises sharply at the end to a high p-phrase boundary tone (H_P) at the right edge of the word. Compare this to the focused constituent *dæorer-o* [d̪ɛ̯ɹɐɹo] 'brother-in-law-GEN-ALSO' in (19) below – the contour is largely

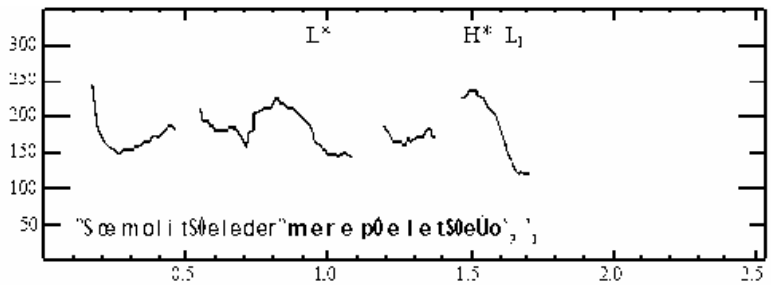
identical, regardless of the type (*i.e.* wh-answer; also) or source (*i.e.* elicited by preceding question; addition of an emphatic clitic) of focus.



(19) didir [dæorer-o]$_{FOC}$ dZonno upohar kinetS^{h}i
[d̪id̪iɹ d̪ɛɔ̯ɹeɹo dʑon:o upohaɹ kinetɕ^{h}i]
sister's **also**-b-in-law's for gift I bought
'I bought a present for sister's *brother-in-law*, **too**.' (L&F-C 28b, §4, p. 134)
Presupposition: I bought a present for sister's *x*, *x* ≠ brother-in-law.

Given these two examples, it would appear that the addition of an emphatic clitic onto a p-phrase simply triggers the same focus realization seen in cases of focus without the clitic. Lahiri & Fitzpatrick-Cole counter this hypothesis by showing cases where the emphatic clitic is attached to the middle of a focused constituent, which is possible if the focused constituent is made up of more than one word, or one morphologically complex word, where the stem of the focused constituent to which the clitic attaches could potentially be an independent p-phrase (independently determined by examining reduplicative patterns and other phenomena). For example, the phrase *mere p^{h}eletS^{h}e* [meɹe p^{h}eletɕ^{h}e] 'beat to death' is made up of the perfective infinitival verb *mere* 'having beaten' and the finite verb *p^{h}eletS^{h}e* 'has thrown'; these two words are obligatorily parsed as one p-phrase (and the segmental and suprasegmental phenomena confirm this), in

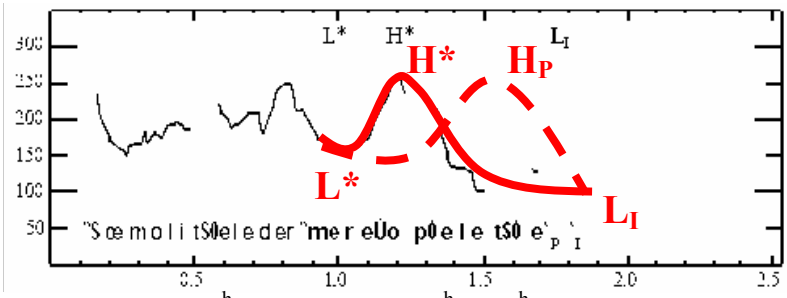
order to be interpreted as meaning ‘beat to death’ as opposed to ‘having beaten, (s)he threw’. Still, because the stem *mere* can in other situations be a p-phrase on its own (with the meaning ‘having beaten’), the emphatic clitic *–o* ‘also’ is permitted to attach to it, leaving the second half of the compound (*i.e.* $p^h eletS^h e$) to the right of the clitic. Thus, the pitch tracks in (20) and (21) below are both ways to produce ‘**also** *beat to death*’.



(20) Sæmoli tS^heleder [mere p^heletS^{h}e-o]$_{FOC}$
[ʃɛmoli tɕ^{h}eleḍeɹ meɹe p^{h}eletɕ^{h}eo]
Shyamoli the boys beaten threw-**also**
‘Shyamoli **also** *beat* the boys *to death*.’ (L&F-C 30b, §4, p. 137)

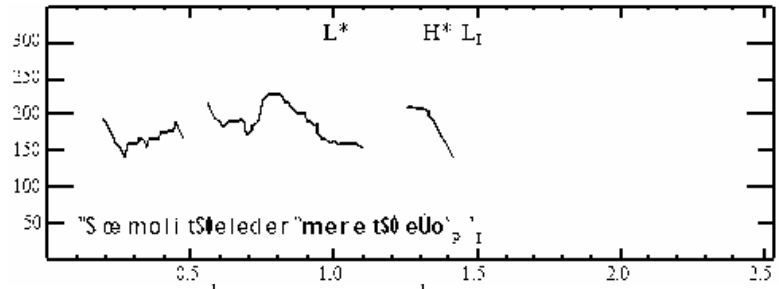
In (20), the focused compound verb *mere* $p^h eletS^h e$ shows the expected focus realization: a low pitch accent (L*) on the initial syllable *me*, and a rising pitch towards the right edge of the focus domain. The final fall is due to the low i-phrase boundary tone (L_I). Once again, it appears that the clitic *–o* only triggers the normal focus realization onto its host word. However, Lahiri & Fitzpatrick-Cole show that this is not the case; if the clitic *–o* attaches to the stem of the compound (*i.e. mere*), the focus domain is still understood to last through the end of the compound (*i.e.* all the way through $p^h eletS^h e$), although the peak in pitch no longer occurs near the end of this domain, as in (21) below. This suggests that the high pitch is attracted to the clitic *–o*, and not to the right edge of

the focus domain *mere-o p^{h}eletS^{h}e*. The otherwise expected pitch contour, with the pitch maximum borne on the right edge of the focus domain, is traced with a dotted line.



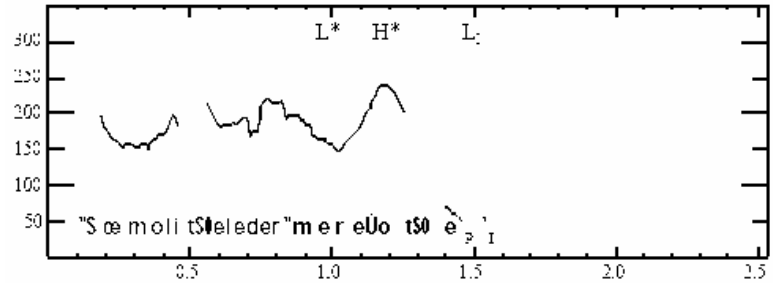
(21) Sæmoli tS^{h}eleder [mere-o p^{h}eletS^{h}e]$_{FOC}$
[ʃɛmoli tɕ^{h}eleḏeɹ meɹeo p^{h}eletɕ^{h}e]
Shyamoli the boys beaten-**also** threw
'Shyamoli **also** *beat* the boys *to death*.' (L&F-C 30c, §4, p. 137)

Similar patterns are seen even within words such as *mere-tS^h-e-**o*** (beat-PERF-3RD-**also**) and *mere-**o**-tS^h-e* (beat-**also**-PERF-3RD), which both mean '(s)he has beaten'. In this case, the clitic can either attach to the full word with stem and affixes attached, or to the verb stem that could potentially be its own p-phrase (*i.e. mere* 'beaten'), thus stranding the bound affixes (*-tS^h* 'PRESENT' and *–e* '3RD PERSON') to the right of the clitic, as shown in (22) and (23) below.



(22) Sæmoli tS^{h}eleder [mere-tS^{h}e-o]$_{FOC}$
[ʃɛmoli tɕ^{h}eleḍeɹ meɹetɕ^{h}eo]
Shyamoli the boys beat-**also**
'Shyamoli **also** *beat* the boys.' (L&F-C 29b, §4, p. 136)

As with (20), the verb in (22) appears to show the default focus pitch contour, with a low pitch accent (L*) on *me*, followed by a rise in pitch towards the end of the focus domain, followed by a fall due to the low i-phrase boundary tone (L_I). This is contrasted with the verb in (23) below, where the clitic *–o* attaches to the stem, and appears to bring the high pitch with it, leaving the following suffixes with low, falling pitch due to the low i-phrase boundary tone (L_I).



(23) Sæmoli tS^{h}eleder [mere-o-tS^{h}e]$_{FOC}$
[ʃɛmoli tɕ^{h}eleḍeɹ meɹeotɕ^{h}e]
Shyamoli the boys beaten-**also**
'Shyamoli **also** *beat* the boys.' (L&F-C 29c, §4, p. 136)

Lahiri & Fitzpatrick-Cole explain this attraction of high tone to the emphatic clitic by positing two different focus realizations: the "focus frame" pattern (*i.e.* L*…H_P) described in Hayes & Lahiri (1991) is used for focus domains that do not include emphatic clitics, while a different pattern is proposed for focus domains that do include emphatic particles. This second pattern involves a low pitch accent (L*) on the initial syllable (as with the first type of focus realization), rising up to a high pitch accent (H*) introduced by the clitic itself. The high boundary tone (H_P) is still underlyingly present, but undergoes H_P deletion triggered by the OCP violation – the existence of two adjacent high tones (H* and H_P). This is schematized in (24) below.

(24)

L* H* H_P L_I

H_P deletion ↓

L* H* Ø L_I

Sæmoli tS^{h}eleder [mere-o-tS^{h}e]$_{FOC}$

[ʃɛmoli tɕ^{h}eleḍeɹ meɹeotɕ^{h}e]

Shyamoli the boys beaten-**also**

'Shyamoli **also** *beat* the boys.' (L&F-C 29c, §4, p. 136)

This analysis is not uncontroversial; Lahiri & Fitzpatrick-Cole acknowledge that it is unusual for a non-tonal language to encode a particular tone in the lexical representation of exactly two morphemes (*i.e.* the high pitch accent on the emphatic clitics). However, they compare this explanation with other analyses involving recursive prosodic structure and floating tones, and find the lexical tone analysis to be superior. Although the high tone on the clitics is labeled as a high *pitch accent* (H*), it is not associated with stressed syllables in the same way that the other pitch accents in the language are; Lahiri & Fitzpatrick-Cole explicitly state that the clitics are not stressed and could not attract a

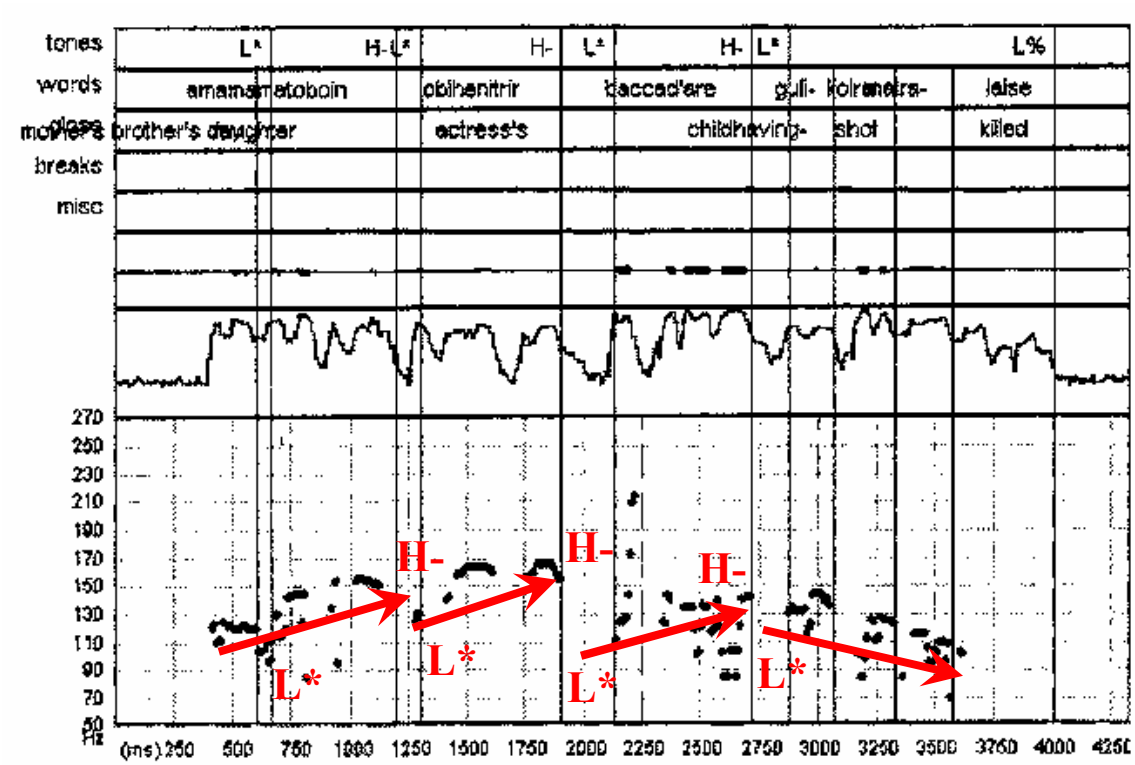
hypothetical postlexical high pitch accent (H*). By proposing both L*…H_P and L*…H*, the authors suggest that there are two related types of focus marking that, despite their surface similarity in sentences where focus particles coincide with p-phrase boundaries, explain the early focus high tone realization when focus particles and p-phrase boundary do not align.

8.3 Michaels & Nelson (2004)

In an unpublished term paper for an Intonation course at UCLA, Michaels & Nelson (2004) set up an Intonational Phonological model of Eastern Bengali, based on a corpus of 91 sentences collected from one speaker. They assume that only stressed syllables bear pitch accents, but that other syllables may bear phrasal and boundary tones. Michaels & Nelson find two pitch accents and three tonally-marked prosodic units in their data, described in more detail below.

As was found in Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole's (1999) description of Standard Kolkata Bengali, Michaels & Nelson find a rising pitch contour repeated over the course of each sentence in Eastern Bengali, as illustrated below in **(25**). This default rise is analyzed as a low pitch accent (L*), borne on the stressed syllable of the word, and high phrase accent boundary (H-), realized as rising pitch across the remainder of the phrase. Each rising contour (L*…H-) is considered an intermediate phrase (ip), equivalent to the phonological phrase (P-phrase). Given the rhythmic nature of this rising ip contour, Michaels & Nelson speculate that the number of ips in the same

sentence can vary depending on factors such as speech rate, although they do not provide evidence to support this hypothesis.



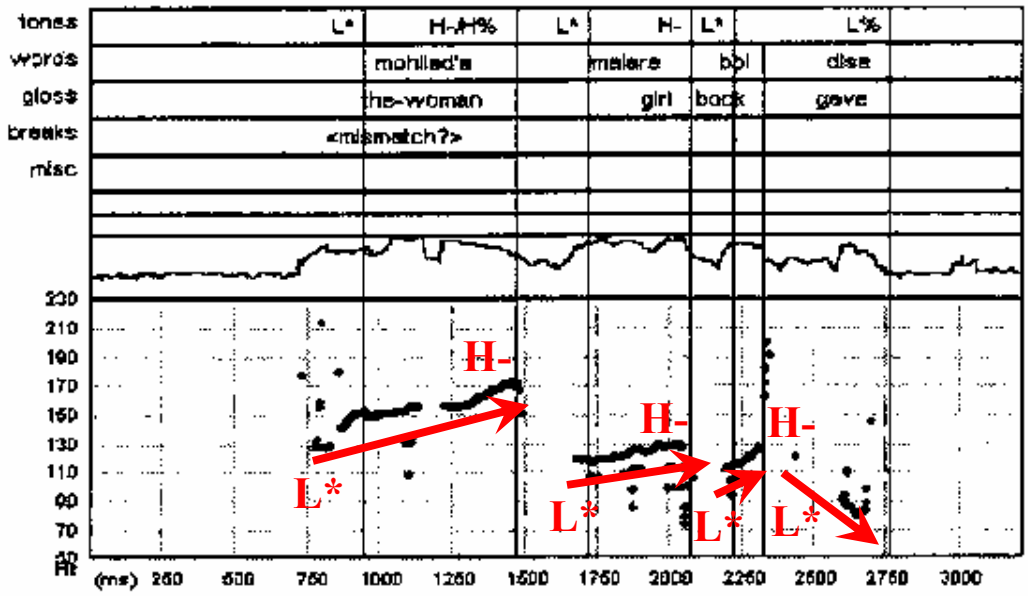
(25) amar mamatoboin obhinetrir baccad'are guli-koira maira-laise
[amaɹ ˈmamat̪obojn ˈobʰinet̪ɹiɹ ˈbatɕːadaɹe ˈgulikojɹa majɹalajse]
my cousin actress's the child having shot has killed
'My cousin killed the actress's child.'[39] [M&N 3, §2.1, p. 6]

Note in (25) that all ips bear a low pitch accent (L*); Michaels & Nelson agree with previous studies that L* is the default pitch accent. Also confirming previous studies, Michaels & Nelson find that the high phrase accent (H-) is not found on the final phrase of the sentence. Instead, the contour of the final ip is generally falling; Michaels &

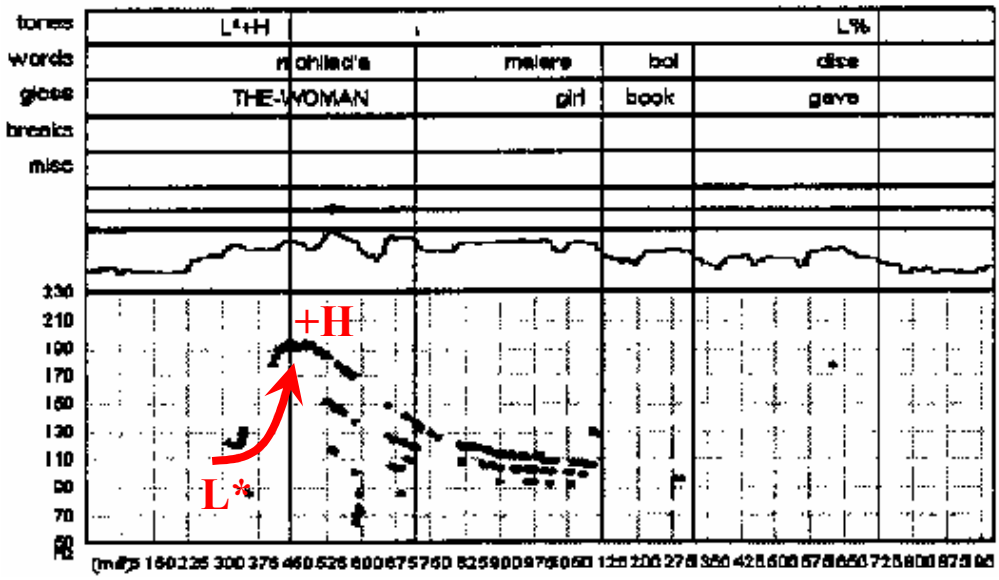
[39] A more accurate translation of this sentence might be 'My cousin [specifically "mother's brother's daughter"] has shot and killed the actress's child.'

Nelson agree with previous studies that the contour of the final ip falls towards a low intonation phrase (ip) boundary tone (L%), associated with declarative sentences. While Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole (1999) attribute the difference between the contours of final and non-final phrases to a fundamental distinction between the tonal inventories of head and nucleus P-phrases, Michaels & Nelson do not posit an underlying difference to final and non-final phrases, but propose that the high phrase tone (H-) of the final ip is overridden by the boundary tone of the intonation phrase (IP), disallowing "stacked" boundary tones such as H-L%. They claim that the tone of a smaller prosodic unit will always be overridden by the tone of a larger prosodic unit if they are aligned to the same location in the text. Although Hayes & Lahiri describe tonal interaction between two high boundary tones *in sequence*, Michaels & Nelson are the first to mention *concurrent* tonal overriding.

Michaels & Nelson find an additional pitch accent not explicitly described in any previous analysis: the rising pitch accent (L*+H). Michaels & Nelson claim that this pitch accent is used only on focused constituents, including wh-words. It is characterized phonetically by rising pitch during the stressed syllable and expansion of pitch range, often leading to higher pitch on the trailing +H portion of the L*+H than the high phrase accent H-. Compare the pitch tracks of the two productions of [mohilada majaɹe boj d̪ise] 'The woman gave the girl the book' under neutral focus (26) and with narrow focus on the subject [*mohilada*] '*the woman*' (27). Notice how the default sequence of ips bearing smooth rising contours in (26) is replaced by a very different pattern in (27).



(26) mohilad'a maiaɹe boi dise
[ˈmohilada ˈmajaɹe ˈboj ˈd̪ise]
the woman girl book gave
'The woman gave the girl the book.' [M&N 21, §2.6.1, p. 22]



(27) *mohilad'a* maiaɹe boi dise
[ˈ*mohilada* majaɹe boj d̪ise]
the woman girl book gave
'*The woman* gave the girl the book.' [M&N 22, §2.6.1, p. 22]

Even at first glance, it is clear that the rising pitch contour on the focused constituent [*mohilada*] '*the woman*' does not resemble the rising pitch contour on focused constituents described in models of Kolkata Bengali. The focused constituent [*mohilada*] '*the woman*' bears the rising pitch accent (L*+H), realized as strongly rising pitch during the stressed syllable [mo], reaching a peak well before the edge of the word. This contrasts with the focus contour (*i.e.* L*…H_P) of Hayes & Lahiri's (1991) and Lahiri & Fitzpatrick-Cole's (1999) data (*cf.* examples 14, 20), and also with Michaels & Nelson's default pattern of Eastern Bengali (*i.e.* L*…H-), where the highest pitch in the word is realized at its right edge. Also peculiar to this pitch accent is the requirement that it not be followed by any additional pitch accents within the IP. This combination of pitch expansion during the focused word and following lack of pitch accents (*i.e.* post-focal

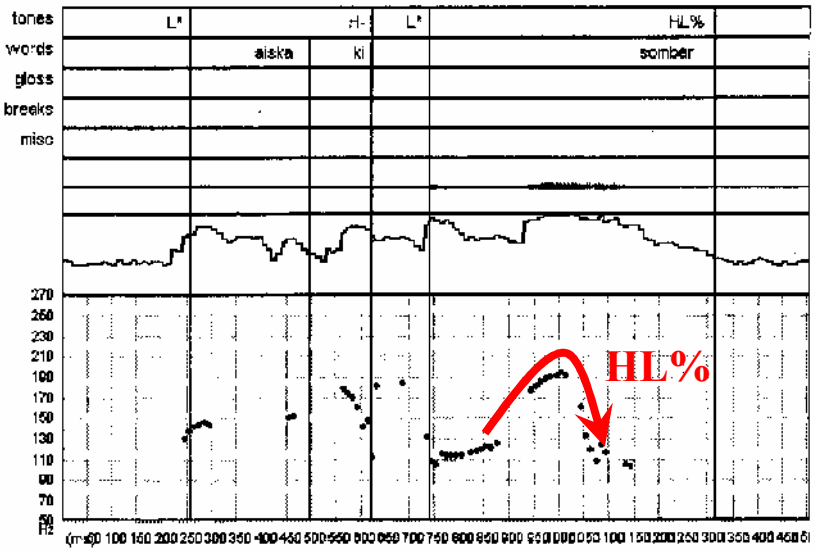
tone deletion) serves to highlight the focused word both metrically and in terms of phrasing. Note that although this "early rise" and its label (L*+H) are not explicitly mentioned in previous analyses of Bengali—the focus contours presented in Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole (1999) do not even resemble those of Michaels & Nelson (2004)—the pitch contours described in Chatterji's (1921) example of focus domain variation illustrated in (5) appear to match the patterns observed by Michaels & Nelson. Ray *et al.* (1966) also associate contour tones with focus, although they do not provide examples. The two pitch accent types found in Michaels & Nelson (2004) and their respective uses are summarized in Table 13.

Pitch accent	**Use**
Low (L*)	Neutral
Rising (L*+H)	Focused (includes wh-words)

Table 13. Pitch accent types described in Michaels & Nelson (2004).

Michaels & Nelson assume a two-phrase model of tonally marked structure for Eastern Bengali, as Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole (1999) do for Standard Kolkata Bengali. The intonation phrase (I-phrase) and phonological phrase (P-phrase) of the latter models roughly correspond to Michaels & Nelson's intonation phrase (IP) and intermediate phrase (ip), respectively. As previously mentioned, non-final ips are always marked by a high phrase accent (H-); IPs, on the other hand, can bear one of three boundary tones—low (L%), high (H%), and falling (HL%)—reflecting the sentence type.

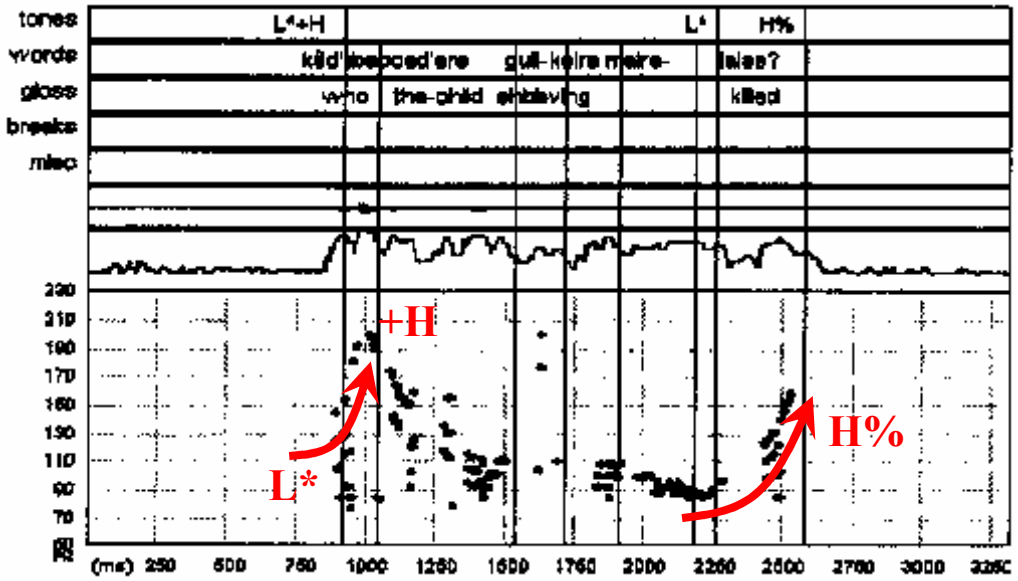
As shown in previous examples, declarative sentences bear the low boundary tone (L%). Yes/no questions, however, are marked with a falling IP boundary tone (HL%), confirming the findings of Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole (1999). This boundary tone is characterized in Michaels & Nelson (2004) by an F_0 rise followed by a sharp fall during the IP-final syllable, as shown below in (28).



(28) aiska ki śombar
['ajska ki 'ʃombaɹ]
today Q Monday
'Is today Monday?' [M&N 13, §2.3.2, p. 15]

In contrast to declaratives and yes/no questions, wh-questions in Michaels & Nelson's model can bear one of two IP boundary tones—high (H%) and low (L%)—depending on whether or not the sentence includes neutral or narrow focus.[40] Neutral wh-questions take the high boundary tone (H%), as illustrated below in (29).

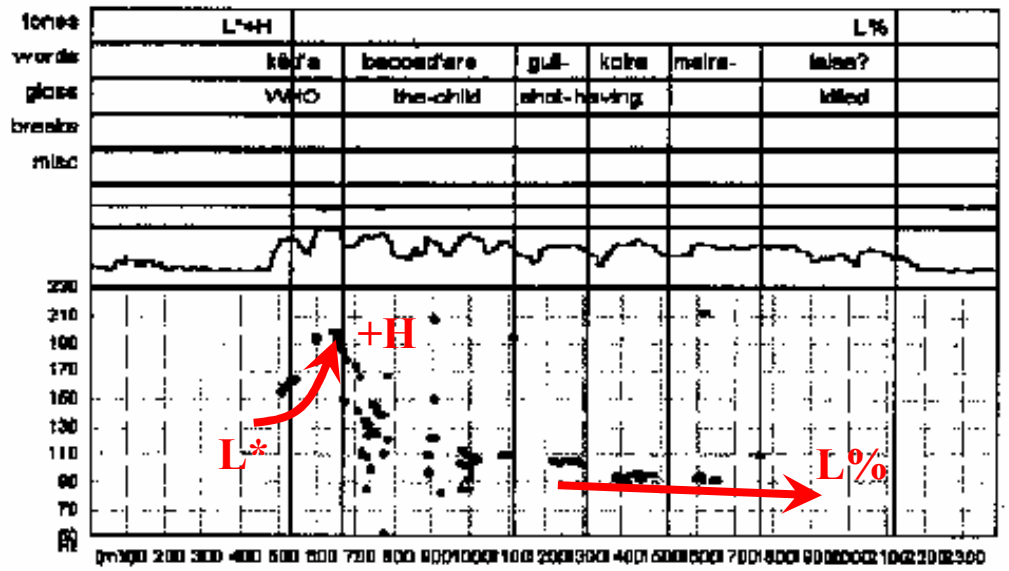
[40] The factors influencing when one would use a narrow focus *vs.* a neutral focus wh-question are complex. In some situations, prior context seems to trigger narrow focus (L%): if Speaker A sees Speaker B walking



(29) kêd'a baccad'aɹe guli-koira maira-laise
[ˈkɛda batɕːadaɹe gulikojɹa majɹalajse]
who the kid having shot killed
'Who shot the child?' [M&N 26, §2.6.2, p. 27]

Narrow-focused wh-questions, however, bear the same low boundary tone (L%) borne on focused declaratives, as shown below in (30).

out a door, she may ask [ko̪tʰaɛ̯ dʑaɔ̯] 'Where are you going?' using neutral focus (H% in the Michaels & Nelson model). However, if Speaker B begins by saying [dʑatɕʰːi] 'I'm going', Speaker A may ask [ko̪tʰaɛ̯ dʑaɔ̯] 'Where are you going?' with narrow focus (L%). In other situations, however, a speaker may use narrow focus (L%) even without any prior speech, *e.g.* upon entering a room and finding a broken window, a speaker may exclaim [dʑanalata ke bʱaŋlo] with narrow focus (L%).



(30) KÊD'A baccad'aɹe guli-koira maira-laise
[ˈkɛda batɛːadaɹe gulikojɹa majɹalajse]
WHO the kid having shot killed
'WHO shot the child?' [M&N 27, §2.6.2, p. 27]

The three IP boundary tone types and their respective uses, as observed by Michaels & Nelson (2004), are summarized below in Table 14.

IP boundary tone	Uses
Low (L%)	Neutral declaratives Focused declaratives Focused wh-questions
High (H%)	Neutral wh-questions
Falling (HL%)	Yes/no questions

Table 14. Boundary tone types described in Michaels & Nelson (2004).

In summary, the Michaels & Nelson (2004) model of Eastern Bengali prosody resembles those of other varieties of Bengali in its adoption of two layers of tonally-marked prosodic structure (*i.e.* IP, ip), where each prosodic unit bears a tone at its right

boundary (*i.e.* T-, T%). The smaller phrase (*i.e.* the intermediate phrase or ip) hosts one pitch accent, which is low (L*) in most cases. The Michaels & Nelson model can be distinguished from previous studies, however, in its analysis of focus. While studies of other varieties of Bengali attribute the rising contour on focused constituents to a low pitch accent and high boundary tone, Michaels & Nelson propose a bitonal pitch accent (L*+H) for focused constituents in East Bengali. Furthermore, Michaels & Nelson are the first find a distinction between neutral wh-questions, which bear a high IP boundary tone (H%), and focused wh-questions, which bear the same low IP boundary tone (L%) borne on focused declaratives.

8.4 Jun (2005)

Drawing from Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole (1999), Jun (2005) classifies (what can be assumed to be a Central dialect of) Bengali as an accentual phrase (AP) language, thus positing that the P-phrase of Hayes & Lahiri (1991), Lahiri & Fitzpatrick-Cole (1999), and Truckenbrodt (2003) (and, by extension, the MaP of Selkirk 2006 described in §8.5, the ip of Michaels & Nelson 2004, and presumably the stress group of Ray *et al.* 1966) corresponds to an AP, the approximately word-sized prosodic phrase defined as the domain of exactly one pitch accent. Her summary of prosodic phrasing is consistent with previous studies, positing only two tonally-marked levels (*i.e.* IP and AP). Despite previous suggestions by Hayes & Lahiri (1991) and Michaels & Nelson (2004) that the P-phrase/ip can host any number of pitch accents (including zero), Jun's classification of Bengali as an AP language entails that each of these phrases can

host only one pitch accent; while ips are not defined by the number of pitch accents contained, APs are always the domain of exactly one pitch accent. Jun also classifies Bengali as a non-tonal (although she does mention Lahiri & Fitzpatrick-Cole's lexically-specified tone for emphatic clitics), non-lexical pitch accent, non-stress accent language, assigning pitch on heads and edges only postlexically, thus adopting the classification of Bengali described in Ladd (1996), reproduced below in Table 15, while countering the stress-accent classification suggested in other aforementioned studies of Bengali.

		Phonetic typology	
		Stress accent	*Non-stress accent*
Lexical typology	*Lexical pitch*	example: Swedish	example: Japanese
	Postlexical pitch only	example: English	example: Bengali

Table 15. Ladd's (1996) typology of accent type.

As Jun (2005) is concerned with typological observations, there is no new data provided to support her classification of Bengali in the intonational typology. As discussed more thoroughly in Chapter Three, the current study concurs with Jun's classification of Bengali as a non-tonal, non-lexical pitch accent language, assigning pitch on prosodic heads and edges only postlexically. The current study also reclassifies Bengali as a stress-accent language based on phonological alternations and pitch accent attraction (§6), confirming many previous analyses of the language. This would place Bengali in the stress accent/post-lexical pitch cell of Ladd's typology (Table 15), along with English. Furthermore, the current study adopts Jun's Accentual Phrase (AP) as the

smallest tonally-marked prosodic unit in Bengali, while also positing an Intermediate Phrase (ip).

8.5 Selkirk (2006)

Using the sentence reproduced in (31) below, Selkirk (2006) reanalyzes Hayes & Lahiri's (1991) data by applying features from her Prosodic Hierarchy to the phrasing of focused constituents, and by using both morphophonological and morphosyntactic OT constraints to derive the tonal patterns described in previous studies of the intonation of Standard Kolkata Bengali.

(31) ami raǰar čʰobir ǰonno ʈaka anlam
[ami ɹadʑaɹ tɕʰobiɹ dʑonːo taka anlam]
I king's pictures for money gave [sic]
'I gave [sic] money for the king's pictures.'[41]

8.5.1 Prosodic hierarchy

Selkirk's prosodic hierarchy, shown in Figure 4 below, is composed of an Intonation Phrase (IP), which is composed of at least one Major Phrase (MaP), which is composed of at least one Minor Phrase (MiP), which is composed of at least one Prosodic Word (PWd), which can be further decomposed into feet, syllables, and lastly, moras.

[41] The sentence should be translated 'I *brought* money for the king's pictures', not *gave* ([anlam] means 'I brought', while [d̪ilam] is 'I gave').

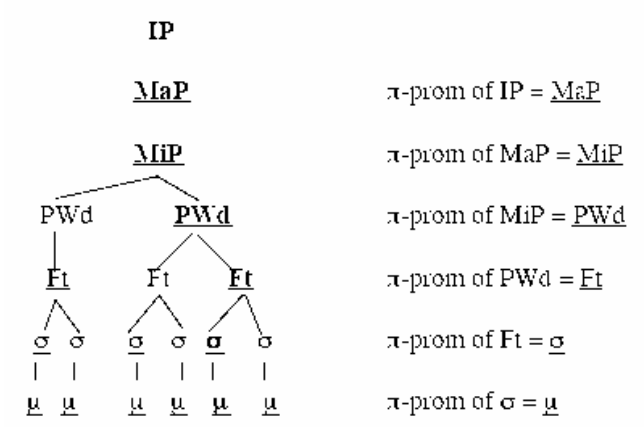


Figure 4. Selkirk's prosodic hierarchy

In her reanalysis of Hayes & Lahiri (1991), and by extension, of Ray *et al.* (1966) and Lahiri & Fitzpatrick-Cole (1999), Selkirk groups words in Bengali into phrases, which group into one large sentence-level phrase. While she labels the sentence-level phrase the Intonation Phrase (IP), like previous analyses, she analyzes the P-phrase of Hayes & Lahiri (1991) (and, by extension, the p-phrase of Lahiri & Fitzpatrick-Cole 1999 and the Stress Group of Ray *et al.* 1966) as a Major Phrase (MaP), defined as the morphosyntactic maximal projection phrase, located in the hierarchy immediately below the IP. The MaP is associated with what she calls "big" focus (or FOCUS) – contrastive focus, in other words. Like Hayes & Lahiri, Selkirk makes a connection between the P-phrase and what others would call an ip; in her analysis of Japanese, Selkirk defines Beckman & Pierrehumbert's (1986) and Pierrehumbert & Beckman's (1988) ip as a MaP, and also defines their AP as a MiP. A comparison of three prosodic hierarchies is given in Shattuck-Hufnagel & Turk's (1996) prosody tutorial, reproduced below in Figure 5.

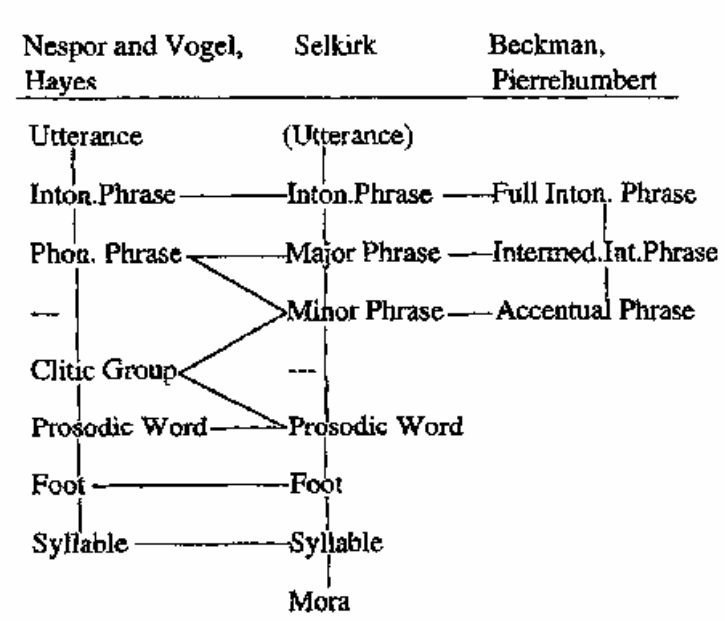


Figure 5. A comparison of three prosodic hierarchies (Shattuck-Hufnagel & Turk 1996, Fig. 2, p. 14).

Following Shattuck-Hufnagel & Turk's analysis, Selkirk's MaP indeed maps to the ip of the intonation-driven prosodic models of Beckman and Pierrehumbert, and to the P-phrase of other syntax-driven prosodic models, such as those of Nespor & Vogel (1986) and Hayes (1989). As represented by the two association lines attached to the P-phrase in Figure 5, however, the P-phrase of the syntax-driven group's hierarchy can map to either of two different phrase types in the other hierarchies, depending on the language and analysis. While Selkirk makes the claim that Hayes & Lahiri's P-phrase refers to the MaP/ip in Bengali, Jun (2005) claims that the P-phrase is equivalent to the MiP/AP in Korean.

8.5.2 Theoretical objective

Selkirk's main objective in reanalyzing previous studies of Bengali intonation lies in the theory that particular kinds of phonetic phenomena should not be directly related to a morphosyntactic phenomenon such as focus. She states that the appearance of particular tonal morphemes, the appearance of particular pitch accents, the addition of unexpected prosodic boundaries, the appearance of nuclear stress, non-phonemic vowel lengthening, and other prosodic phenomena should not be derived by focus-prosody interface constraints in any language. They should instead fall out naturally from constraints independently governing the prosody of the language, and that the only phonetic feature that should be directly related to focus is metrical prominence (*i.e.* some particular level of stress).[42]

The main problem Selkirk finds with previous models of Bengali intonation is the appearance of a P-phrase boundary to the right of the focused constituent, as in (13), where the focused constituent *kágoǰəlake* 'newspaperman' projects a high P-phrase boundary tone (H_P) to the right. Since Bengali normally is left-edge-prominent (*i.e.* metrical prominence is left-headed, in that stress regularly occurs on the leftmost syllable of the leftmost non-clitic word in a phrase), and since prosodic breaks are normally aligned with prominent syllables (Truckenbrodt 1995), there is no motivation for a prosodic break in the aforementioned focus sentence to be projected to the right of a focused word.

[42] See Calhoun (2007) for a perception study of the relation between prominence and focus.

Having stated the theoretical problem, Selkirk proceeds to reanalyze the high P-phrase boundary tone found to the right of a focused element as a floating $[H]_{FOC}$ tone[43], and also succeeds in deriving by OT constraint interaction many of the tones posited by previous authors as phonemic or morphemic. I briefly discuss these analyses below.

8.5.3 OT analysis of phrasing and underspecified tones

By drawing more heavily on the OCP and by employing violable constraints in an OT framework, Selkirk derives many of the tones that must be treated as underlying in previous analyses. Selkirk notes, as did Hayes & Lahiri and Lahiri & Fitzpatrick-Cole, that certain combinations of tones do not occur on the surface in Bengali. Adjacent identical tones (*e.g.* two high tones in a row) are shown to not exist in Kolkata Standard Bengali, explaining why high plateaus in pitch are unattested, and why focused (underlying $L^*\ldots H_P\ldots H_I L_I$) and default yes/no interrogatives ($L^*\ldots H_I L_I$) are largely homophonous. As previous researchers have proposed that the OCP keeps such strings from occurring, Selkirk extends this principle by proposing that many tones are underspecified for height (*i.e.* H or L), and the OCP derives the height based on surrounding tones. I summarize here the rationale behind her proposal, while overlooking the details, as they will be largely irrelevant for the purpose of my proposal: the fact that the basic tonal pattern of a MaP in Bengali is L*…H – with polar opposite tones on either side – is not a coincidence; while the boundary tones are underlyingly high (H), the pitch

[43] It is unclear how positing a focus-specific tonal morpheme abides by Selkirk's interpretation of Focus Prominence Theory, as it does not fall naturally from independent constraints on the language's prosody.

accents are underlyingly unspecified for tonal quality. The OCP, however, ensures that the tone of the pitch accent is the opposite of the surrounding boundary (H) tones, thus deriving the default low pitch accent (L*). The boundary tones, in turn, are derived by a constraint that aligns floating H tones to the right edges of MaPs: Align-R (MaP, H). This predicts that all MaPs will project a H tone on their right edge; this is surface true, except for final MaPs – in Selkirk's sentence (31), the verb *anlam*, for example, does not have a H tone on its right edge. Selkirk explains this by proposing that the constraint against contour tones (*i.e.* *ContourTone) forbids the coexistence of the MaP-R H tone and the IP-R L boundary tone (to be discussed in §8.5.4 below). Since she does not explicitly analyze other sentences, it is not clear exactly how *ContourTone would prevent Hayes & Lahiri's I-phrase rising boundary tone L_IH_I from being simplified to L_I.[44]

8.5.4 Focus and tonal morphemes

The main objective of Selkirk's analysis is, as I mentioned earlier, resolving the issue of focus phrasing. Like her reanalysis of the non-focus-related high P-phrase boundary tone (H_P) as a floating H tone that docks to the right edges of MaPs via OT constraint interaction, Selkirk reanalyzes the focus-related high P-phrase boundary tone (H_P) as a tonal morpheme $[H]_{FOC}$, which has in its lexical representation both the morphosyntactic function of assigning focus and the phonological representation of a high tone.[45] As the morphosyntactic function of assigning focus is part of its underlying representation, the

[44] She does, however, posit Realize $[HL]_{QUES}$ to preserve falling contours in yes/no questions, and thus it is presumable that another constraint Realize $[LH]_{CONT}$ could preserve rising contours in continuation rises.

[45] Selkirk also posits a declarative tonal morpheme $[L]_{DECL}$, which I will not be discussing here.

tonal morpheme $[H]_{FOC}$ aligns to the focused constituent of the utterance. It is described as a sort of suffix in its morphosyntactic subcategorization, and thus automatically seeks out the right edge of the focused word and docks there. Then a constraint, namely Align-R ($[H]_{FOC}$, MaP), projects the right edge of a MaP to this tonal morpheme. In positing this representation and these constraints, Selkirk is able to derive both the prosodic boundary and the H tone found at the right edge of a focused phrase without making a direct connection between the morphosyntactic focus feature and the phonological constraints that derive prosodic phrasing.

In order to avoid cluttering the focus-prosody interface with non-independently-motivated processes such as the projection of a high boundary tone to the right of a focused constituent, Selkirk reanalyzes the structure of her example sentence (*i.e.* [ami ɹadʑaɹ tɕʰobiɹ dʑonːo taka anlam] 'I gave [sic] money for the king's pictures'), a focus declarative sentence with a large left-branching direct object (*i.e.* [ɹadʑaɹ tɕʰobiɹ dʑonːo taka] 'money for the king's pictures'). Specifically, she proposes the existence of tonal morphemes, and describes the interaction of eleven OT constraints governing the alignment of those morphemes with various morphosyntactic boundaries. In doing so, Selkirk is able to explain the assignment of tones across MaPs, justify the existence of a boundary at the right edge of the focused constituent, and derive the dephrasing of post-focal words.

8.6 Summary

In adopting an autosegmental-metrical representation of intonational phonology, the five intonational phonology models of at least two forms of Bengali (*i.e.* Hayes & Lahiri 1991; Lahiri & Fitzpatrick-Cole 1999; Michaels & Nelson 2004; Jun 2005; Selkirk 2006) revolutionized the way in which Bengali pitch contours were analyzed. By distinguishing pitch accents and boundary tones, as well as by distinguishing boundary tones for different levels of prosodic structure, the massive tonal inventories of pre-intonational phonology models were reduced to combinations of L and H tonal targets aligned to stressed syllables and the right edges of two prosodic units above the word level. I summarize in Table 9 the prosodic units described in intonational phonological models of Bengali.

Hayes & Lahiri (1991)	**Lahiri & Fitzpatrick-Cole (1999)**	**Michaels & Nelson (2004)**	**Jun (2005)**[46]	**Selkirk (2006)**[47]
I-phrase	i-phrase	IP	IP	IP
P-phrase	p-phrase	ip	AP	MaP

Table 16. Summary of prosodic units and disjuncture types described in various Intonational Phonology studies of Bengali prosody.

[46] Although she does not explicitly propose an intermediate phrase (ip) in Bengali, Jun's prosodic hierarchy allows for such a prosodic unit, as has been done in her analysis of Korean.

[47] Although she does not explicitly propose a minor phrase (MiP) in Bengali, Selkirk's prosodic hierarchy allows for such a prosodic unit, as has been done in her analysis of English.

In Chapter Three, I introduce the current study, providing a detailed description of the data collected in a series of experiments, along with a discussion of findings to support the claims of the current model.

CHAPTER THREE
INTONATIONAL PHONOLOGY OF BENGALI

In this chapter I describe in detail the intonational phonological model and accompanying transcription system (*i.e.* Bengali Tones and Break Indices or B-ToBI) of the prosody of the standard form of the Bengali language as spoken by the subjects of the experiments described in §9 below. This form of the language will be identified in the current study as Bangladeshi Standard Bengali. The model is presented in §10; the tonal inventory, the interactions between tones, and comparisons of similar tones are described in §10.1, and a discussion of phrasing and break indices is provided in §10.2.

9 METHODS

The tonal inventory, prosodic structure, and phonological rules proposed in the current intonational phonological model are based on data collected in a series of experiments conducted in 2006. I include below a general characterization of the dialect background of the subjects, as well as a brief description of each experiment.

9.1 Subjects

A total of 29 subjects participated in the various experiments carried out for the current study, although not all subjects participated in all tasks. Most speakers were born and grew up in parts of Bengal that are now in modern-day Bangladesh. Of the 20 speakers who participated in Experiment I[48] (where the majority of the data supporting

[48] Sex of the subjects participating in Experiment I: 9 male, 11 female. Ages ranged from late teens to 60s.

the current model is found), nine identified as coming from Northern Bengali-speaking districts, ten from Eastern Bengali-speaking districts, and one from a Central Bengali-speaking district. As Northern and Central Bengali are classified together in Grierson (1928) as part of the Western Branch of Bengali dialects, this splits the subject pool in half—ten speakers from districts where Western Branch dialects are spoken and ten speakers from districts where Eastern Branch dialects are spoken. In this sense, the form of Standard Bengali spoken by the subjects of the current study should presumably balance out some of the contrasting influences from the Western and Eastern Branches. However, as the majority of the districts of subject origin (18/20) are found in what is now Bangladesh (even though these Bangladeshi districts are split across the Western-Eastern Branch boundary), I cannot claim that the current model captures the diversity of Indian varieties of Standard Bengali. Thus, I characterize the standard language spoken by the subjects of the current study as Bangladeshi Standard Bengali, noting additional speaker-specific details where appropriate.

9.2 Experiment I – Production of scripted sentences

Experiment I records a total of 57 scripted sentences, consisting mostly of sonorants to aid in pitch tracking, produced by 20 subjects each. Sixteen of these 57 sentences included (near-) homophonous sentences with different expected intonational patterns elicited by the existence and placement of sentence-level particles, word-level focus particles, and/or punctuation marks. In these 16 recordings, the sentence [monoaɹa ɹomilake nie elo] 'Monoara brought Romila,' was repeated with different intonations,

signaled by a question mark ([monoaɹa ɹomilake nie elo]? ‘Did Monoara bring Romila?’), or particles such as -[na] ([monoaɹa-na ɹomilake nie elo]? ‘Didn’t Monoara bring Romila?’), -[naki] ([monoaɹa-naki ɹomilake nie elo]? ‘So is it true that Monoara brought Romila?’), -[o]~[-ɔ̰] ([monoaɹa *ɹomilake-o* nie elo] ‘Monoara brought *Romila*, too.’), and others. Certain particles triggered focus realization, while others signaled a change in IP boundary tone (*e.g.* for yes/no questions, echo questions, *etc.*). These 16 sentences were recorded mainly to collect the inventory of boundary tones in the language.

Thirteen sentences were recorded as examples of corrective focus. The subject would first produce an “incorrect” sentence such as [monoaɹa ɹumuke nie elo, t̪aj na]? ‘Monoara brought Rumu, right?’, which would be followed immediately by a correction (also produced by the subject), [na, na. monoaɹa *ɹomilake* nie elo] ‘No, no. Monoara brought *Romila*.’ The focused answer was controlled for length, to test whether the number of segments, syllables, or words affected the realization of corrective focus.

Another ten sentences were sentence pairs involving wh-questions and their corresponding answers. The answers were all [ɹumu nepaleɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni] ‘Rumu couldn’t remember the names of the gardeners of the queen of Nepal’, but the preceding wh-question would trigger focus realization on only one part of the answer. For example, one of the question-answer pairs goes as follows: [ɹumu nepaleɹ ɹaniɹ kon koɹmiḏeɹ namgulo mone ɹakʰt̪e paɹe ni]? ‘Which workers of the queen of Nepal did Rumu not remember the names of?’, which would trigger focus on [maliḏeɹ]

‘of the gardeners’ in [ɹumu ɹaniɹ *maliḏeɹ* namgulo mone ɹakʰt̪e paɹe ni] ‘Rumu couldn’t remember the names of the *gardeners* of the queen of Nepal.’ These ten sentences were collected in order to answer questions of focus domain size and realization type.

The remaining sentences elicited phenomena including lists, embedded clauses, strings of intermediate phrases, focus realization within and outside an embedded clause, and the four-way distinction in break size (word, AP, ip, IP). These were elicited in order to measure duration at large prosodic breaks, as well as investigate the interaction of pitch accents and different boundary tone types. See the Appendix for a full set of sentences recorded. Measurements and comparisons made between subsets of the data set are described in more detail in later sections.

Each subject first sat with the experimenter and read the randomized list of scripted sentences to himself or herself. Then, after the subject was familiarized with the sentences, he or she read them individually into a USB microphone attached to a laptop. The subject was allowed to re-record a sentence if he or she felt that it was produced disfluently. The recordings were made in WaveSurfer (Sjölander & Beskow 2005) and were analyzed in Praat (Boersma & Weenink 2005).

9.3 Experiment II – Naturalistic production of a story

As a study of naturalistic speech, Experiment II has three main goals: (1) confirmation of the validity of the proposed tones and phrasal structure in connected

speech, (2) identification of the nonstandard dialect influences on individual speakers' production of Bangladeshi Standard Bengali, and (3) illustration of the prosodic patterns of those nonstandard varieties. In Experiment II, a total of 29 subjects were given the short children's book "Frog, where are you?", composed of detailed illustrations but no written text. Each subject was told to browse through the book and familiarize him- or herself with the story. The experimenter then left the room, having asked the subject to tell the story in his or her own words (in whatever dialect he or she felt most comfortable speaking) into a USB microphone attached to a laptop, using the pictures as a guide. By offering the subjects no written material as a guide, it was possible to record something very close to spontaneous speech, avoiding the complications associated with scripted speech (*e.g.* reading pauses at line breaks or before difficult words, disfluencies within words, misinterpretation of intended meaning or focal domain, *etc.*). The recordings, which were made in Audacity, normally lasted between five and ten minutes. Selected excerpts of these recordings were later annotated in Praat (Boersma & Weenink 2005) using the B-ToBI model of transcription described in §10.

9.4 Experiment III – Cross-dialectal comparison of focus enclitics

Experiment III is a brief investigation of the interaction between the two focus enclitics -[i]~[j] 'only', 'indeed' and -[o]~[ɔ̰] 'also', 'even' and various boundary tones of higher prosodic units (*i.e.* ip, IP, to be introduced in §10.1.2), produced in both Bangladeshi Standard Bengali and nonstandard Eastern Bengali. The sentence [monoaɹa

ɹomilake meɹe felːo] ‘Monoara killed Romila’ was produced once without clitics, once with the clitic -[i]~[j] ‘only’, ‘indeed’ attached to the first half of the complex verb (*i.e.* [meɹe] ‘having beaten’), once with the clitic attached to the second half (*i.e.* [felːo] ‘dropped’), once with the clitic -[o]~[o̞] ‘also’, ‘even’ attached to the first half, and once on the second half, and once with narrow focus elicited without clitics on the entire complex verb (*i.e.* [meɹe felːo] ‘killed’). The same variations were applied to the sentence [ɹomilake meɹe felːo monoaɹa], a scrambled form of the original sentence. By moving the subject to the end of the sentence, the emphatic clitics could be separated from the sentence boundary. This list of variations in clitic existence/placement and word order was repeated in a second list in Eastern Bengali, using the base sentence [monoaɹa ɹomilaɹe majɹa falajlo] ‘Monoara killed Romila’. One bi-dialectal subject was recorded for this task. The subject sat with the experimenter and read the two lists (*i.e.* Standard, Eastern) of scripted sentences to herself. Then, after being familiarized with the sentences, she read them into a USB microphone attached to a desktop computer. The subject was allowed to re-record sentences she felt were produced disfluently. The recordings were made in Audacity and were analyzed in Praat (Boersma & Weenink 2005).

10 MODELING THE INTONATIONAL PHONOLOGY OF BENGALI

The intonational model I propose for Bangladeshi Standard Bengali adopts the theoretical framework of the Autosegmental-Metrical theory of Intonational Phonology, as described in Chapter Two. The system used to transcribe the prosody is called Bengali Tones and Break Indices or B-ToBI, due to its representational similarity to ToBI

transcription systems in use for other languages (*e.g.* MAE_ToBI for Mainstream American English, Beckman & Ayers Elam 1997; G-ToBI for German; J-ToBI for Japanese; see Jun 2005). The prosodic model and B-ToBI transcription system include a full inventory of pitch accents and boundary tones, along with descriptions of the phonetic realization of these entities, and the phonological interaction between them. In addition to describing the pitch accents, boundary tones, and their variants, I test hypotheses to compare the current model to previous prosodic models of different varieties of Bengali, using acoustic measurements as support for the current model's characterization of Bangladeshi Standard Bengali.

10.1 Tones

As with the Intonational Phonology models of other languages, the current model of Bangladeshi Standard Bengali recognizes multiple layers of prosodic structure marked by tone. The smallest tonally marked phrase, or accentual phrase (AP), is the basic unit of Bengali intonation, and roughly corresponds to a single word or small group of words. The largest tonally-marked prosodic unit is the intonation phrase (IP), which often spans an entire sentence. Between these two units is the intermediate phrase (ip), which demarcates certain phrases and clauses. In this section, I describe in detail the tones associated with each prosodic unit. In §10.1.1, I introduce the phonetic and distributional characteristics of three tonal patterns in the basic accentual phrase (AP), and move on to the phonetic and distributional characteristics of higher prosodic domains (*i.e.* ip, IP) in §10.1.2. The interactions between these tones are illustrated in §10.1.3. The

characteristics of each of the boundary tones (*i.e.* Ta, T-, T%) corresponding to the three tonally-marked prosodic units (*i.e.* AP, ip, IP) are compared directly in §10.1.4. The findings of this intonational phonological model with respect to the postlexical tonal inventory are summarized in §10.1.5.

10.1.1 The accentual phrase (AP)

The basic unit of Bengali intonation is the accentual phrase (AP), which is defined as the prosodic phrase projected by a single pitch accent. The relationship of pitch accent to AP is taken to be one to one, in that every pitch accent is the head of an AP, and every AP must host exactly one pitch accent. The only points in the AP to which a tone can be associated are the metrically most prominent syllable (*i.e.* the phonologically stressed syllable) and the AP's right edge, as schematized in Figure 6 below.

T* Ta T* Ta T* Ta T* Ta T* Ta T* Ta[49]

[pɹó̪tʰom]AP [ʃɔ́bd̪ota]AP [t̪aɹ pɔ́ɹeɹ]AP [ʃɔ́bd̪eɹ]AP [áge]AP [áʃe]AP

first the word its following of word before comes

'The first word comes before its following word.'

Figure 6. Schematic illustration of a sentence with six APs.

[49] The AP boundary tone (Ta) of the last word of the sentence would be overridden by a coinciding IP boundary tone (T%), as explained in §10.1.1, under the heading *Pitch accent-AP boundary tone relationship*.

The exact phrasing of a string of words into APs is not entirely predictable. The factors that influence the phrasing of APs include speech rate, word length, and focus, all of which will be discussed to some degree but not with respect to their influence on the variable phrasing of APs; see Truckenbrodt (2003) for an account of this optionality. The six APs schematized in Figure 6 above, for example, could also be produced as three APs, as in Figure 7 below. Note that the number of phrases must match the number of pitch accents; some of the syllables that bore pitch accents or AP boundary tones in Figure 6 lose their tones when the six APs are reduced to three.

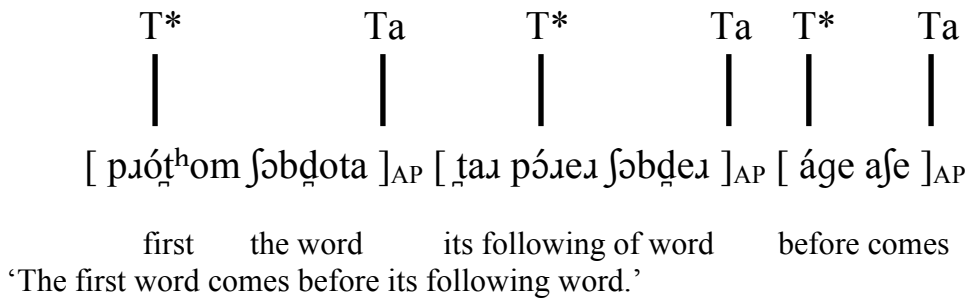


Figure 7. Schematic illustration of a sentence with three APs.

To illustrate this optionality of accentual phrasing, I have included pitch tracks of a sentence containing seven orthographic words, [mumbajeɹ ɹelgaɾi bomahamlaʂ ɹomilaɹ nananani(ɹa)[50] maɹa gelen] 'Romila's grandparents passed away in Mumbai's train bombing', produced by two different speakers. The first speaker splits this string into two ips, and each ip is broken into three APs, giving a total of six APs spanning the string, as shown in Figure 8 below. For now, the reader may ignore the exact labels of each tone,

[50] The animate plural accusative case marker –[ɹa] is optional in this position; the two speakers whose recordings are illustrated here differ in their inclusion of the suffix.

and simply observe where the tonal targets along the pitch track lie, demarcating the metrically prominent positions (*i.e.* stressed syllables) and prosodic boundaries (*i.e.* right edges of tonally-marked phrases) in the string.

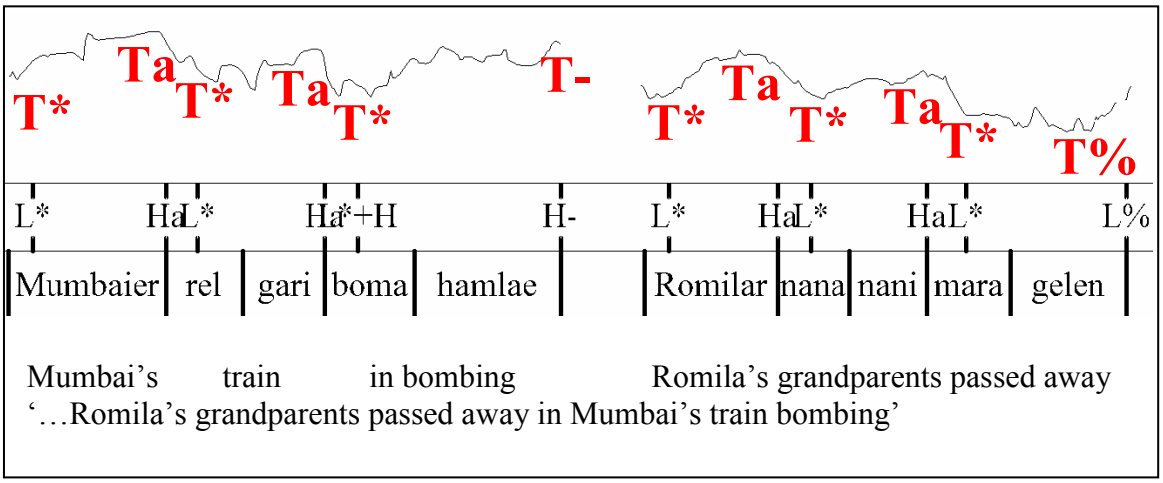


Figure 8. Here, this sentence is split into two ips, altogether composed of six APs, [mumbajeɹ] 'Mumbai's' [ɹelgaɾi] 'train' [bomahamlaɛ̃] 'in bombing', [ɹomilaɹ] 'Romila's' [nananani] 'grandparents', and [maɹa gelen] 'passed away'. [Tu49][51]

The second speaker produced the same string of words with fewer pitch accents and phrase breaks; the seven orthographic words are phrased within one large ip, which is split into only three APs, as shown in Figure 9 below.

[51] Examples from the current study are arranged with the pitch track aligned with two labeling tiers: the tone tier includes labels for all pitch accents and boundary tones, and the word tier, which divides up the segmental string by either orthographic word boundaries or content word boundaries. The word tier uses a shorthand transcription system described in the Appendix. Under the word tier is a rough English gloss of the sentence, followed by a more natural translation. Each example is also given a label in square brackets, with numbers and letters identifying the speaker, stimulus, and experiment.

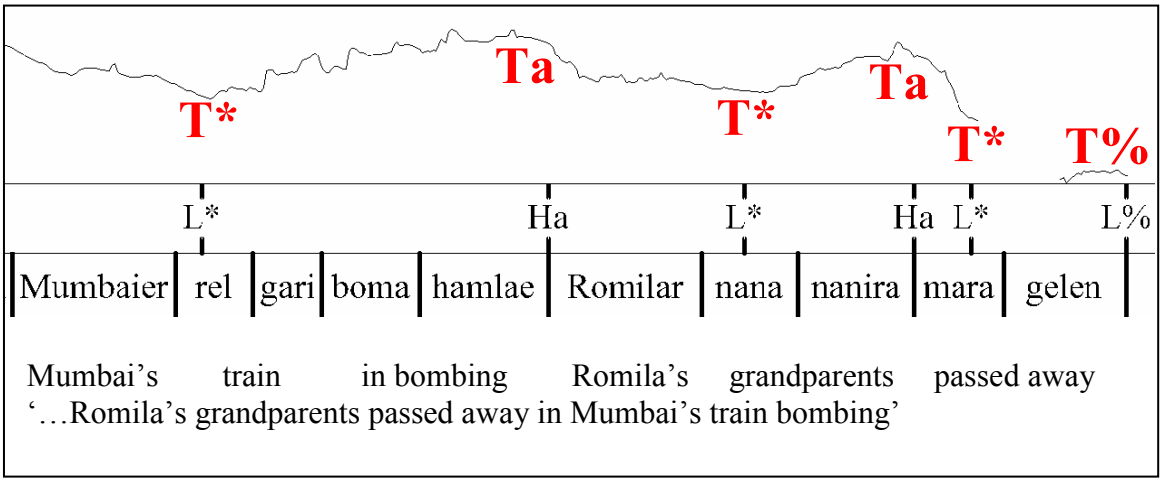


Figure 9. Here, the same string is split into only three APs, [mumbajeɹ ɹelgaɾi bomahamlaɛ̃] 'in Mumbai's train bombing', [ɹomilaɹ nanananiɹa] 'Romila's grandparents', and [maɹa gelen] 'passed away'. Note that the pitch falls across the first word [mumbajeɹ] 'Mumbai's' due to interpolation from an earlier part of the sentence (not shown here) towards the low pitch accent (L*) on [ɹel] 'rail'. [Bo49]

As is exemplified in the preceding schematized illustrations and actual pitch tracks, the location of pitch accents and boundary tones is to some degree variable. However, the variability only extends to phonologically primary stressed syllables and the right edges of phrases larger than a word (*i.e.* AP, ip, IP). Other locations, such as the left edge of the AP or syllables that bear either no phonological stress or only secondary phonological stress cannot be linked to tones. Of course, in practice tones are not solely borne on points in the sentence; the AP boundary tone, for example, cannot always be produced literally "at the right edge". Since tones must be produced during tone-bearing units (TBUs) to be perceived by the listener, the actual realization of a boundary tone can be stretched back to the last TBU in the AP—which we can assume to be the last vowel. In addition, some tones associated to the pitch accent can also be realized on syllables other

than the phonologically primary stressed syllable; this is primarily the case for the bitonal pitch accent (L*+H), described below. Furthermore, all intervening syllables, including unstressed syllables, bear pitch although they do not bear underlying tonal targets, as pitch is understood in the AM theory to be interpolated between tonal targets.

The basic AP

The basic AP in Bengali (*i.e.* one that is non-final within the larger prosodic domain) is made up of exactly two tonal targets: one high (H) target and one low (L) target. These two targets are arranged in only three patterns, two of which are composed of a pitch accent and boundary tone of opposing tonal targets: smooth rise (L*…Ha), smooth fall (H*…La). The third tonal pattern is the sharp rise (L*+H), which does not appear to bear an AP boundary tone; the argument for a lack of an AP boundary tone after the sharp rise is discussed in Chapter Four. In each of the three basic patterns, the first tonal target is borne on the stressed syllable, and the second tonal target is borne further along in the phrase, with the point of realization determined by a number of factors, some of which are described in greater detail in Chapter Four. As the three AP tonal patterns are composed of separate pitch accents (*i.e.* L*, H*, L*+H) and AP boundary tones (*i.e.* La, Ha), they can be broken down into their component parts. However, the three pitch accents and two AP boundary tones do not freely cooccur (there are no observed patterns of the shape L*…La, H*…Ha, L*+H…La, *etc.*), and thus identifying the pitch accent type in each pattern automatically also identifies the AP boundary tone type, if there is a boundary tone at all. Given this distributional characteristic, I begin by describing the

three tonal patterns seen in non-final APs (*i.e.* smooth rise, smooth fall, sharp rise), along with example pitch tracks. I later describe the distribution of pitch accents without boundary tones.

Smooth rise (L*…Ha)

The smooth rise (L*…Ha) can be considered the default AP tonal pattern, as it is by far the most common pattern in Bangladeshi Standard Bengali, and it is not apparently associated with any particular meaning or structure. It is composed of a low pitch accent (L*), which associates to a phonologically stressed syllable, and a high AP boundary tone (Ha), which associates to the right edge of the AP. The low pitch accent (L*) and high AP boundary tone (Ha) respectively highlight the most metrically prominent syllable and the edge of the phrase. Phonetically, this pattern is characterized by an F_0 valley during the phonologically stressed syllable, smoothly rising to an F_0 peak at the right edge of the phrase, as shown in Figure 10.

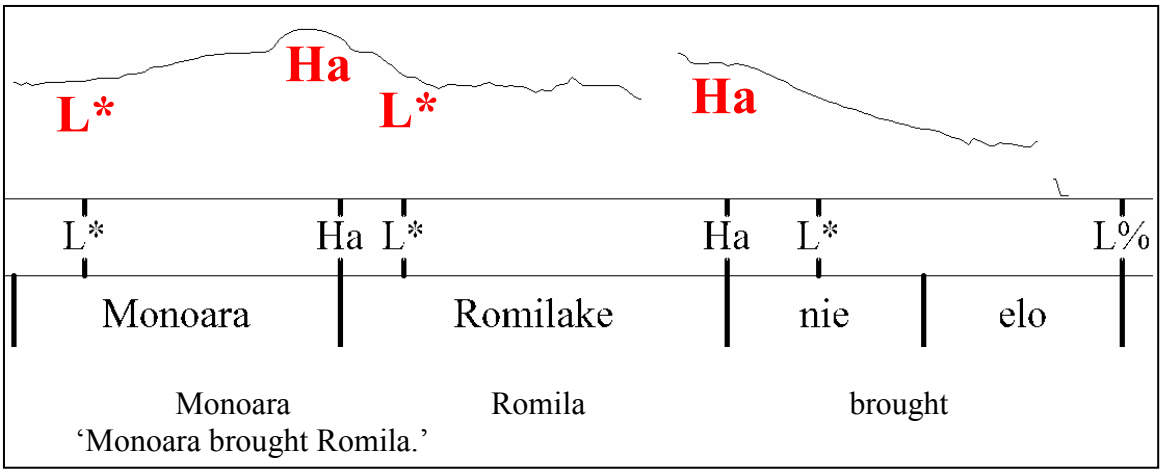


Figure 10. The subject [monoaɹa] 'Monoara' and the object [ɹomilake] 'Romila-ACC' both bear a smooth rise AP tonal pattern, composed of a low pitch accent (L*) and high AP boundary tone (Ha). [Tu01]

The smooth rise pattern of the current model is presumably the same as the head phrase (L*…H_P) of Hayes & Lahiri's (1991) original Intonational Phonological analysis, and similar phrases in subsequent analyses (*i.e.* Lahiri & Fitzpatrick-Cole's L*…H_P, Michaels & Nelson's L*…H-, Jun's L*…Ha[52], Selkirk's L*…H).

In sentences with repeated smooth rise patterns (L*…Ha), one very salient feature of the pitch of the high AP boundary tone (Ha) can be seen. Note in Figure 11 below how each successive high AP boundary tone (Ha) is lower in pitch than that of the preceding

[52] Jun (2005) does not formalize the labels for Bengali tones, but she does identify Bengali as an AP language, suggesting that the default tonal pattern described in previous studies should be formalized as a low pitch accent (L*) and high AP boundary tone (Ha).

AP. This gradual lowering of pitch across a series of smooth rise patterns (L*…Ha) is called Ha downtrend.[53]

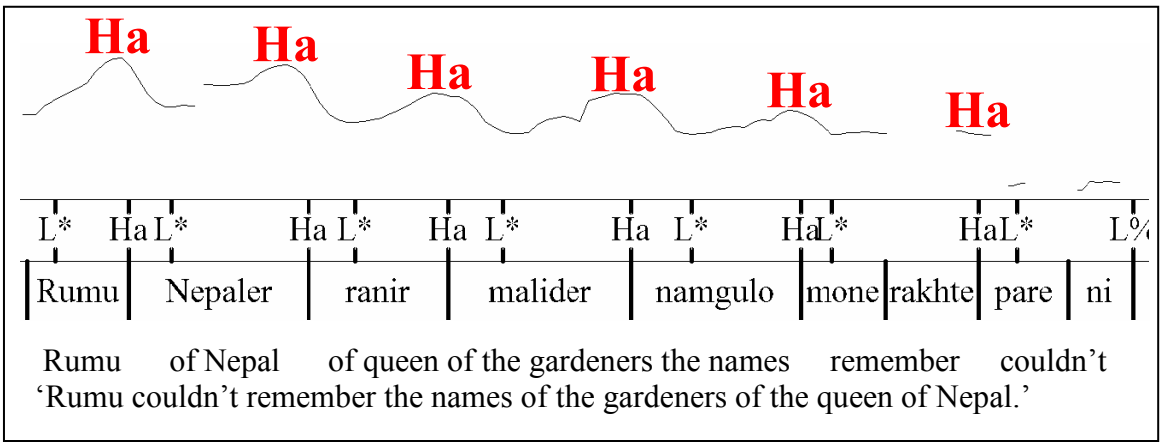


Figure 11. The pitch of the high AP boundary tone (Ha) of each smooth rise AP tonal pattern (L*…Ha) reaches a lower pitch than the preceding AP, following Ha downtrend. [Fa50]

The pitch of each of the six high AP boundary tones (Ha) in the smooth rise patterns (L*…Ha) illustrated in Figure 11 above, starting from the leftmost AP [ɹumu] 'Rumu (a name)', are 320Hz, 302Hz, 250Hz, 246Hz, 210Hz, and 166Hz. Although the slope is not uniform—the pitch of the high AP boundary tone (Ha) following [ɹaniɹ] 'queen's' is only 4Hz higher than the high AP boundary tone (Ha) following [maliḍeɹ] 'of the gardeners'—the general downtrend of successive high AP boundary tones (Ha) is largely consistent. Following the Ha downtrend line can help serve to judge if a string of default/unmarked smooth rise patterns (L*…Ha) is broken by some marked/non-default

[53] See Thorsen (1983) for a discussion of downtrend in Danish, Barjam (2004) for Porteño Spanish, Arvaniti & Baltazani (2005) for Greek, and Jun (2007) for Korean.

material, such as a sharp rise (L*+H), or the high boundary tone of some larger prosodic unit (see §10.1.2). One may also notice that the pitch of each successive low pitch accent (L*) also progressively lowers; this general L* downtrend is however far less regular, and it is not uncommon to see a string of smooth rise patterns (L*…Ha) featuring clear Ha downtrend but violating what could be considered L* downtrend, as in Figure 12 below.

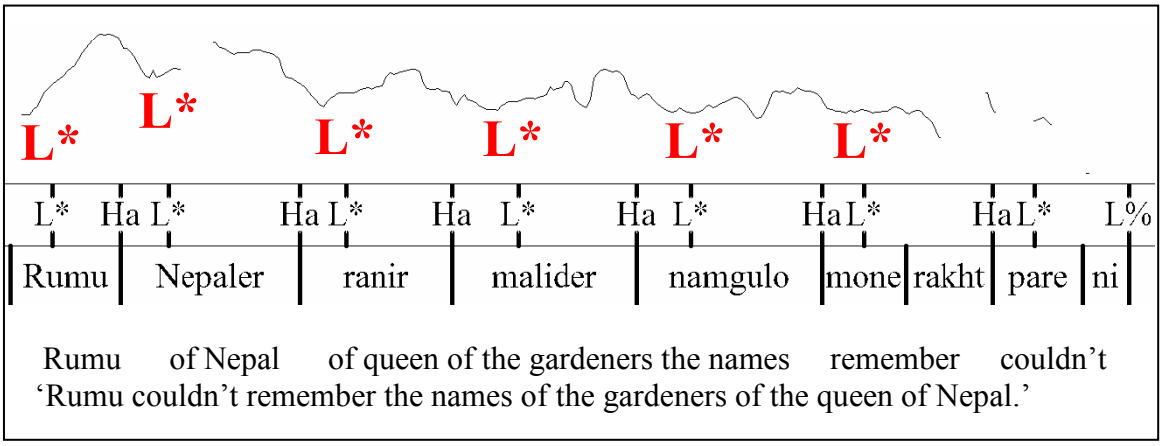


Figure 12. Although the pitch of the high AP boundary tone (Ha) of each smooth rise AP tonal pattern (L*…Ha) reaches a lower pitch than the preceding AP, following Ha downtrend, the low pitch accents (L*) do not regularly follow downtrend. [Ba50]

Ha downtrend can be affected by at least two controllable factors: word size and word type. If a short AP is followed by a much longer AP, the Ha of the longer AP may in fact surpass that of the previous short AP, as is seen in Figure 13 below.[54] Note that the very long AP [naɹaɡ̊ongɔndʑe] 'to Narayanganj' boasts a higher pitch (149Hz) on its high

[54] The fact that word length can affect the height of the high AP boundary tone (Ha) suggests that the pitch contour associated with smooth rises (L*…Ha) may underlyingly bear a default slope whose exact realization can be further affected by additional factors such as word type. However, as the current study is primarily concerned with the fundamental units of Bengali prosody, I do not attempt to model the exact arithmetic slope of any pitch contour.

AP boundary tone (Ha) than the previous AP [amaɹ] ‘my’ (136Hz). Thus, only successive APs of equivalent length can be counted on to consistently follow Ha downtrend.

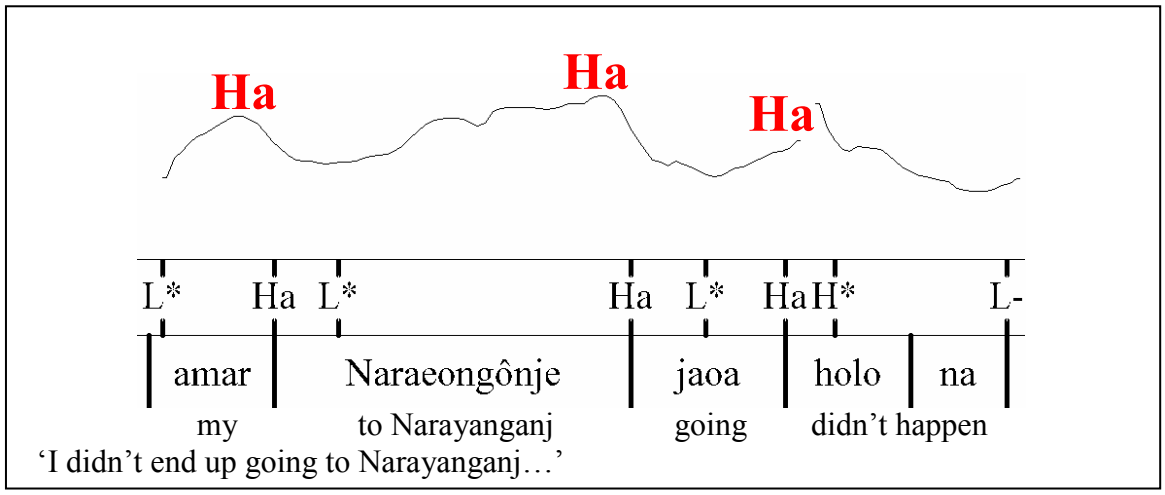


Figure 13. The pitch of the high AP boundary tone (Ha) of each smooth rise AP tonal pattern (L*…Ha) typically reaches a lower pitch than the preceding AP if it is of equivalent size; longer APs can reach higher pitch than preceding shorter APs. [SB37]

The other factor that can affect Ha downtrend is word type. While it is normally predicted that the first AP in a string of successive APs of equivalent length should reach the highest pitch for the high AP boundary tone (Ha), this is not the case when the string begins with an AP composed entirely of a function word. For example, although the first three APs in Figure 14 below are of roughly the same length (in terms of number of syllables), the high AP boundary tone (Ha) of the initial AP [kaɹon] ‘because’ (107Hz) does not reach a higher pitch than that of the following AP [miɹaɹ] ‘Mira’s’ (110Hz), which is a content word.

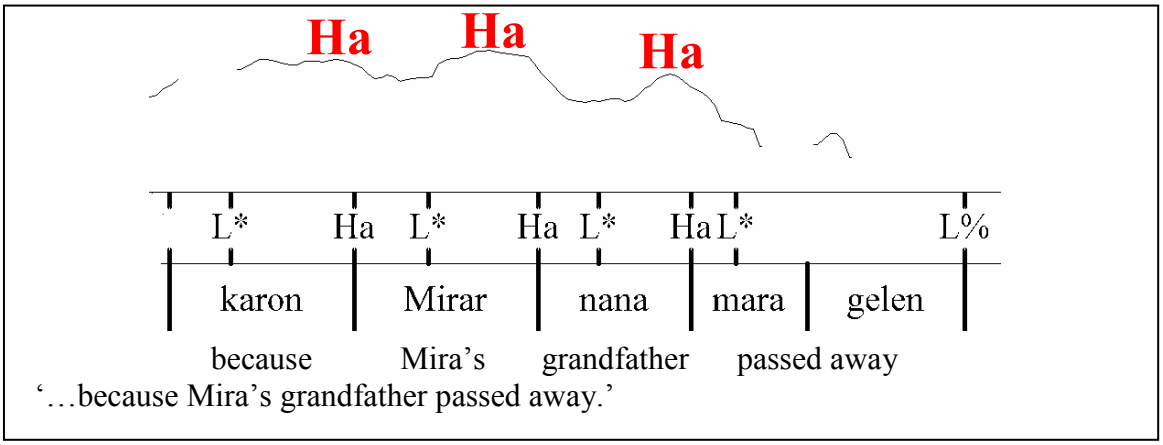


Figure 14. Although the pitch of the high AP boundary tone (Ha) of each smooth rise AP tonal pattern (L*…Ha) typically reaches a lower pitch than the preceding AP, this pattern does not hold for function words such as [kaɹon] 'because'. [SB37]

Smooth fall (H*…La)

Another tonal pattern seen in non-final APs is the smooth fall (H*…La), which is not described in any previous studies on Bengali prosody. The smooth fall pattern (H*…La) is composed of a high pitch accent (H*) and low AP boundary tone (La), exactly the inverse of the smooth rise pattern (L*…Ha). Consequently, the pattern is also phonetically realized as the inverse of the smooth rise: an F_0 peak during a phonologically stressed syllable, followed by a smooth fall in pitch towards an F_0 minimum at the right boundary of the AP. Unlike the smooth rise pattern (L*…Ha), however, the smooth fall pattern (H*…La) appears to be somewhat restricted in its distribution in two regards: first, it is very rare in the recordings of scripted speech, where it is presumed that speakers are using Bangladeshi Standard Bengali. Secondly, the smooth fall pattern (H*…La) only occurs before a high pitch accent (H*)—thus, either

before another smooth fall pattern (H*…La) or before an ip-final high pitch accent (H*) (*i.e.* one without an AP boundary tone realized, due to constraints on tonal interaction described further below). One of the few examples of the smooth fall pattern (H*…La) produced in Bangladeshi Standard Bengali (collected as one of the scripted sentences of Experiment I) is illustrated in Figure 15 below, on the AP [ʃej namgulo] 'those names', preceding the high pitch accent (H*) on the AP [bʱule gelen] 'forgot-HON'.

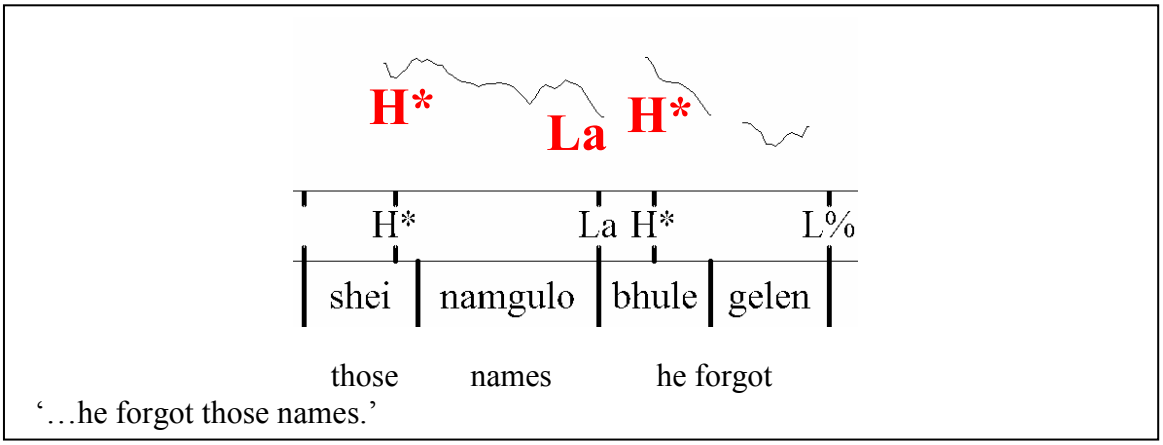


Figure 15. The AP [ʃej namgulo] 'those names' bears a smooth falling tonal pattern, composed of a high pitch accent (H*) and low AP boundary tone (La). [Ba51]

The smooth fall pattern (H*…La) is more common in nonstandard dialects, as recorded in Experiment II, a study of naturalistic speech. In Figure 16, the final AP [kʰaɔ̯aɹ mɔt̪o] 'like eating'[55] bears a high pitch accent (H*), and its preceding AP [digbadʑi] 'somersault' bears the smooth fall pattern (H*…La).

[55] The compound verb [digbadʑi kʰa]- 'do a somersault' uses a very common non-literal use of the verb [kʰa-] 'eat'. Literally, the phrase translates to 'eat a somersault'; this use of 'eat' is found in many

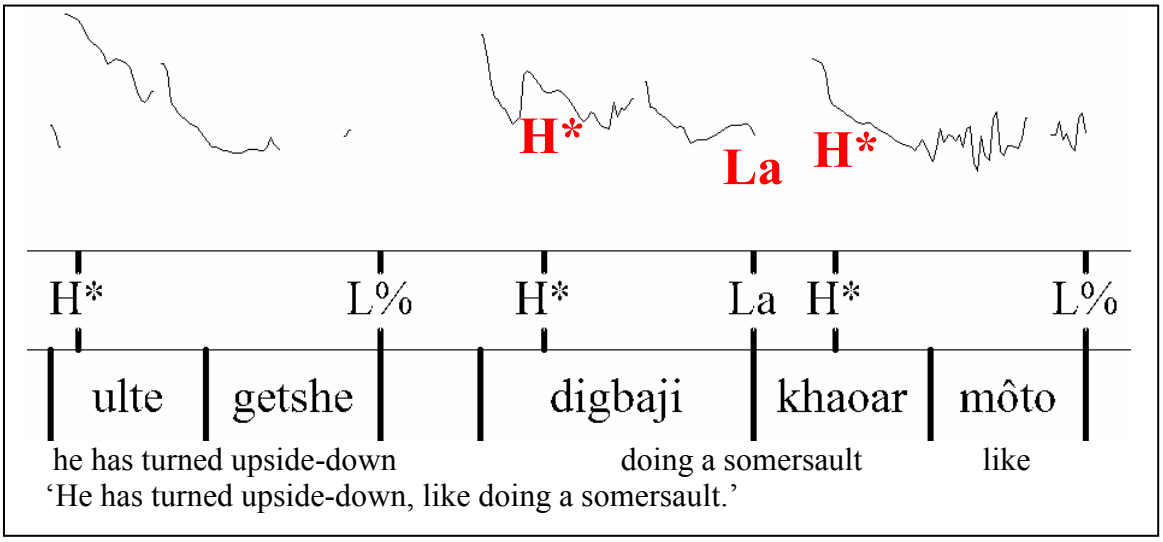


Figure 16. The non-final AP [digbadʑi] 'somersault' bears a smooth fall tonal pattern, composed of a high pitch accent (H*) and low AP boundary tone (La). The phrase was produced by a speaker from Rajshahi Division (in the North Bengali dialect area), using her native dialect in a recording session of naturalistic speech. [BMS170]

Similarly, in Figure 17 below, the high pitch accent (H*) on [fele d̪ise] 'he has dropped' allows the immediately preceding AP [tɕʰeletake] 'the boy-ACC' to bear the smooth fall pattern (H*…La).

expressions denoting passive voice or involuntarily actions (*e.g.* [tʰapːoɹ kʰa]- 'get slapped', [atɕʰaɾ kʰa]- 'fall over', [bɔka kʰa]- 'get scolded').

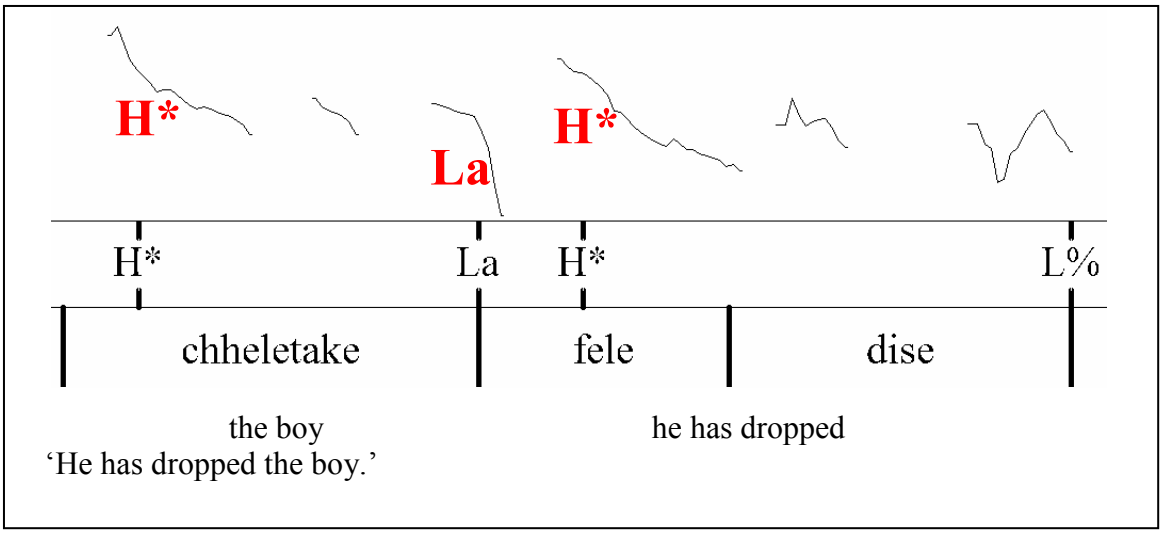


Figure 17. The non-final AP [tɕʰeletake] 'the boy-ACC' bears a smooth fall tonal pattern, composed of a high pitch accent (H*) and low AP boundary tone (La). The phrase was produced by a speaker from Mymensingh District (in Dhaka Division), using her native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. [NnS119]

In Figure 18 below, the final AP [poɹe ɡɛlo] 'fell' bears a high pitch accent (H*), giving its preceding APs [ɛkta pukuɹe] 'in a pond' and [mone hɔe̯ tɕʰeleta] 'it seems the boy' the option of bearing the smooth fall pattern (H*…La). Note that the high pitch accents (H*) and low AP boundary tones (La) become progressively lower in a series of smooth fall patterns, just as low pitch accents (L*) and high AP boundary tones (Ha) become progressively lower in a series of smooth rise patterns. However, due to the low frequency of long strings of smooth fall patterns (H*…La), it is unclear as to how common a putative H* downtrend or La downtrend might be.

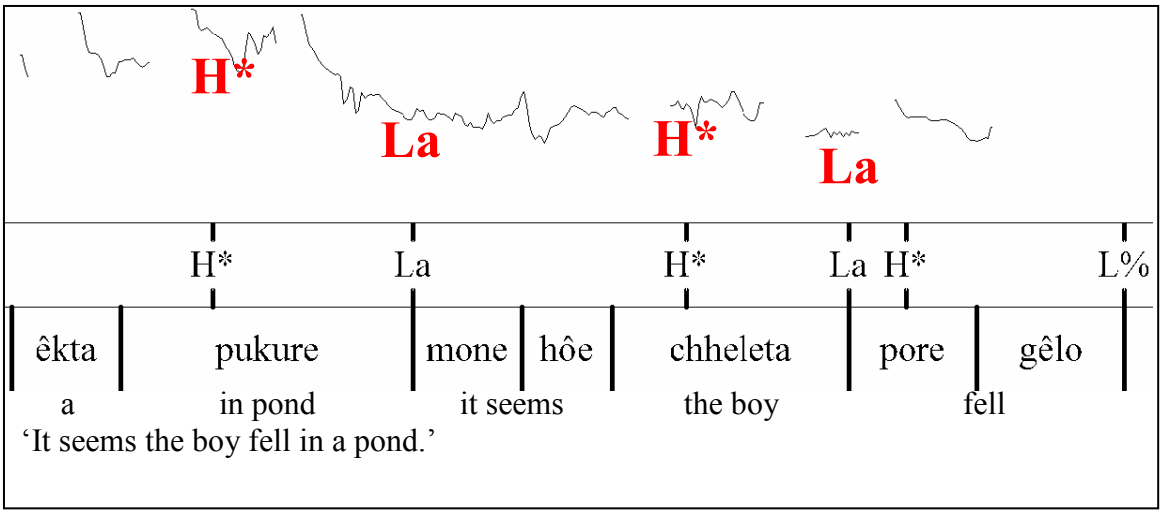


Figure 18. The non-final APs [ɛkta pukuɹe] 'in a pond' and [mone hɔe̯ tɕʰeleta] 'it seems the boy' both bear a smooth fall tonal pattern, composed of a high pitch accent (H*) and low AP boundary tone (La). The phrase was produced by a speaker from Tangail District (in Dhaka Division), using her native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. [ByS184]

The fact that previous descriptions of Bengali prosody were unable to find examples of the smooth fall pattern (H*…La) is most likely due to the dialectal familiarity of the subjects studied. As mentioned in Chapters One and Two, Hayes & Lahiri (1991), Lahiri & Fitzpatrick-Cole (1999), Jun (2005), and Selkirk (2006) base their data primarily from the speech of one speaker from Kolkata and checked against that of three other speakers of the same dialect, while the data for the current study are collected from speakers from various parts of both West Bengal (in India) and East Bengal (modern-day Bangladesh). Indeed, the smooth fall pattern is far more frequent in the data collected in Experiment II, in which subjects were asked to speak casually and fluently as they told a children's story without written prompts. In such a task, speakers presumably spoke in what might closely resemble their native dialect, which for all of the speakers includes influences from both

the standard dialect and a variety of nonstandard dialects. Recordings from Experiment II are full of instances of nonstandard vocabulary (*e.g.* use of /sɛmɹa/ and /pola/ for standard /tɕʰele/ 'boy'), pronunciation (*e.g.* use of denti-alveolar fricatives /s, sʰ, z, zʱ/ or affricates /ts, tsʰ, dz, dzʱ/ in place of standard postalveolar affricates /tɕ, tɕʰ, dʑ, dʑʱ/), and morphology (*e.g.* use of -/t̪es/ instead of standard -/tɕʰ/ for the progressive marker), and it is thus conceivable that the prosodic patterns recorded also reflect those of nonstandard dialects. Therefore, the smooth fall pattern (H*…La) can be listed as part of the tonal inventory of nonstandard eastern Bengali dialects, and only tentatively as part of the inventory of Bangladeshi Standard Bengali.

Sharp rise pattern (L*+H)

While the smooth rise (L*…Ha) and smooth fall (H*…La) AP tonal patterns are composed of two largely independent tones, the sharp rise AP tonal pattern (L*+H) is composed of a single pitch accent with two tonal targets, both anchored at the phonologically stressed syllable. Thus, it can be considered a bitonal pitch accent. Phonetically, it is realized as an F_0 minimum during the phonologically stressed syllable, sharply rising towards a high F_0 plateau spanning the following one or two syllables. Formally, the L* represents the low pitch associated with the phonologically-stressed syllable, while the +H represents the "trailing" high pitch of the following syllable(s).

Unlike the smooth rise (L*…Ha) and smooth fall (H*…La) patterns, the sharp rise pattern (L*+H) is somewhat restricted in the types of APs upon which it may occur; its

presence is typically indicative of certain types of focus (described in much greater detail in Chapter Four), as shown in Figure 19 below.

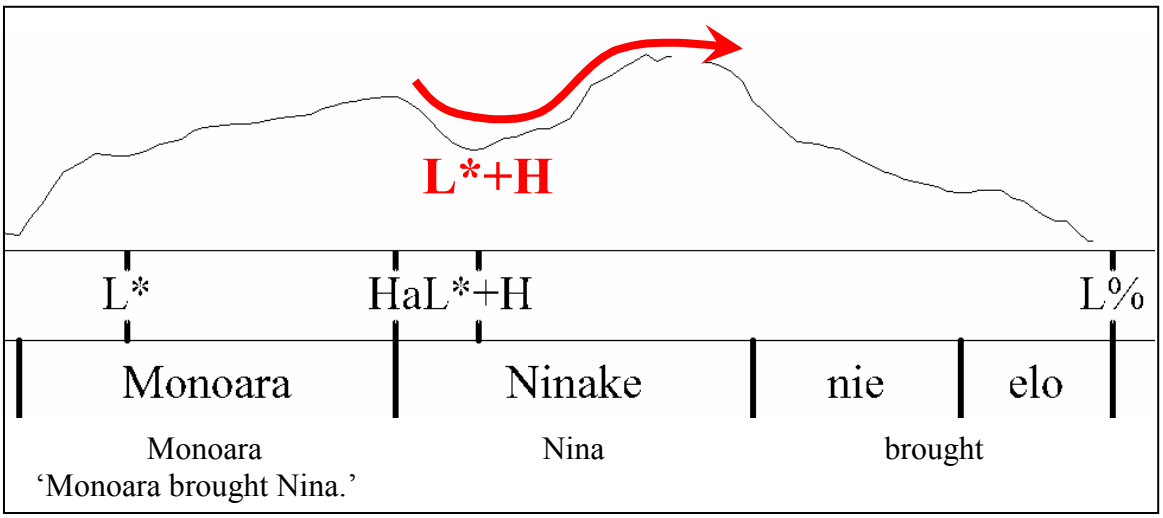


Figure 19. Under corrective focus[56], [ninake] 'Nina-ACC' bears a sharp rising pattern, composed of a rising pitch accent (L*+H). [Ro22]

Since both the smooth rise pattern (L*…Ha) and the sharp rise pattern (L*+H) involve low pitch on the phonologically stressed syllable and a high pitch to its right, it is important to justify the exact differences between the two AP tonal patterns. Compare the focused word [ɹomilake] 'Romila-ACC' in Figure 20 below to the same word, without focus, in Figure 21 further below.

[56] See Chapter Four for a discussion of different types of focus in Bengali.

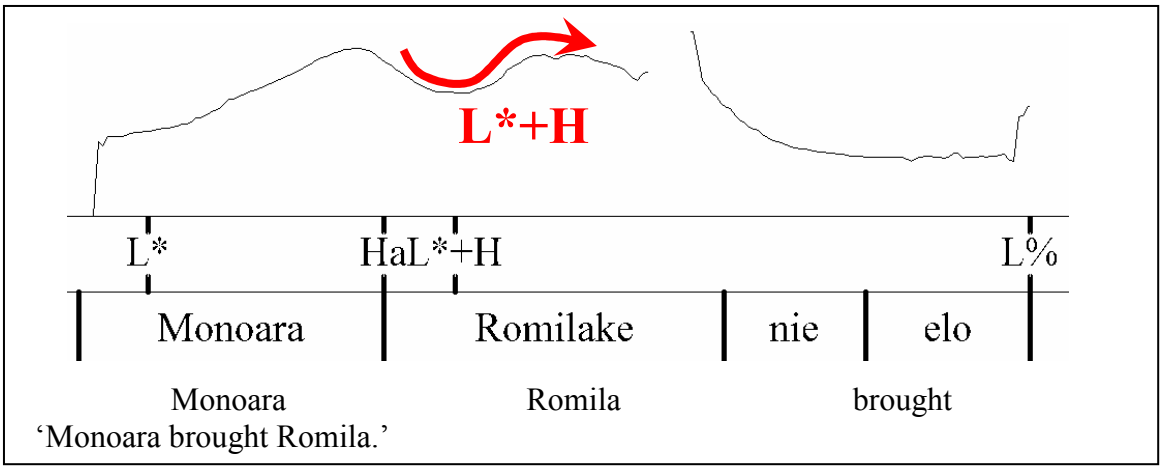


Figure 20. Under corrective focus, [ɹomilake] 'Romila-ACC' bears a sharp rising pattern, composed of a rising pitch accent (L*+H), with pitch extremes at 239Hz (during [o]) and 282Hz (during [il]). [Na23]

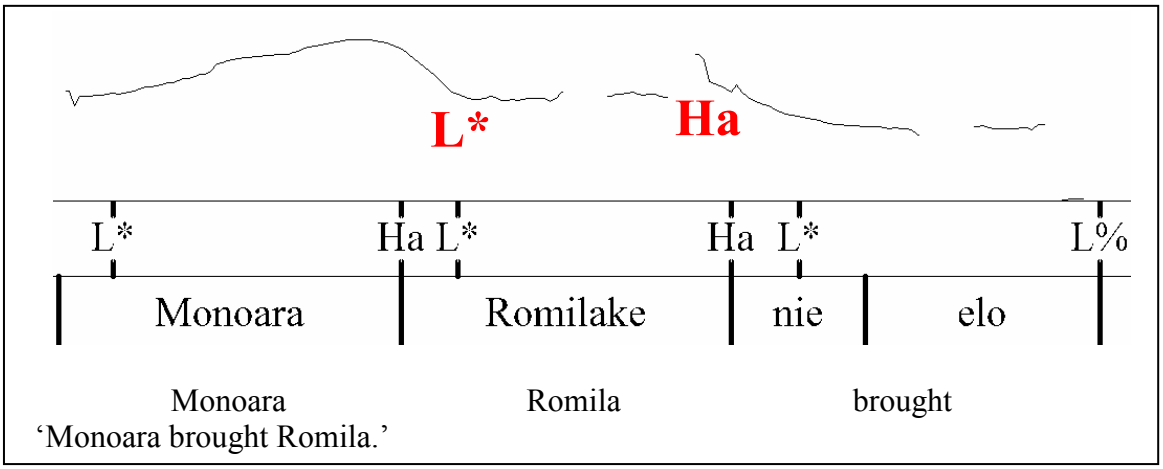


Figure 21. Without corrective focus, [ɹomilake] 'Romila-ACC' bears a smooth rising pattern, composed of a low pitch accent (L*) and high AP boundary tone (Ha), beginning at 219Hz (during [o]) and ending at 228Hz (during [e]). The spike in pitch just before the end of the AP is microprosody due to the voiceless stop [k]; the pitch during the release of the [k] is not taken to be part of the smooth rise. [Na01]

Note the sharper rise in pitch and the early pitch peak occurring partway through the word [ɹomilake] 'Romila-ACC' under corrective focus in Figure 20, as opposed to the smooth rise in pitch and later pitch peak occurring at the right edge of the same word

without corrective focus in Figure 21. Also note that the pitch maximum in the focused word reaches the same level (as in Figure 20) or a higher level (as in Figure 19) as the pitch maximum of the preceding AP, while the non-focused word bears a high AP boundary tone (Ha), which shows Ha downtrend (as in Figure 21). Detailed comparisons between focused and non-focused words (and thus, between smooth rises and sharp rises) are presented in Chapter Four.

Another position in which the sharp rise pattern (L*+H) is commonly found is in the final AP of a very small intermediate phrase (ip)—*i.e.* a phrase instead of a clause—even when the AP on which it appears is not being focused. Despite the lack of corrective or wh-answer focus, the existence of the rising pitch accent (L*+H) is suggested instead by the early pitch peak and violation of Ha downtrend, as shown in Figure 22 below.

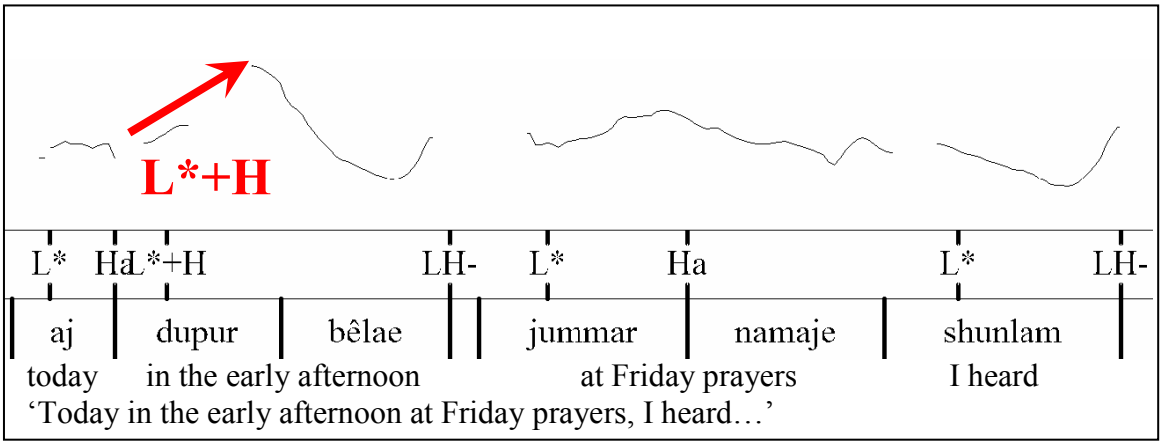


Figure 22. The ip-final AP [d̪upuɹ bɛlaɛ̃] 'in the early afternoon' bears a rising pitch accent (L*+H) despite the fact that it is not under focus. [Sh49]

As mentioned previously, another distinguishing feature of the sharp rise pattern (L*+H) is its ability to break the Ha downtrend. While each successive high AP boundary tone (Ha) in the smooth rise pattern (L*…Ha) lowers in pitch following Ha downtrend, the high trailing portion (+H) of the rising pitch accent (L*+H) is the only AP-level H target whose F_0 is higher than the previous AP-level H target, as shown in Figure 23 below, where the focused word [maliḍeɹ] 'of the gardeners' bears a rising pitch accent (L*+H) with a pitch maximum (114Hz) far higher than that of the preceding high AP boundary tone (Ha) (86Hz).

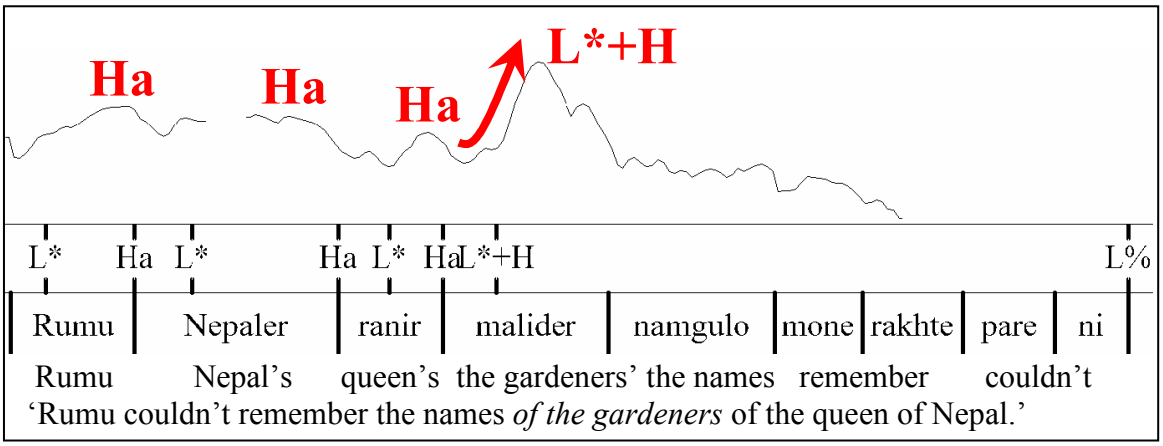


Figure 23. The wh-answer focused word [maliḍeɹ] 'of the gardeners' bears a rising pitch accent (L*+H), which reaches a higher pitch maximum than the preceding APs. [Sf44]

Even when the slope of Ha downtrend is shallow, the sharp rise (L*+H) reaches a much higher pitch than would be expected in a smooth rise (L*…Ha). For example, in Figure 24 below, the Ha downtrend slope is basically flat, with Ha levels at 222Hz, 224Hz, and 222Hz. Nevertheless, the pitch maximum of the sharp rise (L*+H) reaches 246Hz, thereby allowing it to stand out even after a series of rather high Ha values.

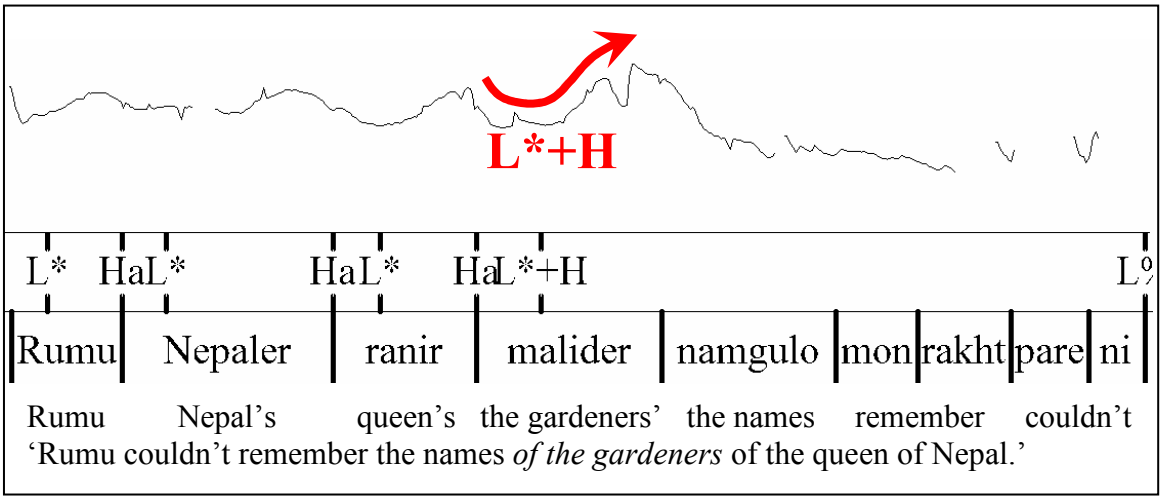
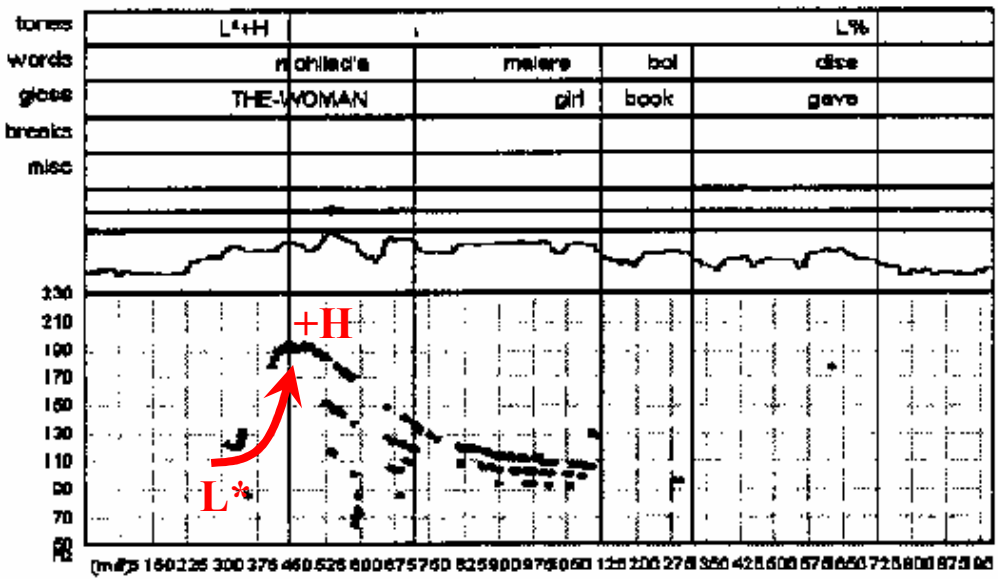


Figure 24. The wh-answer focused element [maliḍeɹ] 'of the gardeners' bears a sharp rise/rising pitch accent (L*+H). [Sh49]

Like the smooth fall pattern (H*…La), the sharp rise pattern (L*+H) is not described in most previous studies of Bengali prosody that identify pitch accent inventories (*i.e.* Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999, Jun 2005, Selkirk 2006). **In fact,** Hayes & Lahiri (1991) explicitly claims that there are no contour pitch accents in Bengali; the only complex tones are due to downstep (L+H*) in their model. The Michaels & Nelson (2004) study of Eastern Bengali, however, proposes a rising pitch accent (L*+H) for corrective focused elements; the pitch contours with which the rising pitch accent (L*+H) is associated are indeed very similar to the pitch contours of sharp rise patterns (L*+H) in the current study. Compare the rising pitch accents (L*+H) of the current study in Figure 23 and Figure 24 above to that of Michaels & Nelson in **(27**) below, repeated from Chapter Two.



(32) *mohilad'a* maiaɹe boi dise

[ˈ*mohilada* majaɹe boj d̪ise]

the woman girl book gave

'*The woman* gave the girl the book.' [M&N 22, §2.6.1, p. 22]

As previous intonational phonological models of Kolkata Bengali (a Central dialect) do not describe sharp rises or include rising pitch accents (L*+H) in their inventories, while the Michaels & Nelson (2004) study of Eastern Bengali does, it is possible that the tonal inventories of both nonstandard and standard varieties of Bengali spoken in eastern regions include a rising pitch accent (L*+H) while the tonal inventory of the standard variety of Kolkata Bengali does not. Another possibility is that Kolkata Bengali in fact has a rising pitch accent (L*+H), but it was not found in the Hayes & Lahiri (1991) and Selkirk (2004) studies, which did not differentiate between different kinds of focus (*e.g.* corrective, wh-answer, encliticized, *etc.*). Lahiri & Fitzpatrick-Cole (1999) differentiated

encliticized *vs.* non-encliticized focus types, but also did not find rising pitch accents (L*+H). This issue is addressed in Chapter Four.

In some instances, the sharp rise pattern (L*+H) can be realized as a shallow mid rise (^L*+H), characterized by mid or high pitch during the phonologically stressed syllable, followed by one or two syllables of sustained high or rising pitch, as shown in Figure 25 and Figure 26 below. The raised pitch of the L target is marked with a carat (^). Note that this is distinct from the upstep diacritic (¡) introduced in §10.1.3.

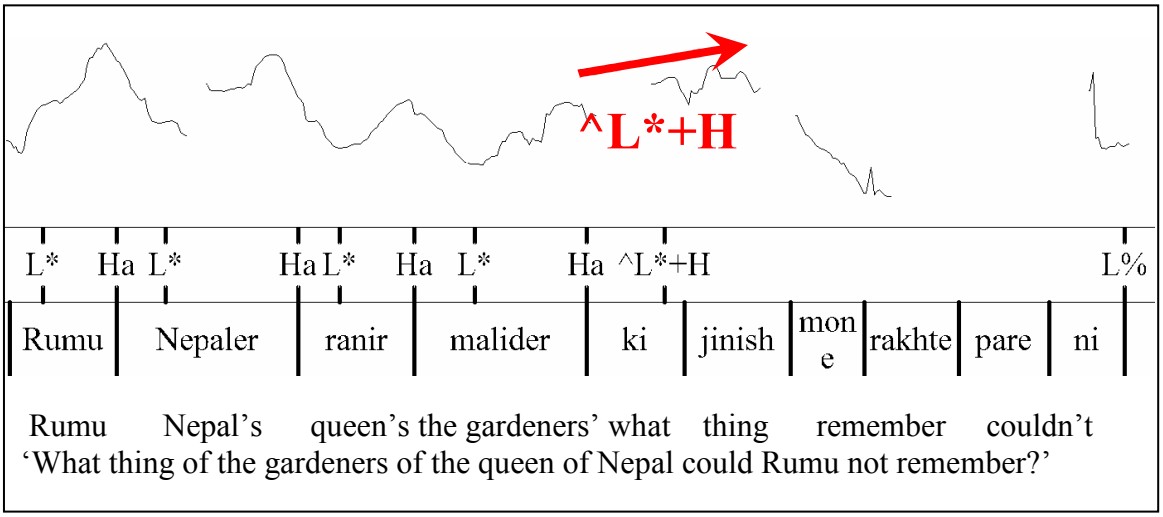


Figure 25. The wh-phrase [ki dʑiniʃ] 'what thing' bears the shallow mid rise variant (^L*+H) of the rising pitch accent (L*+H). (Note the pitch doubling on the IP-final syllable due to creaky voice.) [Fa47]

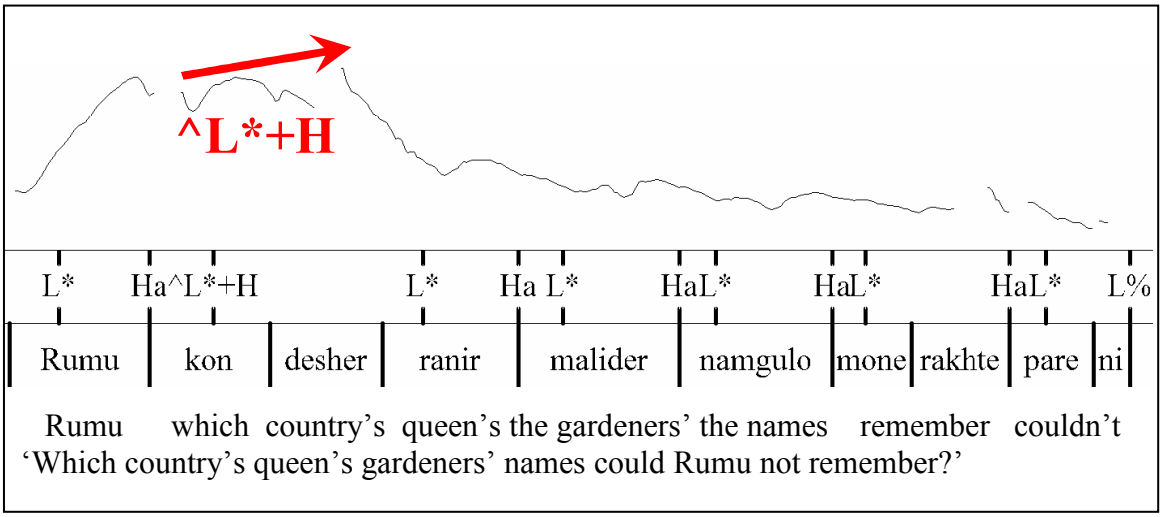


Figure 26. The wh-phrase [kon d̪eʃeɹ] 'which country's' bears the shallow mid rise variant (^L*+H) of the rising pitch accent (L*+H). It is unclear if the shallow mid rise (^L*+H) bears a high AP boundary tone (Ha) here. See Chapter Four for a discussion of the compression of the pitch range following the focused word. [Sh38]

The shallow mid rise variant (^L*+H) of the sharp rise pattern (L*+H) only occurs on wh-words words, which in Bengali always begin with voiceless consonants (*e.g.* [ki] 'what', [kon] 'which'). Considering that the phonetic qualities of voiceless consonants are known to raise pitch to various degrees in several languages (see Jun 1996), it is not unlikely that the shallow mid rise (^L*+H) is a phonetic variant of an underlying rising pitch accent (L*+H) following voiceless consonants. However, as the initial consonant of focused words was not controlled, it cannot be determined at this time if the shallow mid rise (^L*+H) occurs on other focused words (*i.e.* not just wh-words) beginning with voiceless consonants.

Pitch accent-AP boundary tone relationship

As mentioned earlier, two of the three basic AP tonal patterns—smooth rise (L*…Ha) and smooth fall (H*…La)—are composed of two separate tones: pitch accents (*i.e.* low L*, high H*) and AP boundary tones (*i.e.* low La, high Ha). Of these, it is the distribution of the AP boundary tones that is totally predictable. The high AP boundary tone (Ha) can only occur at the right edge of an AP whose head (*i.e.* most metrically prominent syllable) is marked by a low pitch accent (L*), and the low AP boundary tone (La) can only occur at the right edge of an AP whose head is marked by a high pitch accent (H*). There are no exceptions; the existence of a high (Ha) or low (La) AP boundary tone is always predictable. However, this predictability does not work both ways; the distribution of the pitch accents themselves is not predictable, due to a process called *concurrent tonal overriding*, also described in Michaels & Nelson (2004). When an AP is found at the right edge of a larger prosodic domain (*i.e.* the intermediate phrase, or ip), its expected boundary tone (Ta) is overridden—deleted and replaced by—the boundary tone of the higher prosodic domain (T-).[57] Due to this overriding of ip-final AP boundary tones (Ta), one cannot predict whether an ip-final pitch accent will be high (H*), low (L*), or rising (L*+H) by simply observing the surrounding tones; an ip- or IP-final tone (T-, T%) can easily cooccur with any of the three pitch accents (*i.e.* L*, H*, L*+H). As such, it is

[57] An alternative analysis can posit stacking of tones where the edges of tonally-marked prosodic units coincide. Such an analysis would presumably posit that ip boundary tones are composed of both AP- and ip-level tones, and that IP boundary tones are composed of AP-, ip-, and IP-level tones. This could help explain why AP-level tones are always monotonal (Ta), ip-level tones can be monotonal or bitonal (TaTi), and IP-level tones can be monotonal, bitonal, or tritonal (TaT-T%). Among the reasons why I do not adopt this alternative analysis is the fact that it would not be expected given the effect of the OCP on AP-level tones (*e.g.* L*…L% would have to be interpreted in the alternative analysis as L*…LaL-L%, even though a sequence of L*…La is normally prohibited).

important to observe the distribution and realization of the low (L*) and high (H*) pitch accents ip-finally.

In a model that views concurrent tonal overriding by positing that underlying boundary tones are overridden as higher boundary tones are added, we can assume that ip-final low pitch accents (L*) are paired with high AP boundary tones (Ha) like other smooth rise patterns (L*…Ha), but that this underlying high AP boundary tone (Ha) is not pronounced on the surface when an ip boundary tone (T-) overrides it. In a model where boundary tones are simply not projected when they appear in a position where the boundary tone of a higher prosodic phrase will be borne, we can assume that ip-final APs have no AP boundary tones (Ta), both underlyingly and on the surface. This second view, of course, assumes that tones associated with higher prosodic units are projected before AP boundary tones; there is, however, no independent evidence to support this theory. The two views are schematized below in Figure 27 and Figure 28.

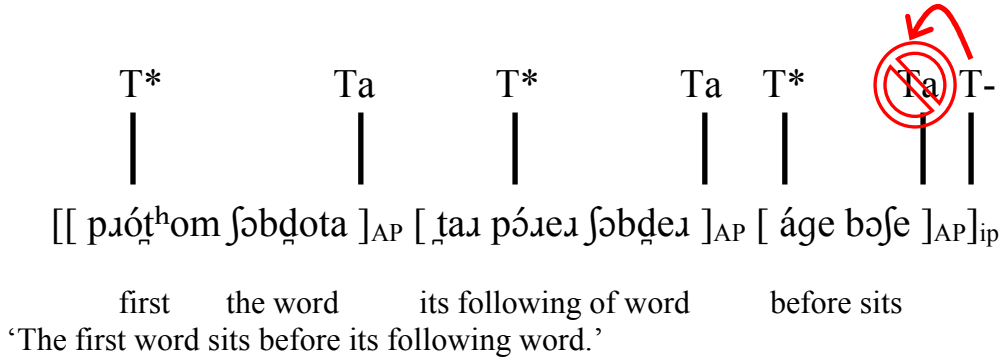


Figure 27. Schematic illustration of a sentence where the ip-final AP projects an AP boundary tone (Ta), which is overridden (illustrated by the arrow and crossed circle) by the concurrent ip boundary tone (T-).

T* Ta T* Ta T* T-

[[pɹóṯʰom ʃəbḏota]$_{AP}$ [ṯaɹ póɹeɹ ʃəbḏeɹ]$_{AP}$ [áge bəʃe]$_{AP}$]$_{ip}$

first the word its following of word before sits

'The first word sits before its following word.'

Figure 28. Schematic illustration of a sentence where the ip-final AP does not project an AP boundary tone (Ta) due to existence of the ip boundary tone (T-) in the same location.

Without making a strong case for either view, I now describe the realizations and distributions of the low (L*) and high (H*) pitch accents.

Low pitch accent (L*)

The low pitch accent (L*) is the most frequently-observed pitch accent in the corpus, both in the smooth rise AP tonal pattern (L*…Ha) and in ip-final position. It is not associated with any particular meaning, and can thus be considered the default pitch accent of the language. It is realized as relatively low pitch during a phonologically stressed syllable. However, as the default pitch accent, it can occur before all sorts of ip- and IP boundary tones, and therefore its realization cannot always be described as a local pitch minimum; a following low or rising boundary tone could easily reach a lower pitch. However, the low pitch accent (L*) is typically realized at a lower pitch than the preceding tone, which is most often a high AP boundary tone (Ha). Examples of ip-final low pitch accents (L*) are provided in Figure 29 and Figure 30 below.

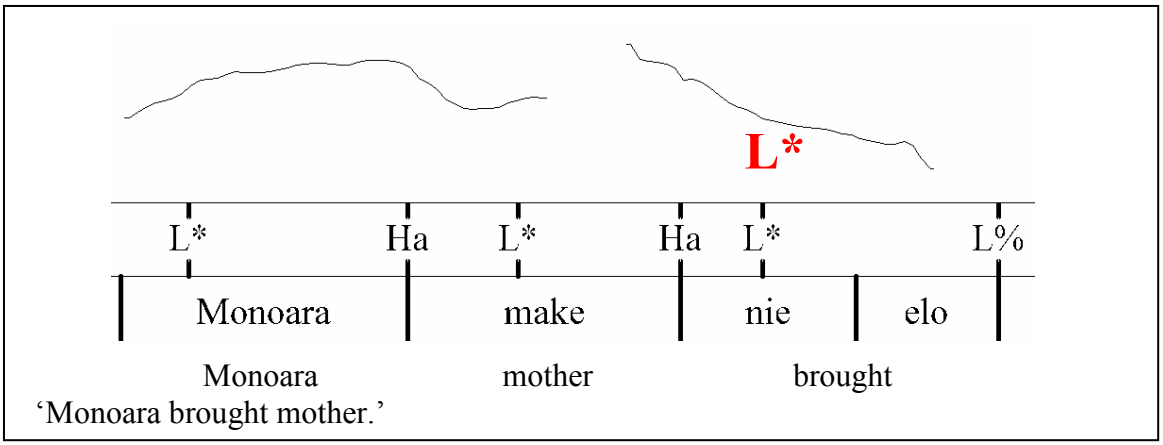


Figure 29. The ip-final AP [nie elo] 'brought' bears a low pitch accent (L*). [Ba19]

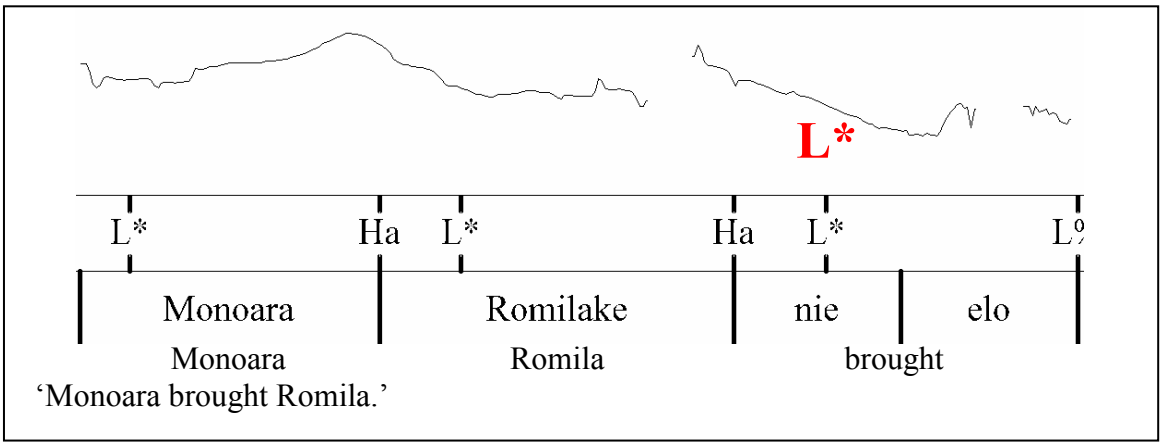


Figure 30. The ip-final AP [nie elo] 'brought' bears a low pitch accent (L*). The irregular pitch track during the word [elo] 'came' is due to creaky phonation. [Bo01]

The low pitch accent (L*) has been described in every intonational phonological model of Bengali, regardless of the variety studied. Thus, the low pitch accent (L*) of the current model can be considered equivalent to other descriptions of a low pitch accent (L*), except for a small difference in the distribution: while Hayes & Lahiri's (1991) low pitch accent (L*) is associated with the stressed syllables of prenuclear P-phrases and

interrogative nuclei only, the low pitch accent (L*) of the current study is the default pitch accent associated with the phonologically stressed syllables of *any* phrase, and is not restricted to position within the larger phrase, or to sentence type. In other words, Hayes & Lahiri (1991) and many subsequent studies maintain that the only instance in which a low pitch accent (L*) can occur I-phrase-finally is in interrogative sentences, while declarative sentences bear a high pitch accent (H*) I-phrase-finally. In contrast, the current analysis finds that almost all phrases bear a low pitch accent (L*), regardless of whether the sentence is declarative or interrogative. Compare for example the following two pitch tracks, composed of the same string of words, with Figure 31 representing the declarative sentence and Figure 32 representing the corresponding interrogative (yes/no) sentence. Notice how in both sentences, the pitch accent borne on the phrase [nie elo na] 'didn't bring' is lower in pitch than the preceding high AP boundary tone (Ha), as well as the preceding low pitch accent (L*).[58]

[58] The sequence of a low pitch accent (L*) followed by a low IP boundary tone (L%) may be similar to the sequence of a downstepped high pitch accent (L+H*) and low I-phrase boundary tone (L_I) described for Kolkata Standard Bengali in Hayes & Lahiri (1991).

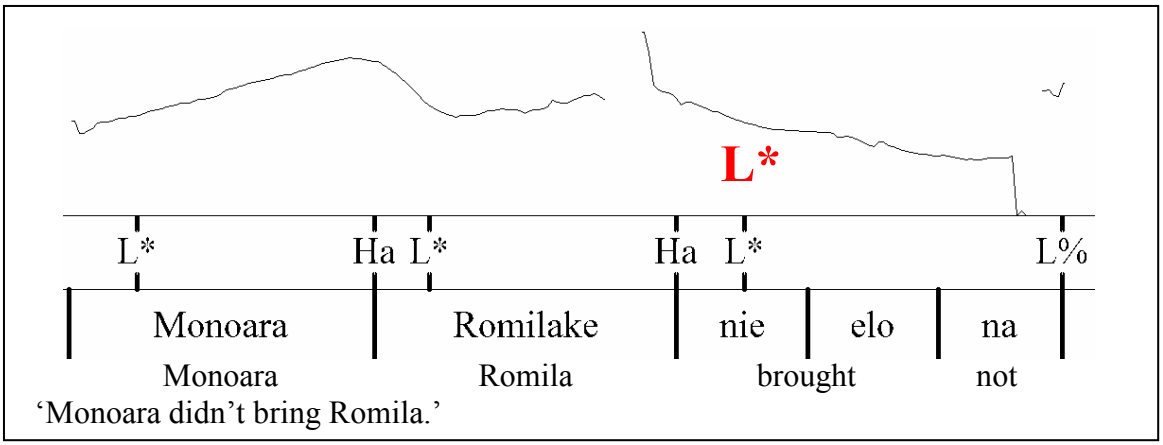


Figure 31. The declarative ip-final AP [nie elo na] 'didn't bring' bears a low pitch accent (L*). [Na06][59]

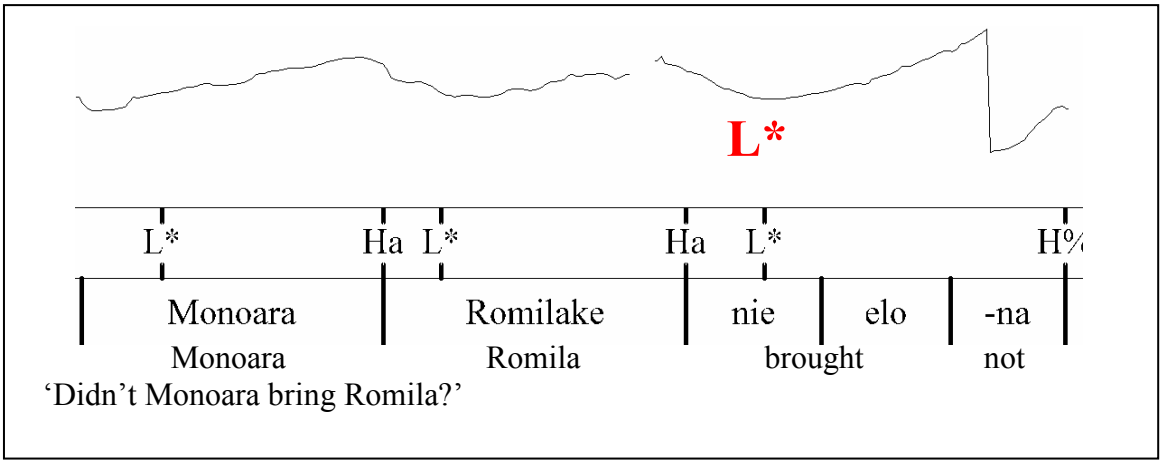


Figure 32. The interrogative ip-final AP [nie elo-na] 'didn't bring' bears a low pitch accent (L*).[60] [Tu06]

[59] This sentence was written "মোনোয়ারা রমিলাকে নিয়ে এল না?" [monoaɹa ɹomilake nie elo na] in Bengali orthography, which can be read as a negative declarative (*i.e.* 'Monoara didn't bring Romila') or a negative yes/no question (*i.e.* 'Didn't Monoara bring Romila?'). A question mark was added to ensure that the negative yes/no question would be elicited. While the speaker in Figure 32 produced the sentence as a negative yes/no question as expected, the speaker in Figure 31 presumably missed the question mark as she read it as a negative declarative.

[60] What looks like a sharp drop-off in pitch during the final syllable of this sentence is actually pitch halving—the reduction of high pitch by 50% in pitch tracking software. Thus, the pitch is in fact rising to a super-high range in the final syllable as part of the high IP boundary tone (H%). The pitch range was not adjusted for this example as the details of pre-IP edge tones would be blurred.

The current model analyzes both the sentences in Figure 31 and Figure 32 as bearing low pitch accents (L*) in the IP-final AP, which is in line with the model proposed for Eastern Bengali in Michaels & Nelson (2004). However, Hayes & Lahiri (1991), Lahiri & Fitzpatrick-Cole (1999), and Selkirk (2006) would have predicted either a high pitch accent (H*) or no pitch accent[61] on the final phrase of the declarative sentence (Figure 31). This may suggest another way in which Western Standard Bengali (*i.e.* Standard Kolkata Bengali), analyzed in Hayes & Lahiri (1991), Lahiri & Fitzpatrick-Cole (1999), and Selkirk (2006) differs from Eastern Bengali, studied in Michaels & Nelson (2004) and that of Bangladeshi Standard Bengali, analyzed in the current model.

Of course, any pitch accent will appear high relative to the low IP boundary tone (L%), which often reaches the floor of a speaker's normal pitch range (see §10.1.2). Considering this, how can we be sure that the pitch accent preceding a low IP boundary tone (L%) is in fact also low, given the fact that low (L*) and high (H*) pitch accents are described as being in complementary distribution in previous models? I support the current model's proposal that low pitch accents (L*) occur in both declaratives and interrogatives, using measurements of the actual pitch during the IP-final stressed syllable. If IP-final APs only take a high pitch accent (H*) in basic declaratives and a low pitch accent (L*) in basic interrogatives, as per Hayes & Lahiri (1991) and subsequent models

[61] Most of the examples of declarative sentences in Hayes & Lahiri (1991) in fact lack the declarative high pitch accent (H*). The reasoning given is that the preverbal word is often focused, causing it to bear the IP-final pitch accent (which is low) instead of the IP-final phrase.

of Kolkata Bengali, there should be a measurable difference in pitch between the last pitch accent of a question and the last pitch accent of a statement. But if IP-final APs take the same pitch accent type—in this case, a low pitch accent (L*)—in both basic declaratives and interrogatives, as per Michaels & Nelson (2004) and the current model, the IP-final pitch accent's F_0 level will not be higher in declaratives than in interrogatives. To test this hypothesis, declarative-interrogative sentence pairs composed of the sentence [monoaɹa ɹomilake nie elo] 'Monoara brought Romila' were compared across the subjects of Experiment I, twelve of whom produced the IP-final pitch accent on the syllable [ni].[62] The F_0 values of twelve subjects were measured and listed in Table 17 below.

[62] Out of 20 subjects recorded for this task, eight had to be excluded from this measurement. Four speakers produced interrogative sentences that could not be used as they applied narrow focus to the preverbal object [ɹomilake] 'Romila-ACC', deleting the pitch accent on [nie] (see Chapter Four). Another two speakers varied in their assignment of an additional pitch accent to the second half of the compound verb [nie elo] 'brought'. The remaining two speakers did not produce a yes/no question at all, instead producing nearly identical declarative sentences for both elicitations.

Speaker	Declarative (…L%)	Interrogative (…HL%)
Da	185	192
Jh	208	220
Sf	73	87
Ba	91	124
Do	193	210
Na	193	208
BM	186	249
Sh	118	169
Re	112	152
Tu	225	244
Bo	240	285
Jo	222	251
Average	**170.5**	**199.3**

Table 17. F_0 values in Hz during the final pitch-accented syllable [ni] in the sentence [monoaɹa ɹomilake nie elo] 'Monoara brought Romila' in both declarative (L%) and interrogative (HL%) contexts. Measurements were compared within speakers.

It turns out that the F_0 value during the phonologically stressed syllable [ni] in the compound verb [nie elo] 'brought' in the statement [monoaɹa ɹomilake nie elo] 'Monoara brought Romila' is in fact no higher than that of the corresponding yes/no question [monoaɹa ɹomilake nie elo?] 'Did Monoara bring Romila?'. On the contrary, the pitch accented syllable of the *interrogative* has a significantly higher pitch than that of the declarative [paired $t(11) = 5.61$, $p < 0.05$]. This is no doubt related to the fact that yes/no interrogative sentences end in a falling IP boundary tone (HL%), which begins with a H tonal target that can raise the pitch of the final pitch accent (see §10.1.3 under *Upstep of*

low pitch accents). In other words, measuring within speaker, the final pitch accent in a declarative is not significantly higher than in the corresponding interrogative, indicating that declaratives in Bangladeshi Standard Bengali in fact do end with default low pitch accents (L*).

Some readers may question whether there is any need to posit a low pitch accent (L*) on an AP preceding the low IP boundary tone (L%). Indeed, in some examples (*e.g.* Figure 30), it may appear that what is labeled as a low pitch accent (L*) could simply be a point along the interpolation of pitch from the high AP boundary tone (Ha) of the ip-penultimate AP to the low IP boundary tone (L%). However, in most cases, it is clear that an additional L target is borne on the first stressed syllable of the IP-final AP, as in Figure 33 below. Without the additional L target, it would be difficult to explain the bend in the pitch track during the stressed syllable [ma].

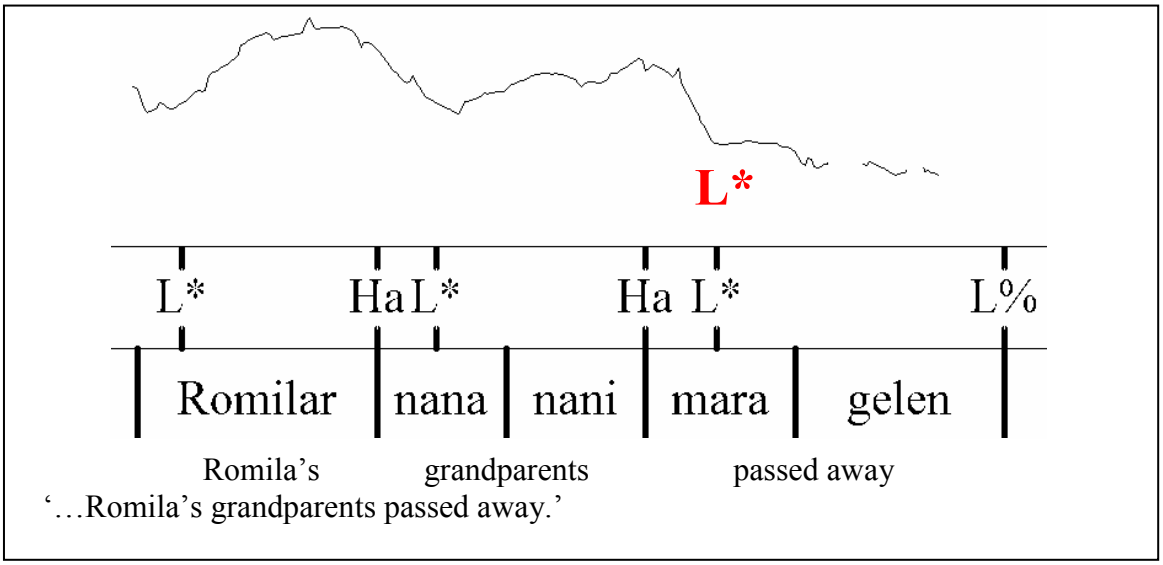
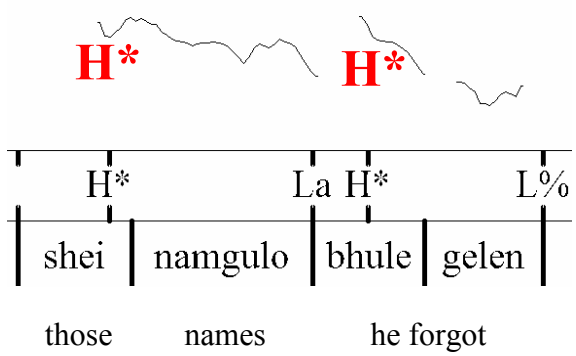


Figure 33. The ip-final AP [maɹa gelen] 'passed away' bears a low pitch accent (L*), whose L target is clearly visibly between the preceding high AP boundary tone (Ha) and the following low IP boundary tone (L%). The irregular pitch tracking during [gelen] is due to creaky phonation. [Tu49]

High pitch accent (H*)

In addition to its use in the smooth fall pattern (H*…La) seen in the naturalistic recordings of nonstandard eastern dialects of Bengali, the high pitch accent (H*) is found in the scripted recordings of Bangladeshi Standard Bengali, typically borne on APs associated with an element of irony, surprise, sarcasm, or otherwise sudden or unexpected information. The high pitch accent (H*) is realized as a local pitch maximum during a phonologically stressed syllable, as in Figure 34 below (repeated from Figure 15), where the ip-final AP [bʰule gelen] 'forgot-HON' bears a high pitch accent (H*) to

signal that it is ironic or surprising that the uncle also forgot the names Rumu couldn't remember.[63]



'…he forgot those names.'

Figure 34. The ip-final AP [bʰule gelen] 'forgot-HON' bears a high pitch accent (H*), to mark unexpected information. Note the H* downtrend across the two APs. [Ba51]

As can be seen in Figure 34, and as mentioned in the description of the smooth fall pattern (H*…La), successive high pitch accents (H*) follow a pattern of downtrend similar to the pattern seen in high AP boundary tones (Ha). Another such example is shown in Figure 35 below. However, because long stretches of successive high pitch accents (H*) are relatively uncommon in the current study's corpus of data, it is not possible to be totally certain of the regularity of H* downtrend without additional data.

[63] The stimuli were not controlled for surprising or ironic information; I am relying on my own intuition and the intuition of my primary consultant to determine the surprise focused status of words.

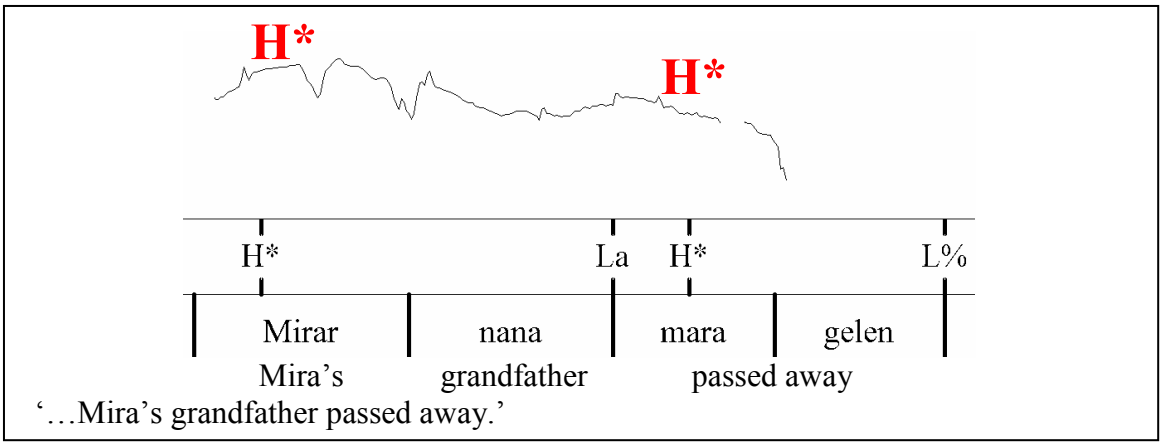


Figure 35. The ip-final AP [maɹa gelen] 'passed away-HON' bears a high pitch accent (H*), possibly marking unexpected information. Note the H* downtrend across the two APs. [By37]

In the naturalistic data of nonstandard dialects, the high pitch accent (H*) was also found in ip-final position even when it is unclear if it is being used to highlight any surprising or unexpected information. As suggested in Figure 36 and Figure 37 below, the high pitch accent (H*) can be considered unmarked in at least some instances in some nonstandard dialects, as can the smooth fall pattern (H*…La).

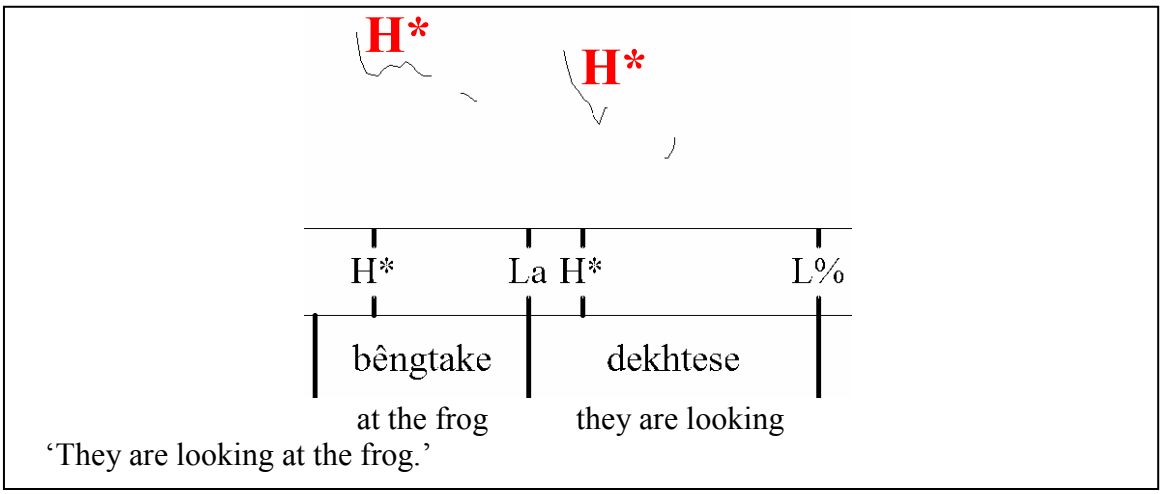


Figure 36. The final AP [d̪ekʰt̪ese] 'they are looking' bears a high pitch accent (H*), despite the lack of sudden or unexpected information. The phrase was produced by a speaker from the Dhaka suburbs, using his native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. Note the H* downtrend across the two APs. [ReS9]

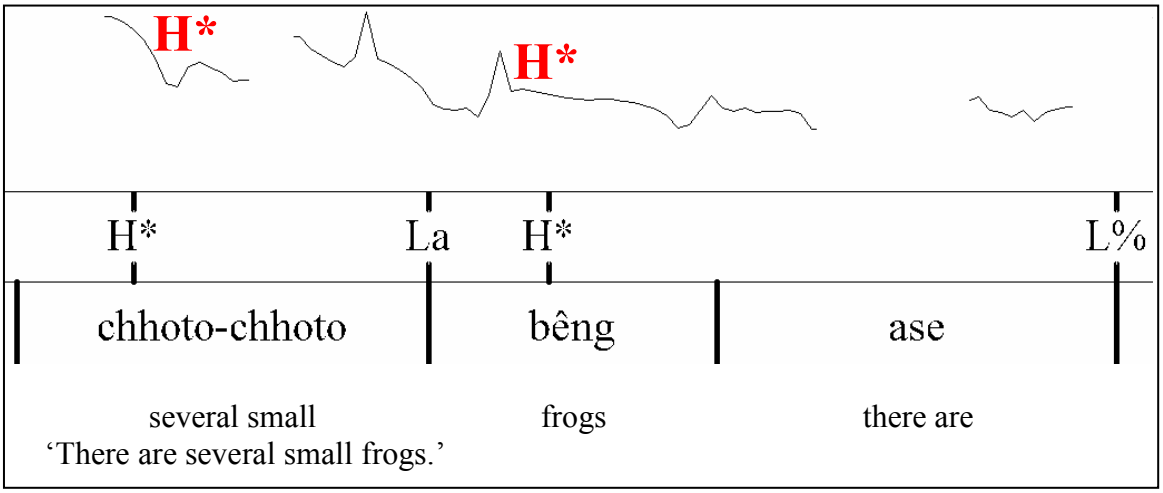


Figure 37. The final AP [bɛŋ ase] 'there are frogs' bears a high pitch accent (H*), despite the lack of sudden or unexpected information. As expected, the preceding AP [tɕʰototɕʰoto] 'several small' bears a smooth fall pattern (H*…La). The phrase was produced by a speaker from the Dhaka suburbs, using his native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. Note the H* downtrend across the two APs. [ReS186]

Although most high pitch accents (H*) are preceded by smooth falling patterns (H*…La), examples of high pitch accents (H*) preceded by smooth rising patterns (L*…Ha) are also found, as in Figure 38 below. As it is not preceded by other high pitch accents (H*), it cannot be determined if an ip-final high pitch accent (H*) following smooth rises (L*…Ha) obeys H* downtrend.

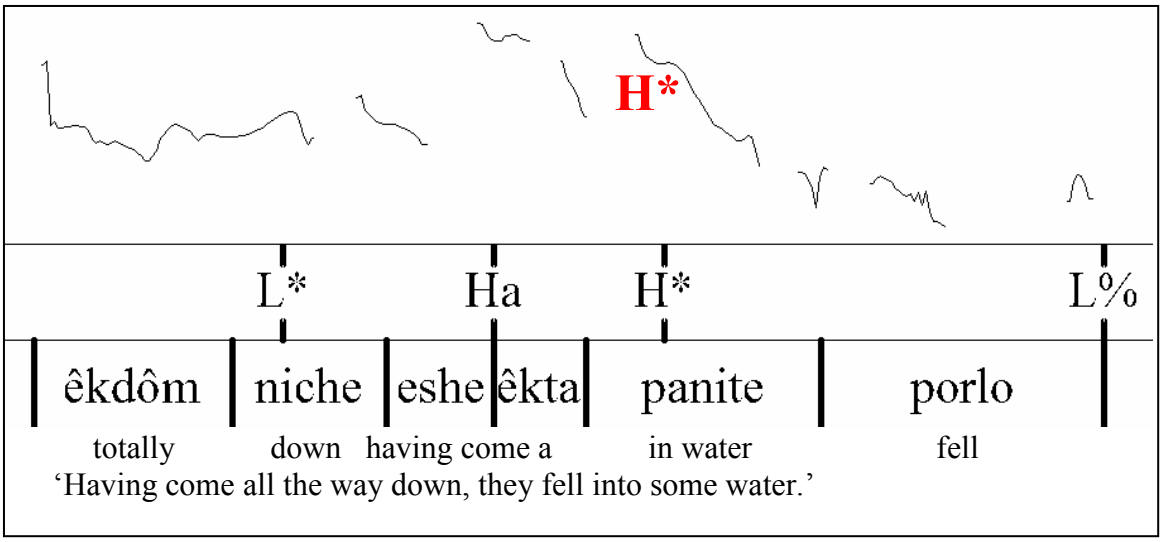


Figure 38. The final AP [ɛkta paniṯe poɹlo] 'they fell into some water' bears a high pitch accent (H*), signaling sudden or unexpected information. The preceding AP [ɛkd̪ɔm nitɕe eʃe] 'having come all the way down' bears a smooth rise pattern (L*…Ha). The phrase was produced by a speaker from Mymensingh District (in Dhaka Division), using her native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. [JhS104]

In Figure 39 below, the ip-final AP [holo na] 'didn't happen' bears a high pitch accent (H*) to signal that the fact that the trip to Narayanganj did not come to fruition is surprising, ironic, or in some way unexpected.

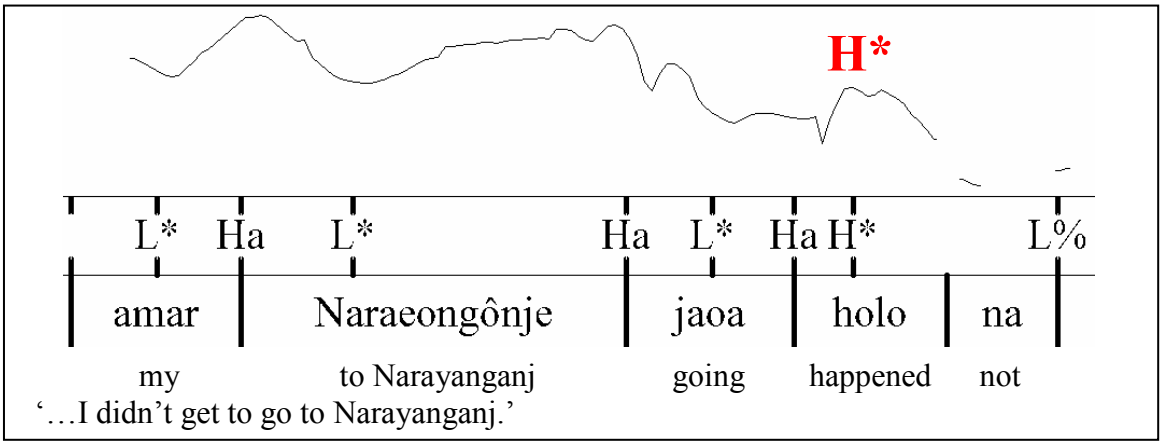


Figure 39. The ip-final AP [holo na] 'didn't happen' bears a high pitch accent (H*), to mark unexpected information. Although not as common as smooth falls (H*…La) in this position, the preceding APs bear smooth rises (L*…Ha). [To34]

In Figure 40 and Figure 41 below, the ip-final AP [ʃɔbaj dʑane] 'everyone knows' bears a high pitch accent (H*) to signal that the speakers feel the hypothetical listener (the subjects were simply speaking into a microphone) would find it surprising, unexpected, or ironic that everyone knows Monoara brought Romila.

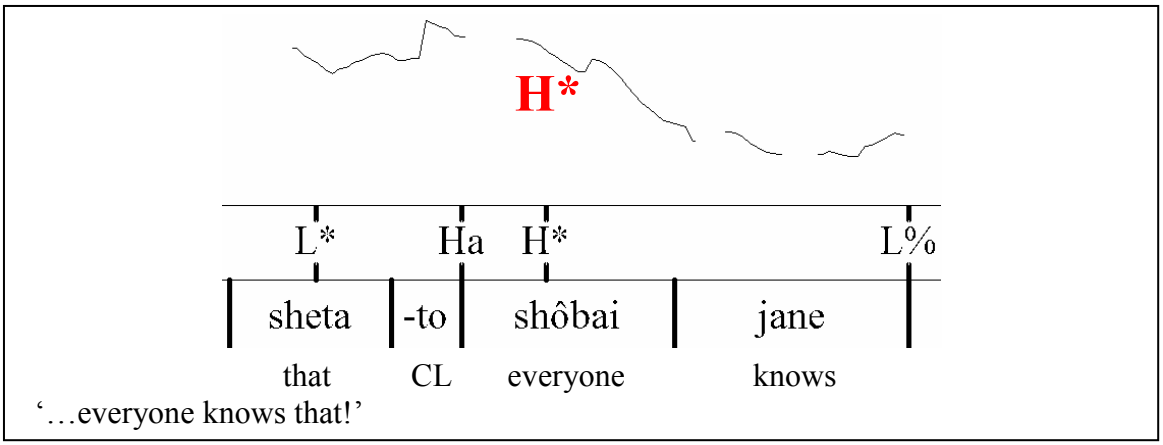


Figure 40. The ip-final AP [ʃəbaj dʑane] 'everyone knows' bears a high pitch accent (H*), marking unexpected information. In this case, the preceding AP bears a smooth rise (L*…Ha). Note that the final syllable of [dʑane] 'knows' does not bear rising pitch; the pitch track is unreliable on that syllable due to creaky phonation. [To15]

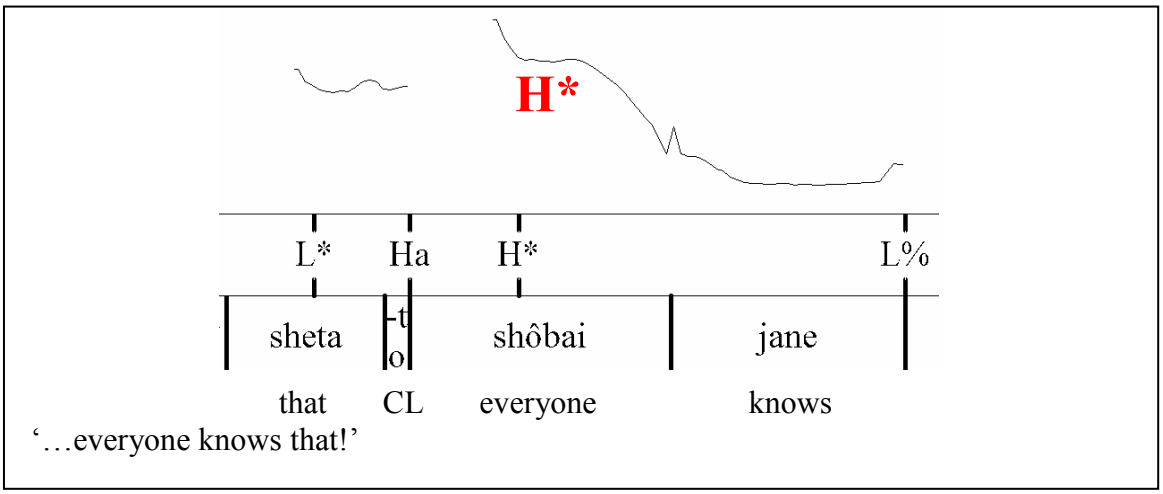


Figure 41. The ip-final AP [ʃəbaj dʑane] 'everyone knows' bears a high pitch accent (H*), marking unexpected information. In this case, the preceding AP bears a smooth rise (L*…Ha). [Re15]

In Figure 42 below, the AP [ze sɛɹada uldaʂ poɹlo] 'that the boy fell turned upside-down' bears a high pitch accent (H*), signaling that the fact that the boy fell turned

upside down is sudden or unexpected. The phrase was produced by a speaker from Netrakona District (in Dhaka Division), using his native Eastern Bengali in a recording session of naturalistic speech.

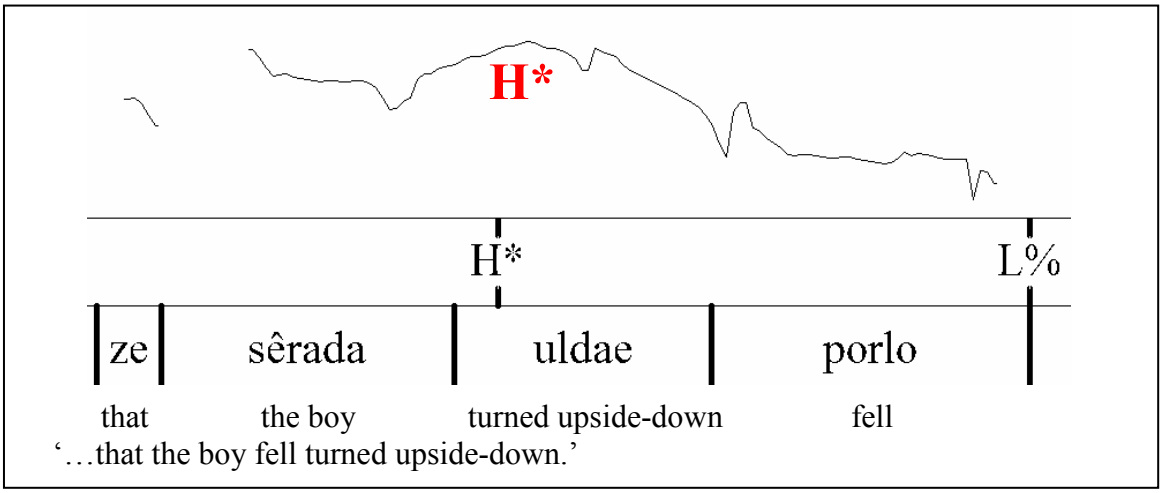


Figure 42. The AP [ze sɛɹada uldaɛ̯ poɹlo] 'that the boy fell turned upside-down' bears a high pitch accent (H*), signaling sudden or unexpected information. The phrase was produced by a speaker from Netrakona District (in Dhaka Division), speaking Eastern Bengali in a recording session of naturalistic speech. [FlS65]

In some situations, it is clear that the ip-final high pitch accent (H*) is violating what appears to be H* downtrend. High pitch accents (H*) that exceed the pitch of the preceding high pitch accent (H*) are labeled with an "f" diacritic, explained more thoroughly in Chapter Four. These focused high pitch accents (fH*) indicate a type of focus, specific to sudden, unexpected, or surprising information. One verb that seems to be very often associated with sudden or unexpected information in both Bangladeshi Standard Bengali and nonstandard eastern dialects is [pɔɹ]~[poɹ]~[pojɹ]- 'fall' ([pɔɾ]~[poɾ]- in Standard Western Bengali), as in Figure 43 below, produced by a speaker

from Mymensingh District (an Eastern Bengali dialect area) in Experiment II. The suddenness and surprise often inherently associated with the act of falling might contribute to its frequent association with the high pitch accent (H*).

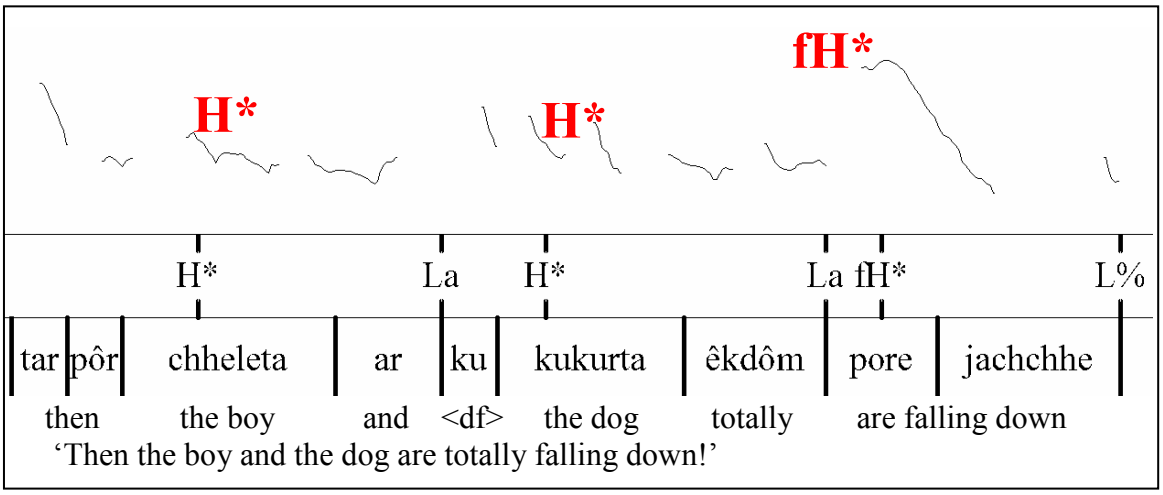


Figure 43. The final AP [poɹe dʑatɕʰːe] 'are falling down' bears a focused high pitch accent (fH*), which violates H* downtrend. The phrase was produced by a speaker from Mymensingh District (in Dhaka Division), using her native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. [JhS98]

Although it is more frequently seen in naturalistic data from Experiment II, both ip-finally and as part of the smooth fall AP tonal pattern (H*…La), the high pitch accent's (H*) use as a marker of irony or new information even in more formal scripted speech (as collected in Experiment I) motivates its inclusion in the tonal inventory of Bangladeshi Standard Bengali as well. The focused high pitch accent (fH*) is described further in Chapter Four.

Summary

The corpus of data collected in the current study suggests that the postlexical tonal inventory of Bangladeshi Standard Bengali includes three pitch accents—low (L*), high (H*), and rising (L*+H)—as well as two AP boundary tones—low (La) and high (Ha). The AP boundary tones can only occur on the right edge of the AP hosting the pitch accent of the opposite tonal target, creating two of the three basic tonal patterns: smooth rise (L*…Ha) and smooth fall (H*…La). A third basic tonal pattern, the sharp rise (L*+H), includes two tonal targets joined as one pitch accent; this pattern bears no AP boundary tone.

The choice of pitch accent is related to pragmatics: the high pitch accent (H*) is associated with unexpected or surprising information, the rising pitch accent (L*+H) with certain types of focused information, and the low pitch accent (L*) in default situations. The rising pitch accent (L*+H) can optionally be raised to a shallow mid rise (^L*+H) when following voiceless consonants. Lastly, the smooth fall pattern (H*…La) appear to be largely restricted to the inventory of nonstandard eastern dialects, leaving its status in Bangladeshi Standard Bengali unclear.

I now move on to describe the tones associated to the edges of higher prosodic domains: intermediate phrases (ip) and intonation phrases (IP).

10.1.2 Higher prosodic domains (ip and IP)

Bengali is a heavily edge-marking language, with three prosodic units marked tonally on their right boundary. From largest to smallest, these three units are the Intonation Phrase (IP), the Intermediate Phrase (ip), and the Accentual Phrase (AP). The boundary tones associated with these three units are transcribed T% (IP boundary tone), T- (ip boundary tone), and Ta (AP boundary tone), where T stands for H, L, or certain combinations thereof. Unlike the boundary tones of other languages (*e.g.* English), T%, T-, and Ta do not cooccur, even when the boundaries to which they are associated coincide. The right edge of an IP, for example, is typically assumed to also be the right edge of an ip and the right edge of an AP, given the Strict Layer Hypothesis (Selkirk 1984, 1986; Nespor & Vogel 1986), and yet only the tone belonging to the edge of the highest prosodic category – in this situation, the IP boundary tone (T%) – will be realized at such a phrase edge. The ip boundary tone (T-) and AP boundary tone (Ta) are overridden. In most situations, there are no tones of the shape T-T% or TaTi in the current model,[64] unlike the tone stacking seen in models of other languages (*e.g.* MAE_ToBI model of American English, CatToBI model of Catalan, GRToBI model of Greek).

Intermediate phrase (ip) boundary tones (T-)

The intermediate phrase (ip) is a grouping of APs that form a tight syntactic unit, often corresponding to a small phrase (*e.g.* postpositional phrase, topic, adverbial) and

[64] See Chapter Four for a discussion of the limited tonal stacking seen in ip-final encliticized focused constituents.

occasionally to a clause (*e.g.* relative clause, if-clause, because-clause). Unlike the AP, the ip does not appear to have a prosodic head (*cf.* nuclear pitch accent as the ip head in English). Most ips are coterminous with an IP, leaving the ip boundary tone overridden. However, non-IP-final ips are frequent in longer sentences, and thus controlling sentences for overall length can motivate the production of more ips. The right edge of a non-IP-final ip is marked by a boundary tone, either high (H-), low (L-), rising (LH-), or falling (HL-), all of which are subject to the *ip boundary tone locality constraint*, which states that the ip boundary tone can only affect the pitch during the ip-final syllable. Each ip boundary tone is described in detail below with example pitch tracks.

High ip boundary tone (H-)

The high ip boundary tone (H-) marks the right edge of smaller ips (*e.g.* topics, postpositional phrases, adverbials), and is characterized by a sharp F_0 rise on the ip-final syllable. In Figure 44 below, the high ip boundary tone (H-) demarcates the ip [ɹomilaɹ nana naniɹa] 'Romila's grandparents' as a topicalized element, which can be translated as something like 'As for Romila's grandparents…'.

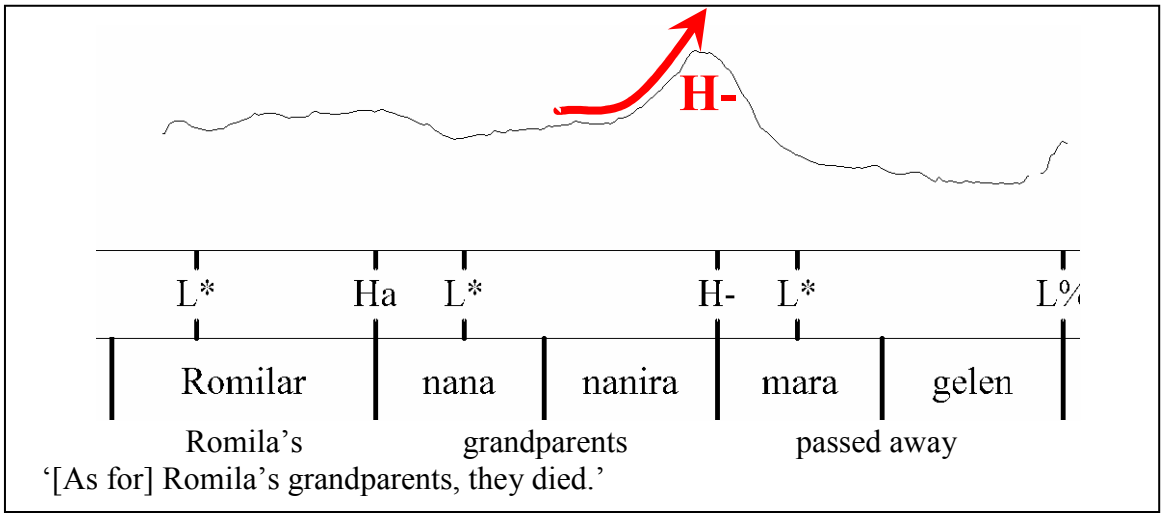


Figure 44. The topicalized phrase [ɹomilaɹ nana naniɹa] '(As for) Romila's grandparents' is marked on its right edge by a high ip boundary tone (H-), realized as a sharp F_0 rise on the ip-final syllable. (This differs from LH-, which includes a dip in F_0 before the final rise). [Na49]

In Figure 45 below, the high ip boundary tone (H-) demarcates the ip [monoaɹa] '(As for) Monoara…' as a topicalized element. Note that this ip is composed of only one AP.

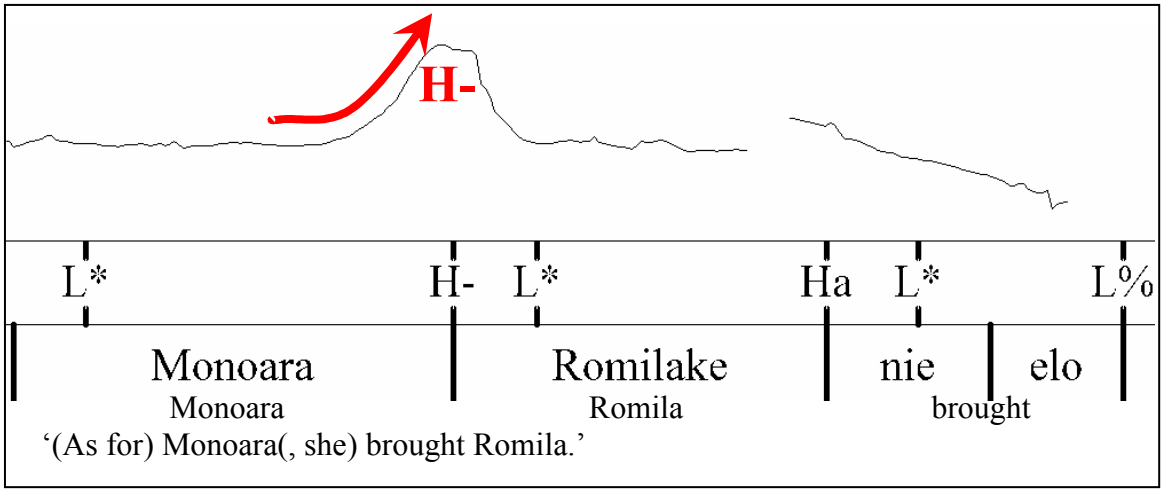


Figure 45. The topicalized phrase [monoaɹa] '(As for) Monoara…' is marked on its right edge by a high ip boundary tone (H-), realized as a sharp F_0 rise on the ip-final syllable. [Do01]

In Figure 46 below, the subject [miɹaɹ nana] ‘Mira’s grandfather’ bears a high ip boundary tone (H-). Compare the smooth rise contour (L*…Ha) on [miɹaɹ] ‘Mira’s’ with the sharper elbow in the following AP [nana] ‘grandfather’.

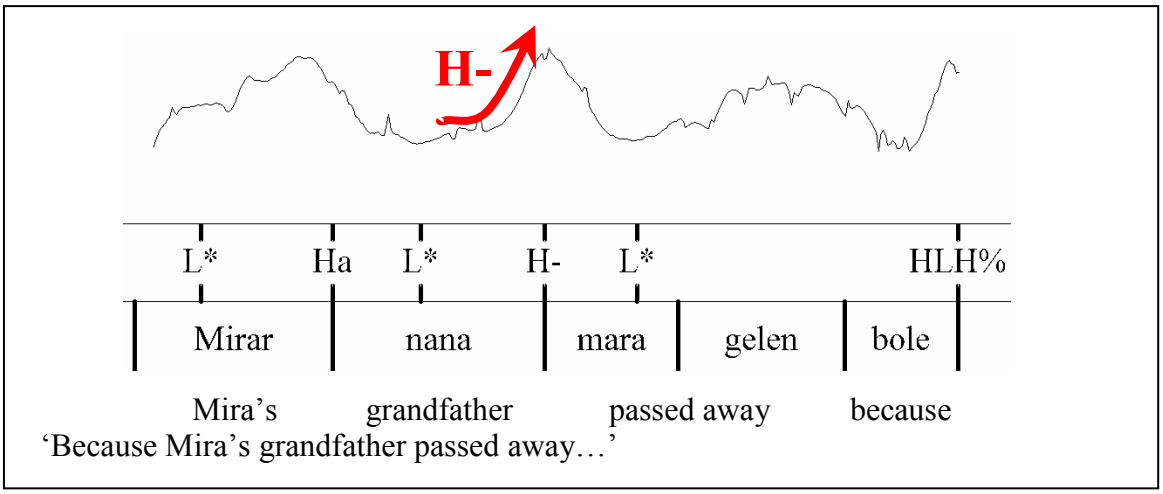


Figure 46. The subject [miɹaɹ nana] ‘Mira’s grandfather’ is marked on its right edge by a high ip boundary tone (H-), realized as a sharp F_0 rise on the ip-final syllable. [To34]

In Figure 47 below, the high ip boundary tone (H-) demarcates a long noun phrase (NP), [ɹumu dʑej namgulo mone ɹakt̪e paɹe ni] ‘the names that Rumu couldn’t remember’.

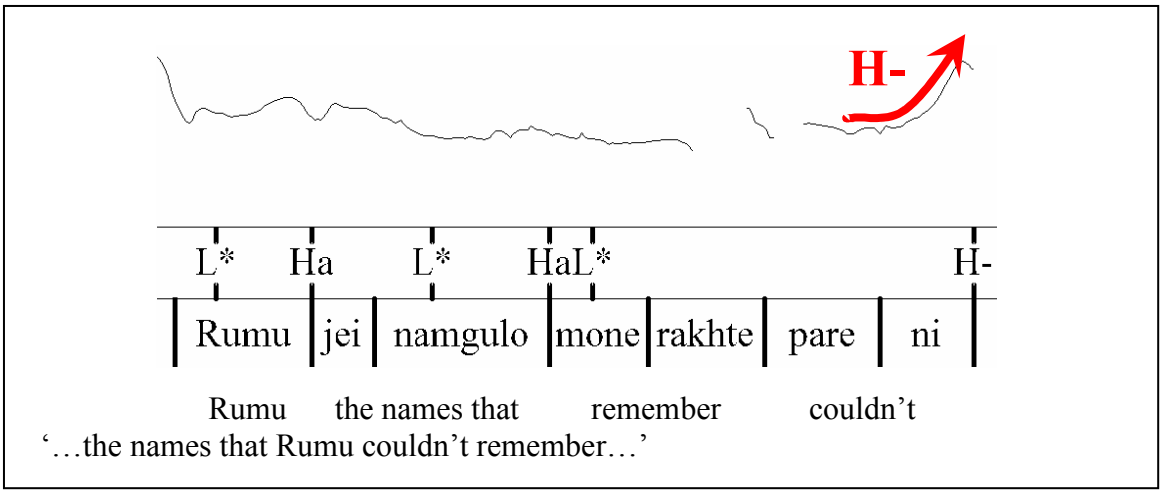


Figure 47. The ip [ɹumu dʑej namgulo mone ɹakt̪e paɹe ni] 'the names that Rumu couldn't remember' is marked on the right edge by a high ip boundary tone (H-), realized as a sharp F_0 rise on the ip-final syllable. [Tu51]

The high ip boundary tone (H-) can be mistaken for the high AP boundary tone (Ha), in that it occurs at the ends of phrases that do not seem to have any particularly uniform syntactic meaning, unlike high IP boundary tones (H%), which always occur at the edges of interrogatives. However, there are significant characteristics distinguishing the high AP and ip boundary tones (Ha, H-), both of which can be seen clearly in the examples above. While high AP boundary tones (Ha) are successively lower in pitch across a sentence due to Ha downtrend, high ip boundary tones (H-) consistently reach a higher pitch than the preceding high AP boundary tone (Ha). In this way, a high ip boundary tone (H-) reaches a higher pitch than a high AP boundary tone (Ha) would in its place. This distinction in relative pitch height (including the distinction between H- and H%) is discussed in greater detail in §10.1.4. Furthermore, the rise in pitch for the high ip boundary tone (H-) is realized primarily *within* the ip-final syllable, due to the ip

boundary tone locality constraint. This creates a pitch elbow between the penultimate and ultimate syllables, while the high AP boundary tone (Ha) causes pitch interpolation across the entire word, creating a smooth rise in pitch from the preceding low pitch accent (L*). The pitch elbow is particularly easy to identify when the high ip boundary tone (H-) occurs far from the preceding pitch accent, as in Figure 48 below.

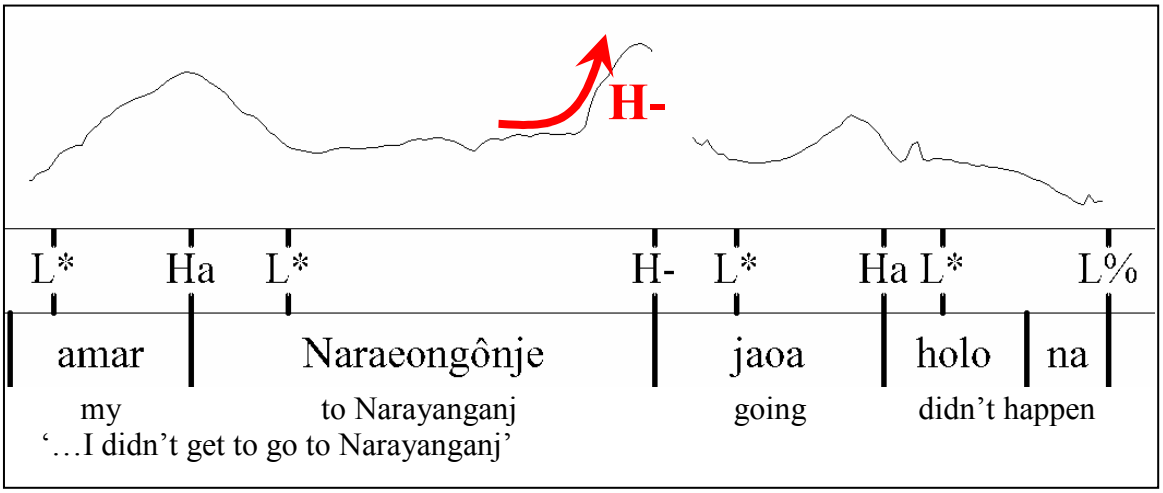


Figure 48. The ip [amaɹ naɹaa̯ongɔndʑe̯] 'my [going] to Narayanganj' is marked on the right edge by a high ip boundary tone (H-), realized as a sharp F_0 rise on the ip-final syllable. [Sh35]

Rising ip boundary tone (LH-)

The rising ip boundary tone (LH-) also occurs at the right edge of long phrases, typically denoting background or known information. It is realized as a fall and rise in pitch during the ip-final syllable, as shown in Figure 49 below.

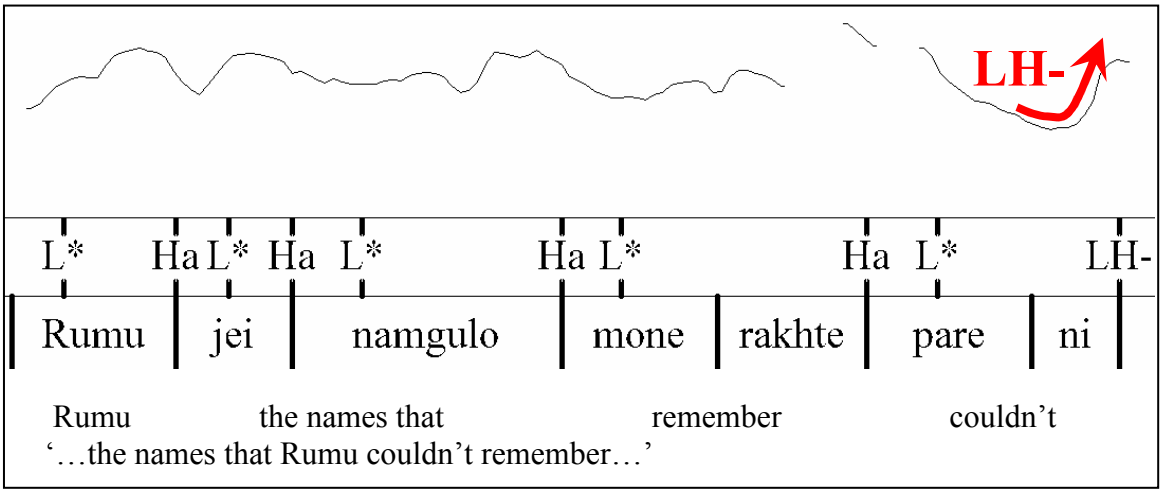


Figure 49. This long noun phrase bears a rising ip boundary tone (LH-). [Ba51]

Figure 50 and Figure 51 are examples of the same sentence produced by two different speakers, both of whom produced rising ip boundary tones (LH-) at the edges of the phrases [adʑ d̪upuɹ bɛlaɕ̞] 'today in the early afternoon' and [dʑumːaɹ namadʑe ʃunlam] 'I heard at Friday prayers'. Note the dipping of pitch from the mid range to achieve the low (L) target of the bitonal boundary tone during the ip-final syllable, in accordance with the ip boundary tone locality constraint.

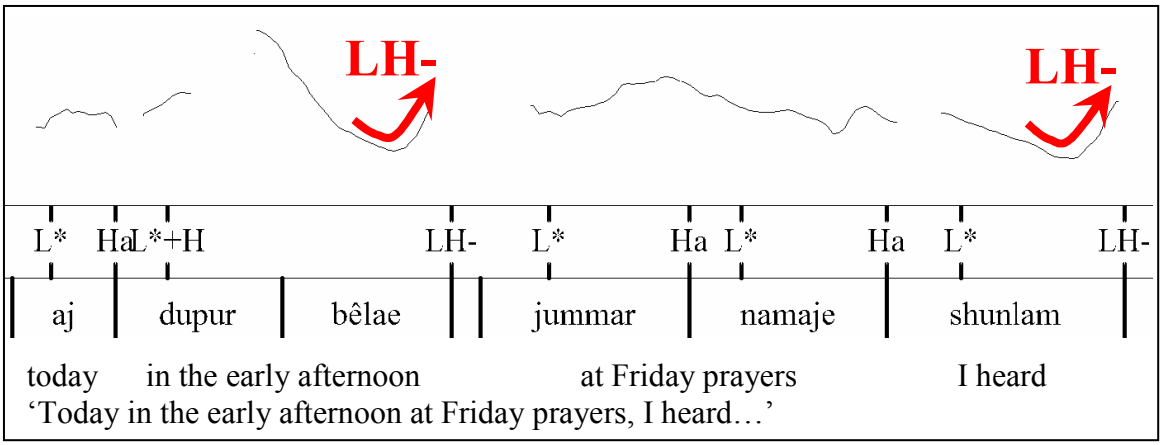


Figure 50. The ips [adʑ d̪upuɹ bɛlaɛ̯] 'today in the early afternoon' and [dʑumːaɹ namadʑe ʃunlam] 'I heard at Friday prayers' both bear rising ip boundary tones (LH-) at their right edge. [Sh49]

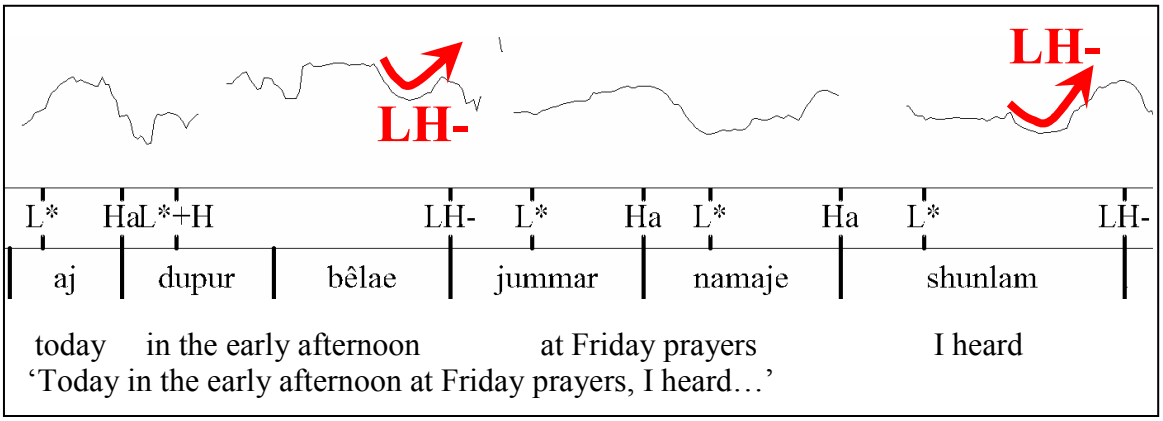


Figure 51. The ips [adʑ d̪upuɹ bɛlaɛ̯] 'today in the early afternoon' and [dʑumːaɹ namadʑe ʃunlam] 'I heard at Friday prayers' both bear rising ip boundary tones (LH-) at their right edge. [Fa49]

Falling ip boundary tone (HL-)

Like the rising ip boundary tone (LH-), the falling ip boundary tone (HL-) occurs at the right edge of long phrases, typically denoting background or known information. It is realized as a rise and fall in pitch during the ip-final syllable, as shown in Figure 52.

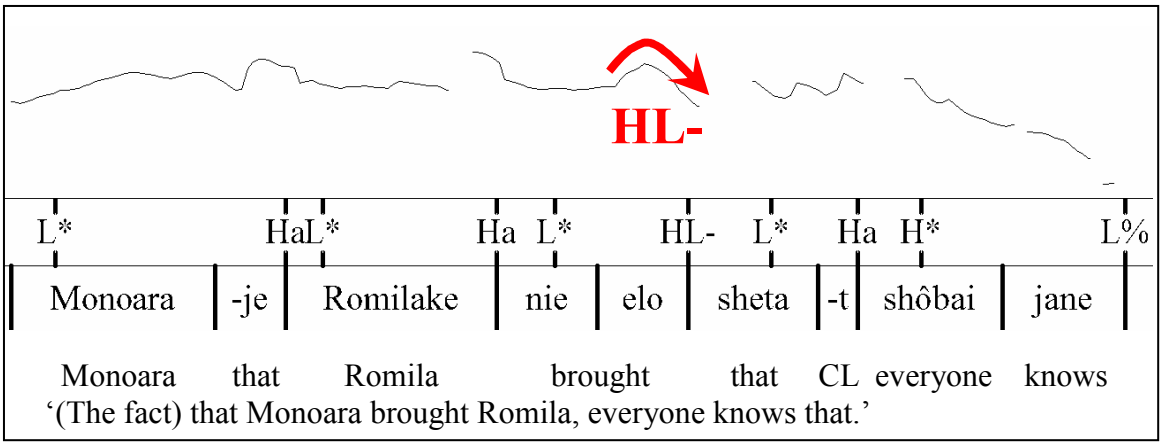


Figure 52. The clause [monoaɹa-dʑe ɹomilake nie elo] '(the fact) that Monoara brought Romila' bears a falling ip boundary tone (HL-) at its right edge. [Do15]

Due to the ip boundary tone locality constraint, pitch is *not* interpolated directly from the previous pitch accent to the H portion of the boundary tone; instead, the pitch of the ip-final pitch accent is either prolonged or slightly interpolated towards the mid range, until immediately preceding the ip-final syllable. However, the first tonal target of the falling ip boundary tone (HL-) can be optionally stretched to the penultimate syllable when the ip ends in a function word, such as the auxiliary verb [gelen] 'went-HON' in the complex verb [maɹa gelen] 'passed away' in Figure 53 below.

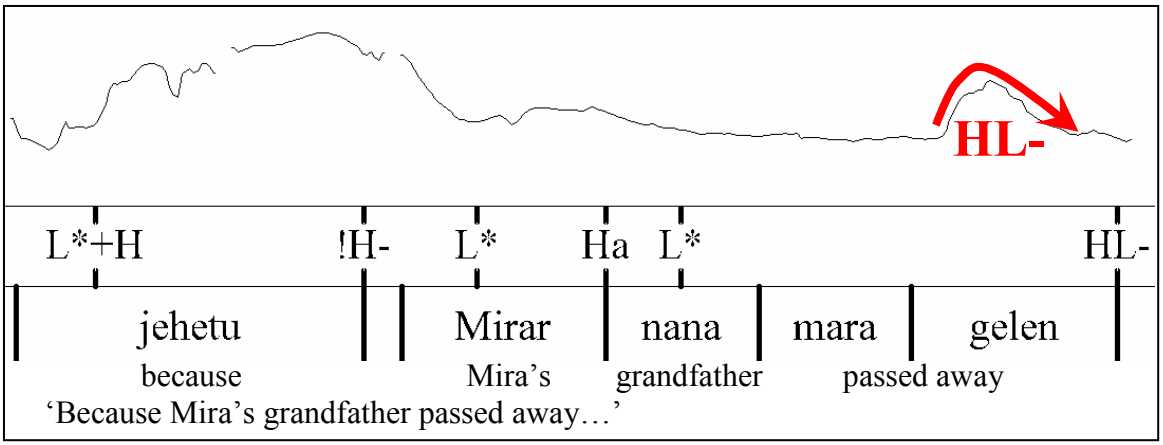


Figure 53. This clause bears a falling ip boundary tone (HL-) at its right edge, stretched across the two syllables of the final function word [gelen] 'went-HON'. See §10.1.3 for a discussion of the downstepped high ip boundary tone (!H-). [Do35]

The falling ip boundary tone (HL-) is not very commonly seen in the scripted data collected in Experiment I; it was found most frequently in the speech of one speaker from Kolkata, with dialect influences from Kushtia District (another Central Bengali dialect area). Further data from that and other regions can reveal the extent of the geographical distribution of this boundary tone.

Low ip boundary tone (L-)

The low ip boundary tone (L-) occurs at the ends of large phrases. Like the other ip boundary tones, it obeys the locality constraint in that it only affects the ip-final syllable. In Figure 54, the low ip boundary tone (L-) follows a rising pitch accent (L*+H), while in Figure 55 it follows a high pitch accent (H*).

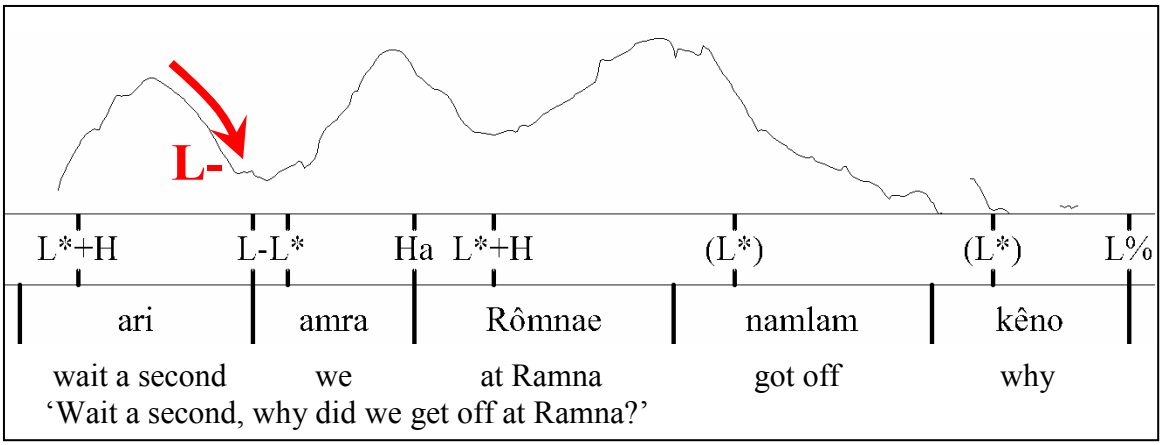


Figure 54. The interjection [aɹi] (roughly equivalent to English 'wait a second') is marked by a low ip boundary tone (L-). See Chapter Four for a discussion of the weakening of pitch accents and loss of AP boundary tones following the focused word [ɹɔmnaɛ̯] 'at Ramna'. [Da48]

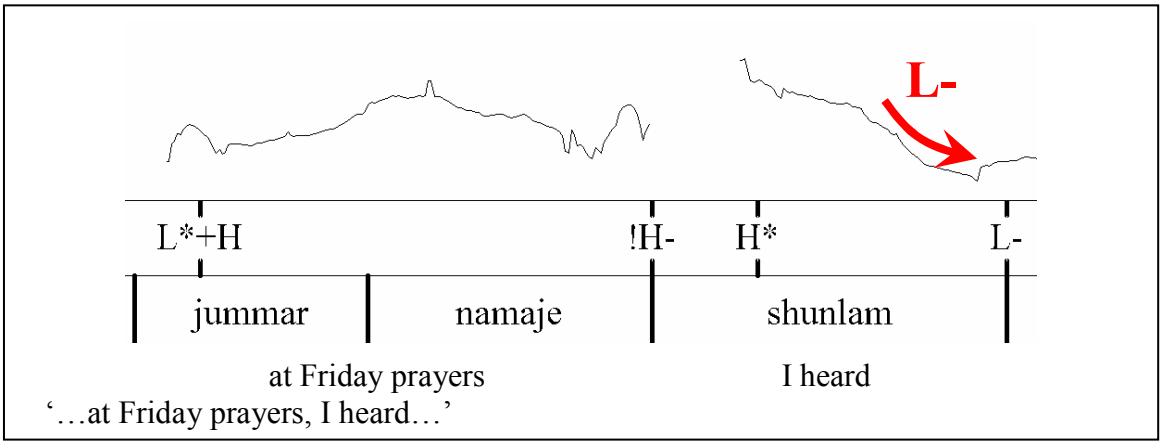


Figure 55. The ip-final word [ʃunlam] 'I heard' bears a high pitch accent (H*) leading into a low ip boundary tone (L-). [Da49]

Although the low ip boundary tone (L-) is easiest to identify following a high (H*) or rising (L*+H) pitch accent, as shown in the examples above, it is also found following the low pitch accent (L*), as in Figure 56 below.

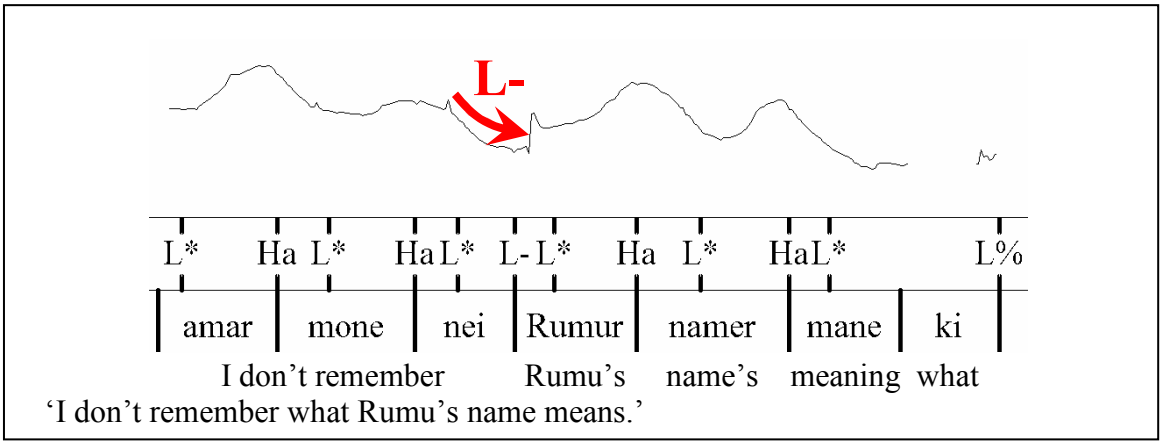


Figure 56. The ip [amaɹ mone nej] 'I don't remember' (*lit.* 'it isn't in my mind') is marked by a low ip boundary tone (L-). [BM32]

As the low ip boundary tone (L-) is rare in the corpus of data collected for the current study, it may be possible to analyze it as a variant of another tone. However, both of the previously mentioned ip boundary tones that include an L target—*i.e.* the rising and falling ip boundary tones (LH- and HL-, respectively)—are normally associated with background or known information, while it is not clear that the low ip boundary tone (L-) can be characterized this way. For now, it can be included as a possible entry in the tonal inventory of Bangladeshi Standard Bengali.

Intonation phrase (IP) boundary tones (T%)

The largest tonally-marked phrase in Bengali is the intonation phrase (IP), which roughly corresponds to a full clause. As with the ip, there is evidence for a tonally-

marked head of this phrase; the phrase is however marked with a following pause, along with one of five boundary tones (*i.e.* low L%, high H%, rising LH%, falling HL%, and dipping HLH%), described below:

Low IP boundary tone (L%)

The low IP boundary tone (L%) marks the right edge of declarative sentences, including focus declaratives. It is also used for wh-questions that bear a rising pitch accent (L*+H) on the wh-word. It is realized as falling pitch interpolated from the IP-final pitch accent, typically followed by an extreme drop in pitch on the final syllable. Pitch tracks often become difficult or almost impossible to read between the IP-final pitch accent and the low IP boundary tone (L%) as this boundary tone is almost universally accompanied by creaky phonation, as shown in Figure 57 and Figure 58 below.

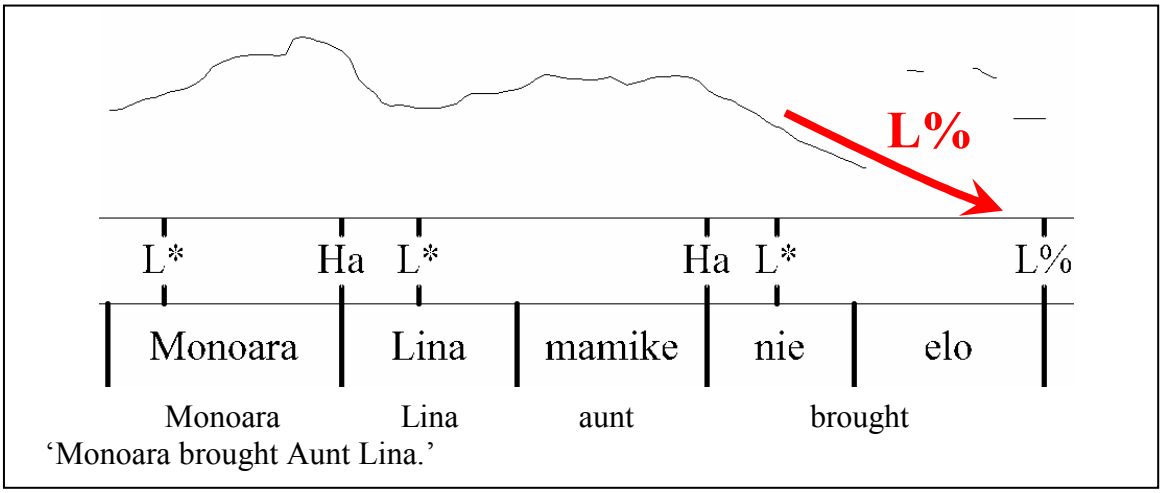


Figure 57. This declarative sentence bears a low IP boundary tone (L%). Note the irregularity of the pitch track during the last two syllable [elo] due to creaky phonation. [Fa24]

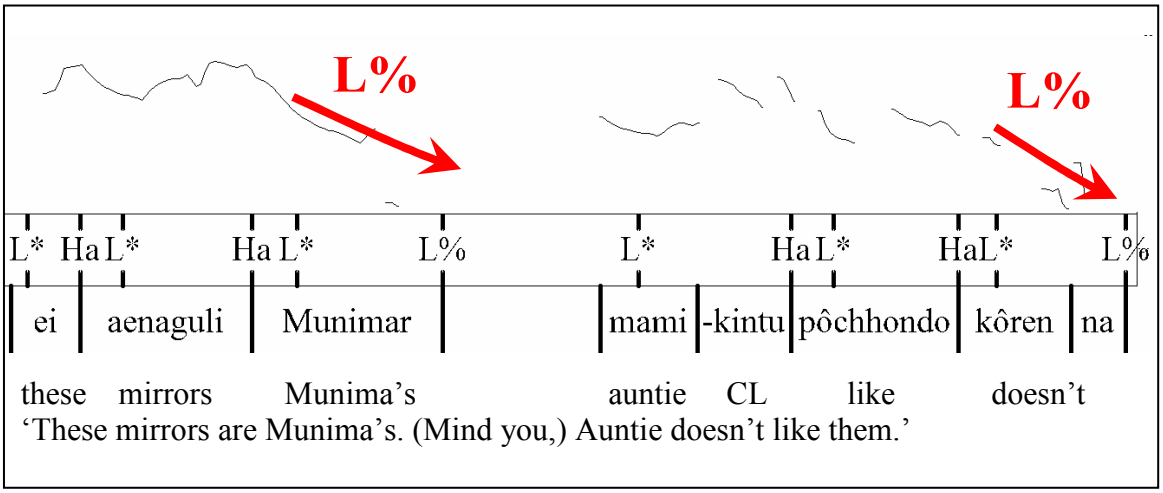


Figure 58. Both the declarative sentences [ej aɛ̯naguli munimaɹ] 'These mirrors are Munima's' and [mami-kin̪tu pɔtɕʰon̪do kɔɹen na] '(Mind you,) Auntie doesn't like them' bear low IP boundary tones (L%). Note the irregular pitch track in both sentences due to creaky phonation approaching the IP boundary. [Da57]

Like the low pitch accent (L*) and high AP boundary tone (Ha), the low IP boundary tone (L%) has been described by all previous studies of Bengali prosody, regardless of the variety studied, and thus I will not go into further detail on its phonetic realization or syntactic distribution.

High IP boundary tone (H%)

The high IP boundary tone (H%) is used for various interrogative sentence types. It is realized as gradually rising pitch from the IP-final pitch accent, with an extreme rise on the final syllable. The "elbow" between the gradual rise and extreme final rise can vary greatly both between and within speakers, and can often appear more or less pronounced depending on the dependability of the pitch tracking software in accurately measuring F_0 at the extremes of a speaker's pitch range, where voice quality changes can affect the pitch track. The most commonly-observed interrogative sentences bearing a high IP boundary tone (H%) are negative yes/no questions (*e.g.* 'Didn't…?' or 'Isn't…?')—often including the enclitic -[na] in either initial (Figure 59) or final (Figure 60) position, or the enclitic -[naki] in initial position (Figure 61).

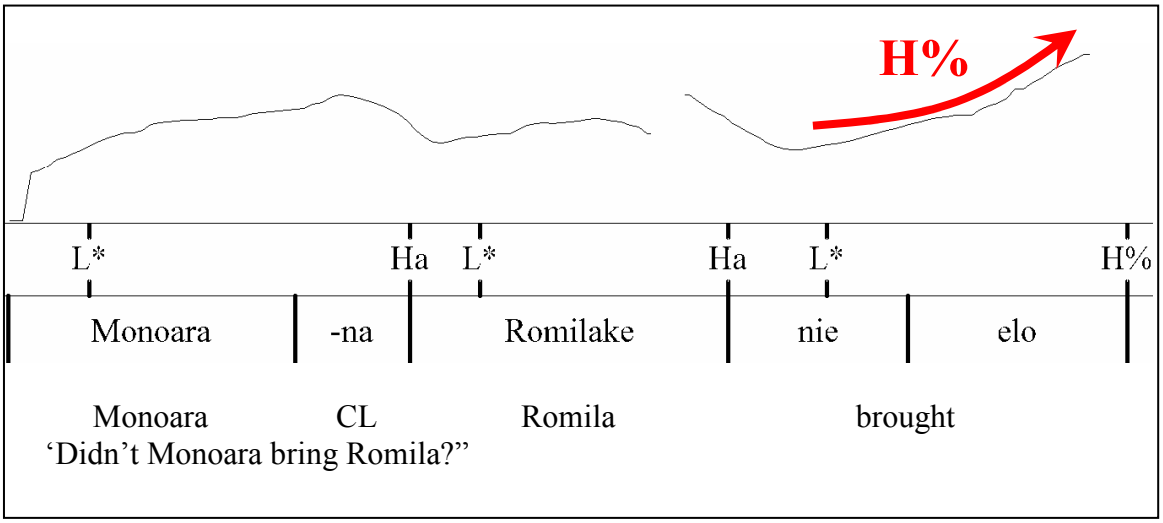


Figure 59. This negative yes/no question bears a high IP boundary tone (H%), realized here as a relatively constant rise. [Az05]

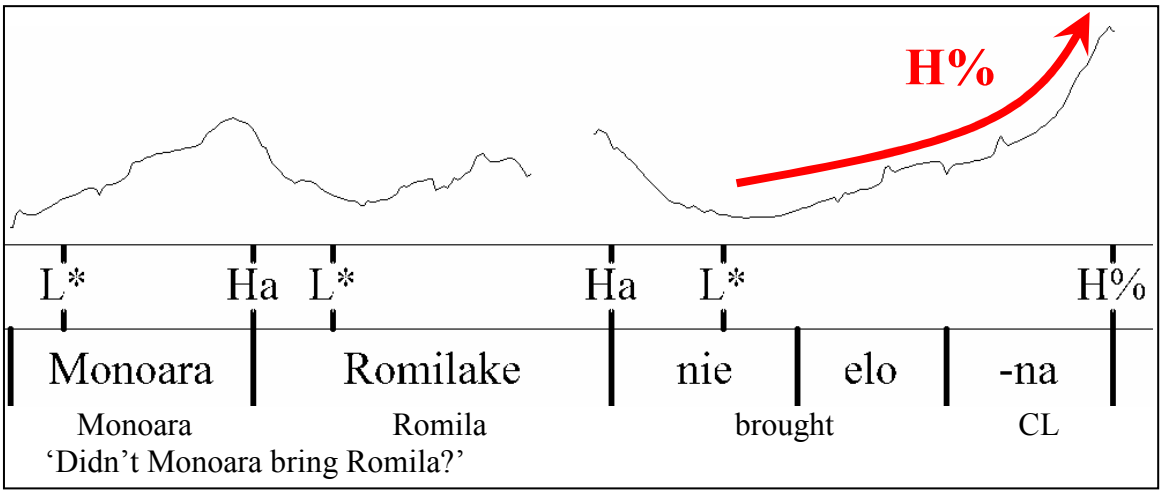


Figure 60. This negative yes/no question bears a high IP boundary tone (H%), realized here with a slight elbow between the gradual rise and extreme final rise. [Fa06]

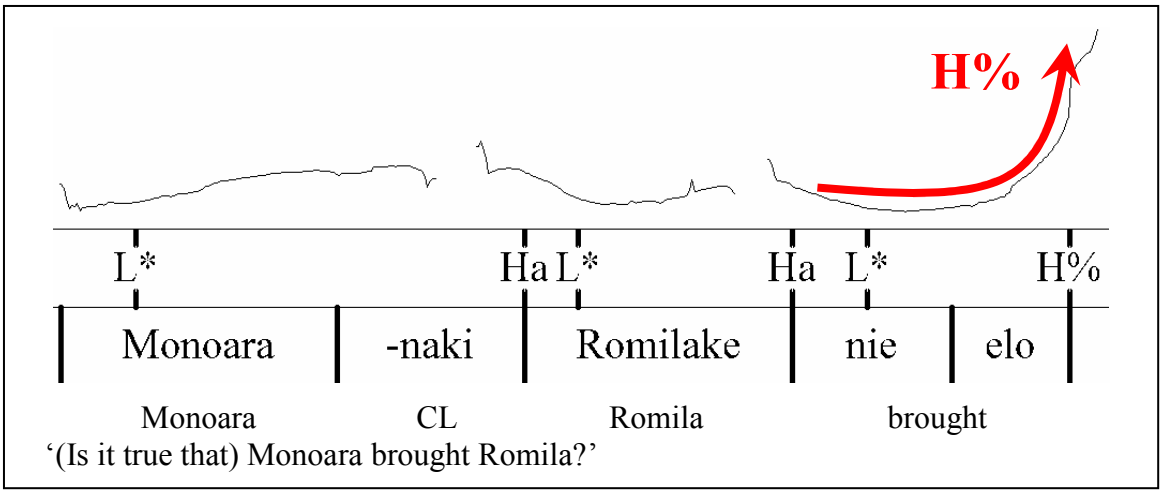


Figure 61. This negative yes/no question bears a high IP boundary tone (H%), realized with a very pronounced elbow between the gradual rise and the extreme final rise [Na07]

Other sentence types bearing a high IP boundary tone (H%) include polite requests (*e.g.* 'May I…?' or 'Would you please…?') as in Figure 62, tag questions (*e.g.* 'Isn't that so?') as in Figure 63, and echo wh-questions (*e.g.* 'You told him *what*?!') as in Figure 64.

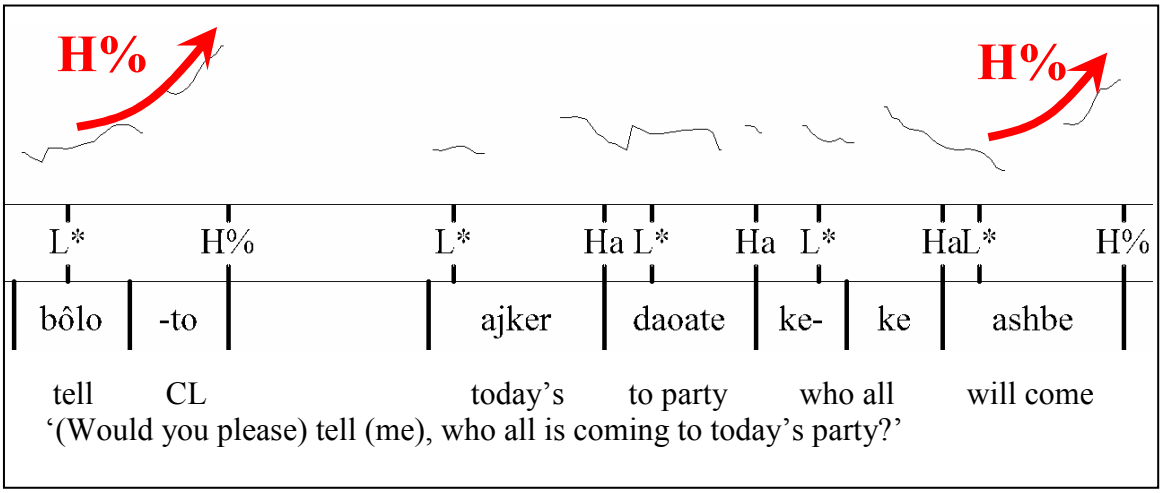


Figure 62. This polite request bears two high IP boundary tones (H%), one after the command itself [bɔlo-t̪o] '(would you please) tell (me)', and one at the edge of the entire

sentence, ending in [aʃbe] ‘will come’. Both tones are realized with little or no visible elbow between the gradual and extreme rises. [Ba31]

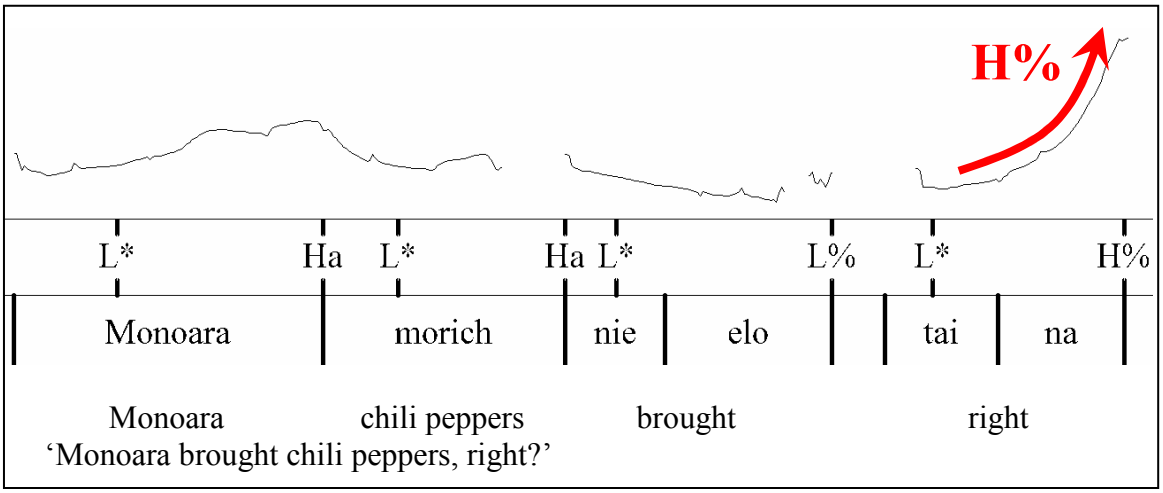


Figure 63. The tag question [t̪aj na] ‘right?’ (*lit.* ‘is [it] not just that?’) bears a high IP boundary tone (H%), realized with a slight elbow between the gradual rise and extreme final rise. [Tu18]

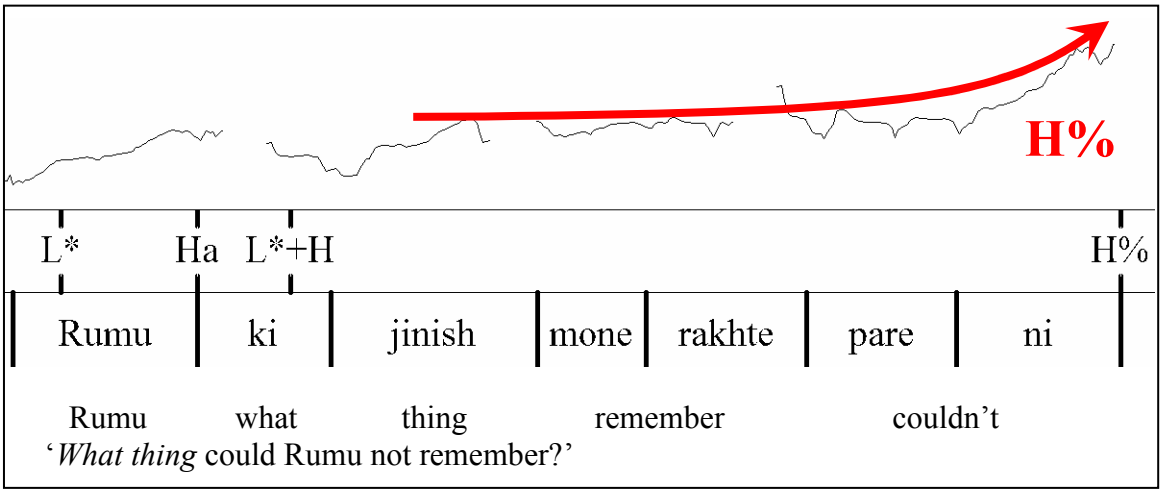


Figure 64. This echo wh-question bears a high IP boundary tone (H%). The lack of AP-level tones following the focused word [ki dʑiniʃ] ‘what thing’ creates a long plateau of high pitch between the rising pitch accent (L*+H) and the high IP boundary tone (H%). See Chapter Four for a discussion of post-focal tone deletion. [To41]

Furthermore, the first member of a set of conjoined sentences such as [kukuɹtaɔ̯ d̪ɛkʰe naj, tɕʰeletaɔ̯ d̪ɛkʰe naj] ‘Neither the dog nor the boy had seen it’ (literally ‘the dog also hadn’t seen it, the boy also hadn’t seen it’) bears a high IP boundary tone (H%), while the second member bears the otherwise expected IP boundary tone—in this case the low IP boundary tone (L%) for declarative sentences—as in Figure 65 below.

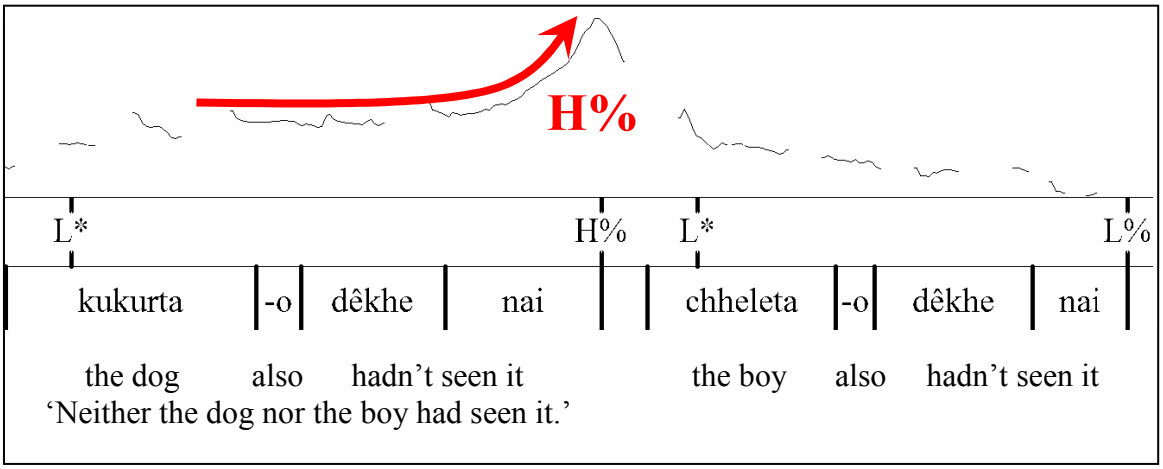


Figure 65. The first member of this set of coordinated sentences bears a high IP boundary tone (H%). This sentence was produced in a hybrid of Eastern Bengali and Bangladeshi Standard Bengali. [FoS50]

Rising IP boundary tone (LH%)

The rising IP boundary tone (LH%) marks the right edge of default wh-questions, as shown in Figure 66 below. It is realized as gradually falling pitch from the IP-final pitch accent, followed by a sudden upward turn on the final syllable, ending mid or high, as in Figure 66 below.

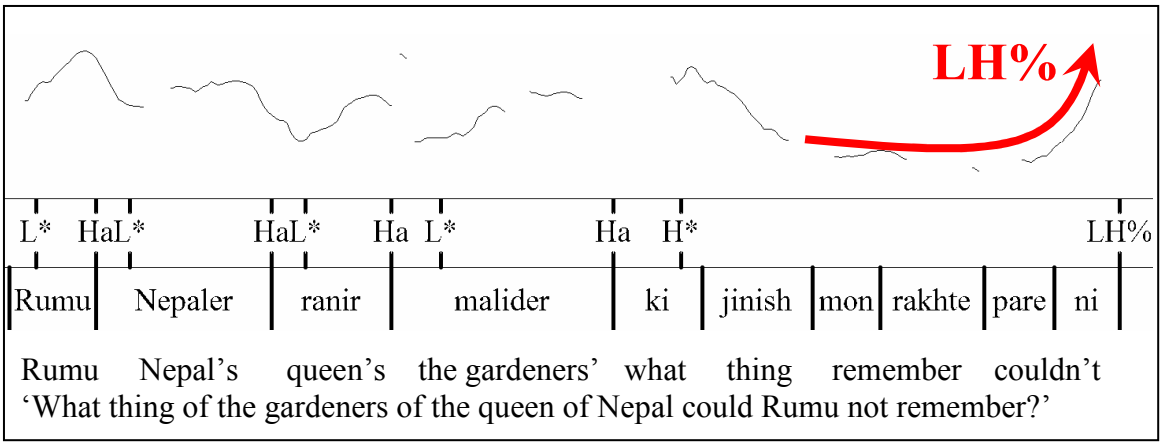


Figure 66. This default wh-question is marked with a rising IP boundary tone (LH%) The lack of AP-level tones following the high pitch accent (H*) in this example clearly reveals the L component of the contour boundary tone. [SB47]

When the rising IP boundary tone (LH%) occurs in close proximity to the preceding pitch accent, it can be difficult to see the L target before the pitch rise on the IP-final syllable. However, when the IP-final pitch accent is separated from the IP's right boundary by several syllables, as in Figure 67 and Figure 68 below, the lowering of pitch to achieve the L target preceding the IP-final syllable is more clearly identifiable. The rising IP boundary tone (LH%) in Figure 67 in particular serves as a good contrast with the high IP boundary tone (H%) in Figure 64 above. Note how the rising IP boundary tone (LH%) involves gradually falling pitch before the final rise, while the high IP boundary tone (H%) involves gradually rising pitch before the final rise.

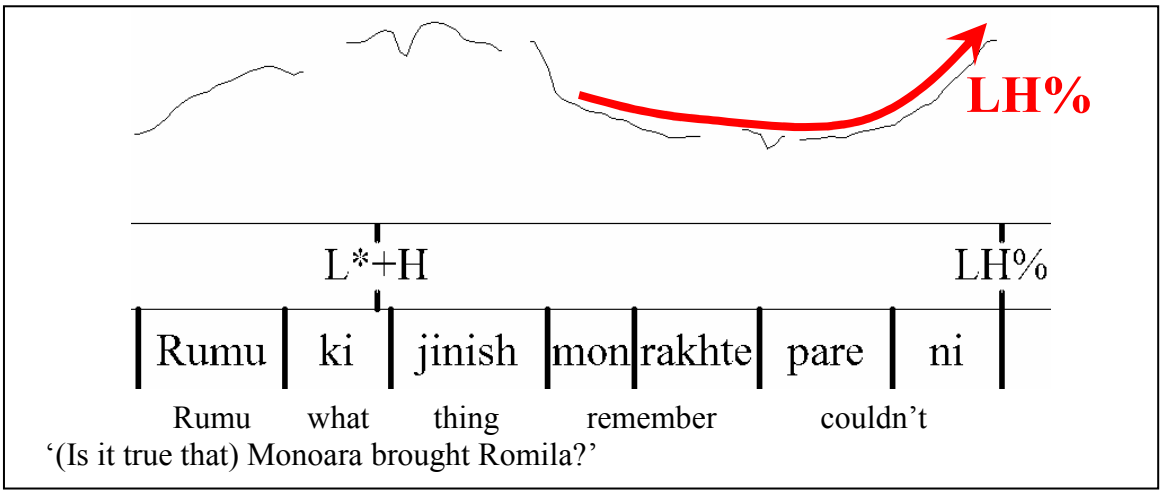


Figure 67. This default wh-question is marked with a rising IP boundary tone (LH%) The lack of AP-level tones following the rising pitch accent (L*+H) clearly reveals the L component of the boundary tone, contrasted with the same sentence produced with a high IP boundary tone (H%) in Figure 64. [Re38]

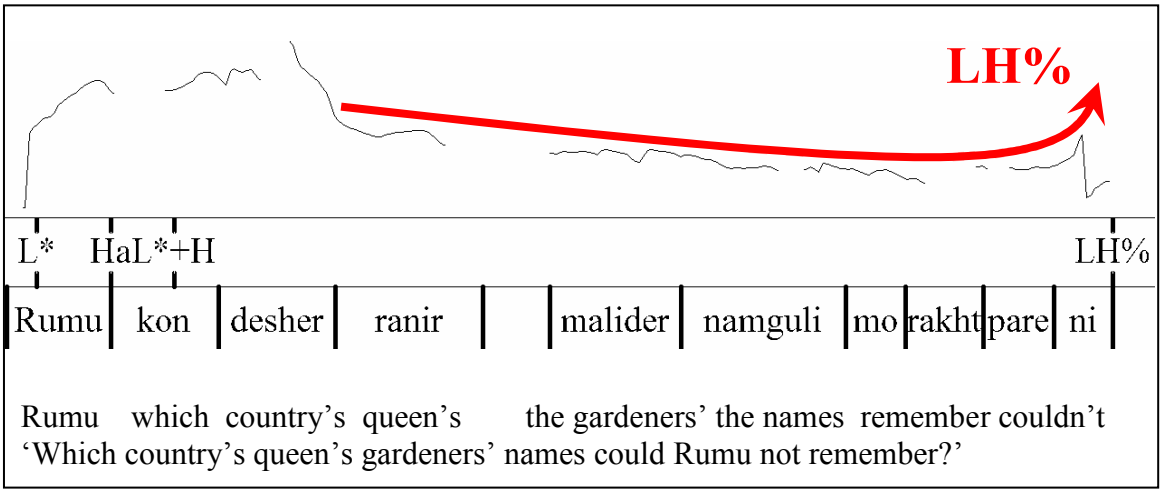


Figure 68. This default wh-question is marked with a rising IP boundary tone (LH%) The lack of AP-level tones following the rising pitch accent (L*+H) in this example clearly reveals the L component of this contour boundary tone. The sharp change in the pitch track during the final syllable [ni] is due to pitch halving, where the tracking software reduces the F_0 measurement by 50%. [SB38]

Even in wh-questions, the rising IP boundary tone (LH%) is less frequent than the low IP boundary tone (L%) in the corpus of data collected for the current study. This is presumably due to the types of wh-questions elicited; most of the wh-questions read by the subjects involved asking for one particular part of a large noun phrase (NP). Given that the asker of the question already knows all the rest of the NP, and is only asking for the one missing piece of information, it would be likely that the asker would produce either the IP boundary tone associated with echo wh-questions (H%) or that of extra-focus wh-questions (L%), with focus realization on the missing information. The rising IP boundary tone (LH%) would be expected in wh-questions produced in an "out of the blue" context, under broad focus.

Falling IP boundary tone (HL%)

The falling IP boundary tone (HL%) is almost exclusively used for yes/no questions. This tone is realized as gradually rising pitch from the IP-final pitch accent to a very high pitch range, followed by a sharp drop during the final syllable, typically falling to mid or low, as shown in Figure 69 below.

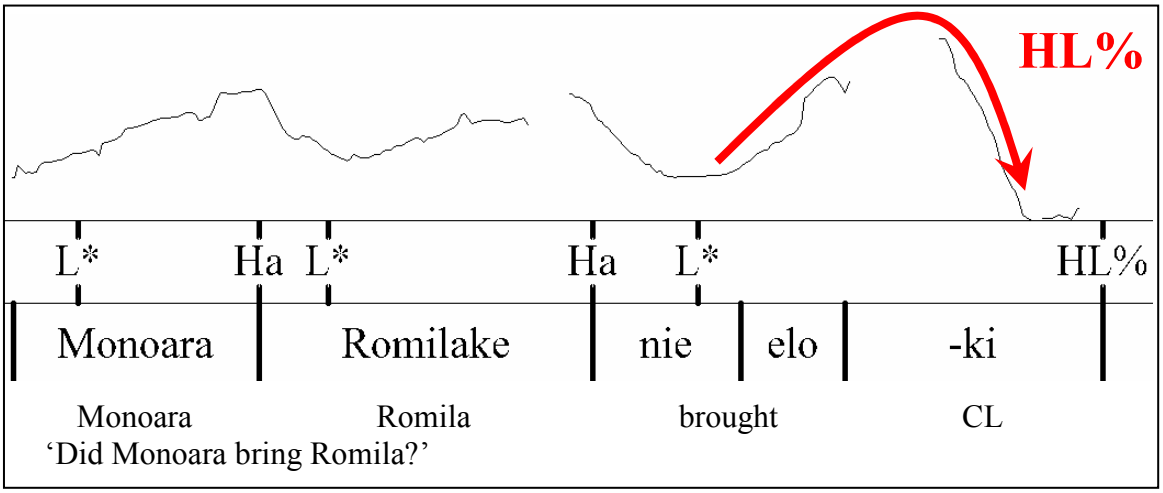


Figure 69. This yes/no-question bears a falling IP boundary tone (HL%). When sentence-initial or -final, the presence of the enclitic -[ki] can indicate yes/no questions. [Fa04]

The L portion of the falling IP boundary tone (HL%) can even be realized in conjunction with creaky phonation, as in Figure 70 below.

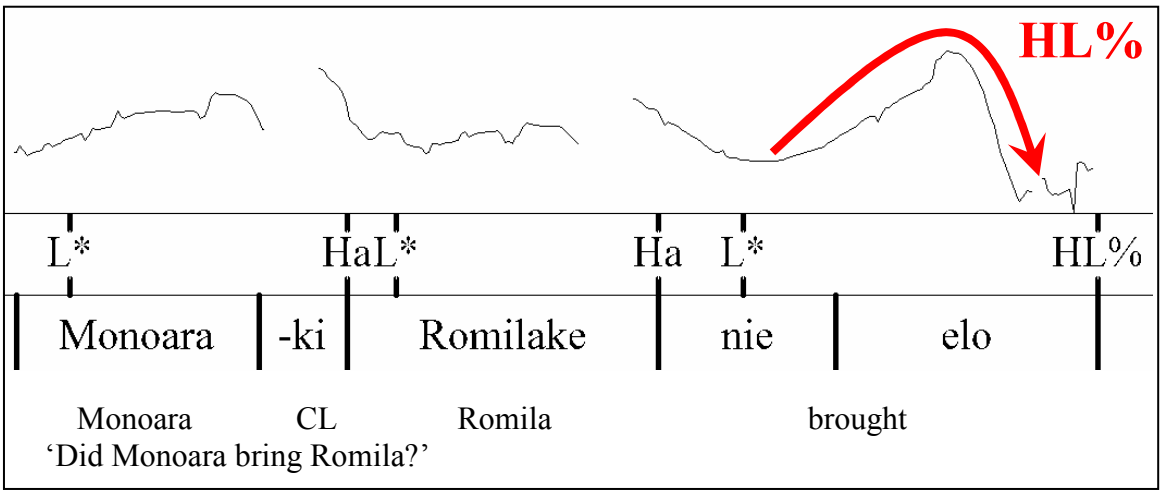


Figure 70. This yes/no-question bears a falling IP boundary tone (HL%). When sentence-initial or -final, the presence of the enclitic -[ki] can indicate yes/no questions. The pitch track becomes choppy at the end of the syllable [lo] due to creaky phonation. [Fa03]

In a few isolated cases, the falling IP boundary tone (HL%) is realized with a sharp rise and fall in the IP-final syllable, as in Figure 71 below, instead of the typical pattern, which involves a gradual rise to the final syllable followed by a sharp drop (*cf.* Figure 69 and Figure 70). In this way, it resembles the falling ip boundary tone (HL-), in that it seems to obey the ip boundary tone locality constraint (*cf.* Figure 52 and Figure 53). In these cases, other features can be observed to distinguish the two boundary tones, such as relative height and upstep of the previous pitch accent (see §10.1.3). See §10.1.4 for a direct comparison of the falling ip (HL-) and IP (HL%) boundary tones.

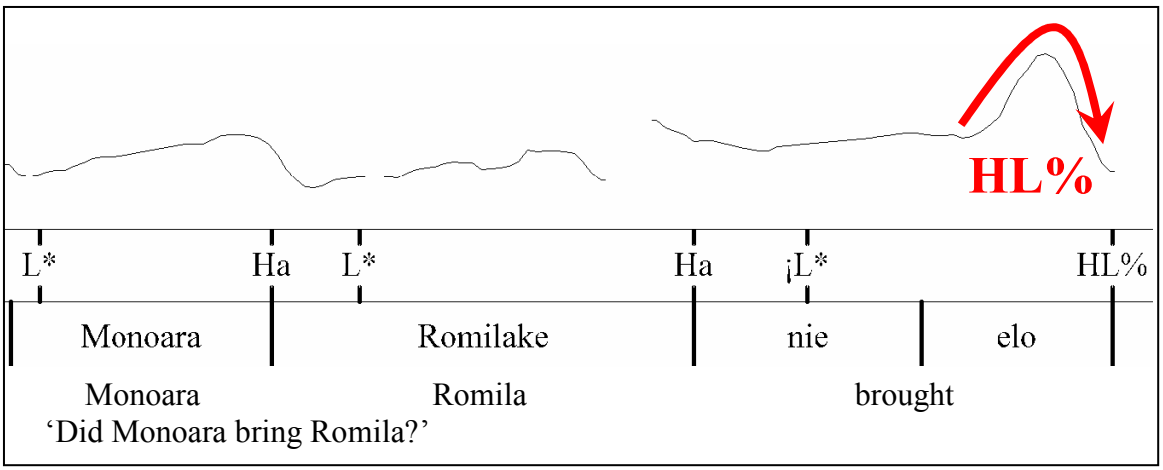


Figure 71. In this example of the yes/no-question, the falling IP boundary tone (HL%) is realized as both a sharp rise and fall in pitch in the IP-final syllable, instead of the more common pattern involving a steady rise in pitch from the IP-final pitch accent followed by a sharp drop in pitch in the IP-final syllable. The upstepped low pitch accent (¡L*) is explained in §10.1.3. [Ba02]

Dipping IP boundary tone (HLH%)

The dipping or falling-rising IP boundary tone (HLH%) marks the right edge of large syntactic clauses that do not occur sentence-finally (*e.g.* relative clauses, because-clauses, if-clauses). In many languages, this boundary tone would correspond to a "continuation rise". With two H tones surrounding an L tone, the dipping IP boundary tone (HLH%) is the only entry in the Bengali tonal inventory to be composed of more than two targets. It is realized phonetically as a rise in pitch from the IP-final pitch accent until the penultimate syllable, and a relatively short fall and rise in pitch within the duration of the final syllable, ending near the middle of the speaker's pitch range, as shown in below in Figure 73. Plausibly due to the quick succession of three tonal targets, the pitch levels reached during this tritonal boundary tone are not as extreme as those seen in the bitonal boundary tones (*cf.* HL%, LH%).

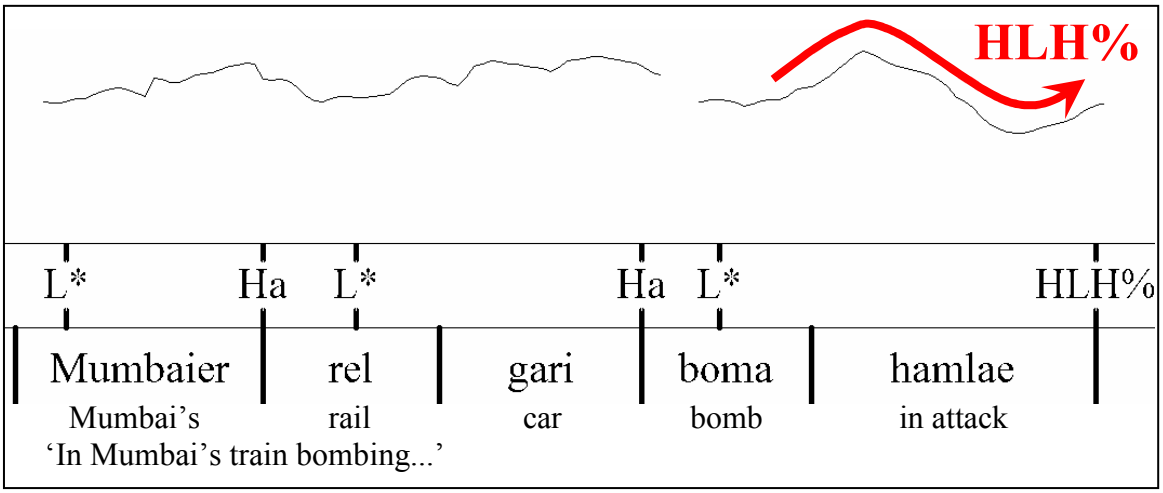


Figure 72. This IP bears a dipping IP boundary tone (HLH%) to indicate that the speaker has not yet completed the full sentence. [Re49]

At first glance, it may be unclear whether the first H target of the dipping IP boundary tone (HLH%) is associated to the IP-final pitch accent or the boundary tone. How can one be certain that the IP-final pitch accent is low (L*) and not, for example, rising (L*+H)? The ambiguity is removed when examining examples such as Figure 73 below, where the IP-final pitch accent is several syllables away from the right boundary of the IP. With the pitch accent and boundary tone so far removed, it is clear that the H target is associated with the boundary tone and not with the pitch accent, as the pitch maximum of the rising pitch accent (L*+H) is typically reached in the second, and in rare cases the third, syllable of the word bearing the pitch accent.

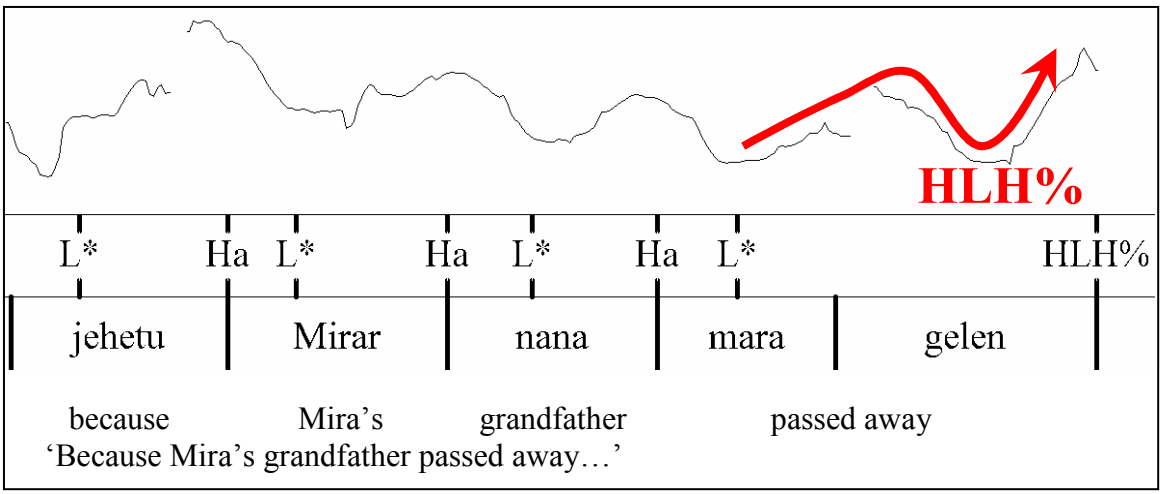


Figure 73. This because-clause is marked on the right edge by a dipping IP boundary tone (HLH%), realized as an F_0 rise after the final pitch accent and a fall and rise on the final syllable. [Fa35]

Additional evidence supporting the claim that the first H target of the dipping IP boundary tone (HLH%) is associated to the boundary tone and not to a rising pitch accent

(L*+H) is relative pitch height. While the sharp rise (L*+H) is normally distinguishable amongst a series of smooth rises (L*…Ha) by an H target that exceeds the pitch of the preceding high AP boundary tone (Ha), the initial H target of the dipping IP boundary tone (HLH%) often does not exceed the pitch of the preceding high AP boundary tone (Ha), as in Figure 74 below.

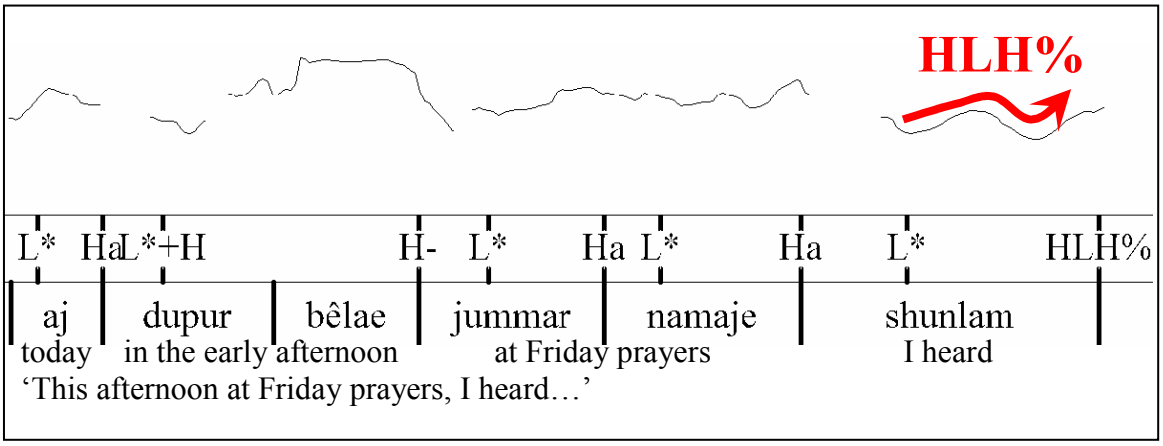


Figure 74. The first H target of the dipping IP boundary tone (HLH%) does not exceed the pitch of the preceding high AP boundary tone (Ha), supporting the claim that it is not associated with the sharp rise tonal pattern (L*+H). [Jo49]

If the IP ends in a function word, as in Figure 75 below, the intermediate L tonal target in the dipping IP boundary tone (HLH%) can be optionally stretched out to the IP-penultimate syllable, pushing the first H tonal target to the antepenultimate syllable. Note how the pitch lowers from the first H during the first syllable of the function word [bole] 'because', instead of during the last syllable as in Figure 73 above (the sentences are the same, but produced by two different speakers).

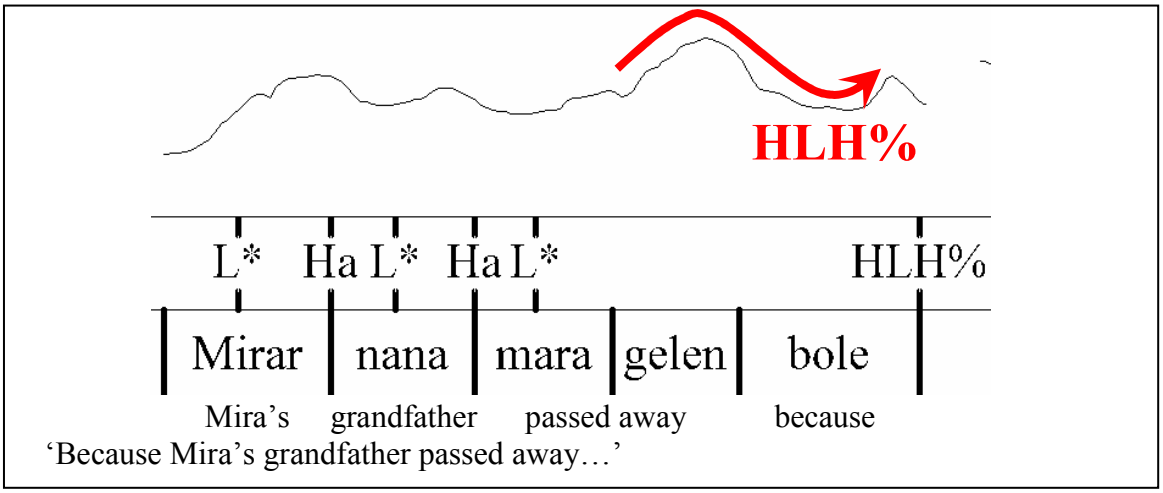


Figure 75. This large non-sentence-final clause is marked on the right edge by a dipping IP boundary tone (HLH%). The L portion of the tone stretches across the initial syllable of the function word [bole] 'because' instead of remaining restricted to the final syllable. [Az34]

Furthermore, the final H target of the dipping IP boundary tone (HLH%) can be optionally undershot in fast, non-emphatic speech—similar to the reduced form of the dipping tone or "third tone" of Mandarin (Huang 2001)—leaving only one H and one L target in the boundary tone, giving HL(H)%. Despite bearing only one H and one L target, this undershot tone is nevertheless distinguishable from the falling IP boundary tone (HL%). As in non-undershot forms, the first H target of the dipping IP boundary tone (HLH%) is realized during the penultimate syllable. The IP-final syllable only bears the L target, instead of the rising contour expected for non-undershot forms of the dipping IP boundary tone (HLH%), as shown in Figure 76 below. In contrast, both the H and L targets of the falling IP boundary tone (HL%) are realized during the final syllable (see

Figure 69, Figure 70, and Figure 71). A direct comparison between the two tones is provided in §10.1.4.

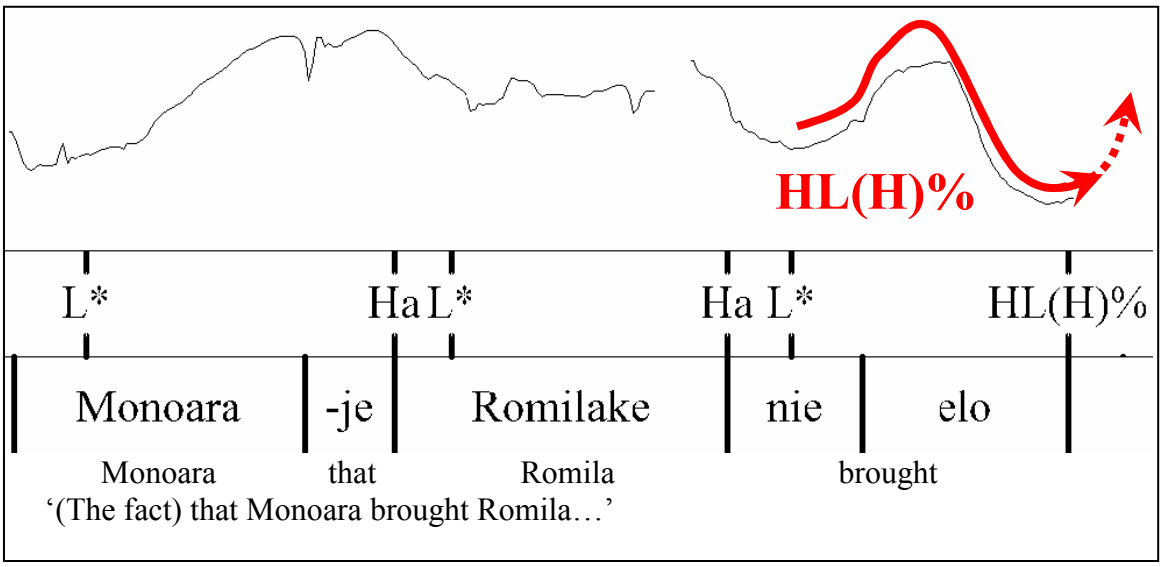


Figure 76. This large non-sentence-final clause is marked on the right edge by a dipping IP boundary tone (HLH%), whose final H target is not reached. [Ro15]

The L target of the undershot dipping IP boundary tone (HL(H)%) can even reach the low pitch range associated with the low IP boundary tone (L%), including the accompanying creaky phonation, as in Figure 77 below. It is nevertheless still distinct from the falling IP boundary tone (HL%) due to the early realization (*i.e.* during the penultimate syllable) of the H target (see §10.1.4).

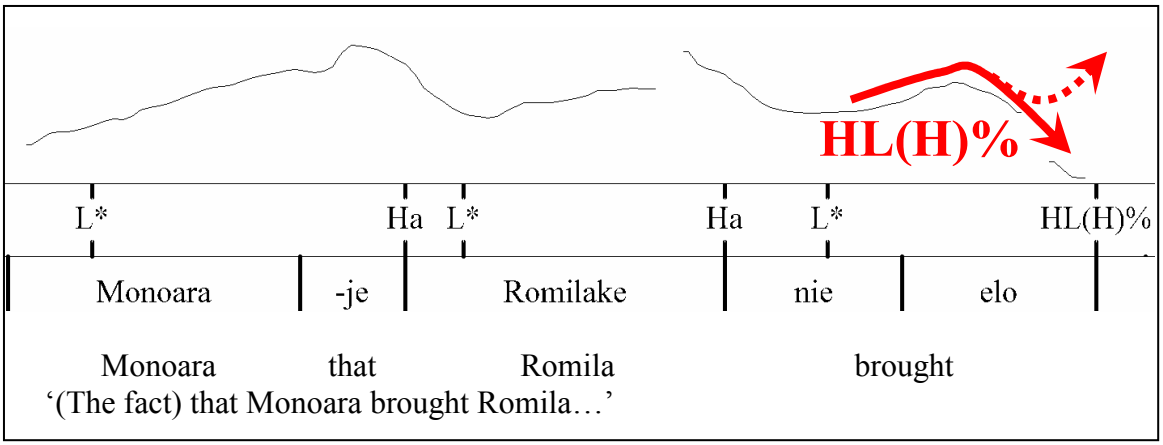


Figure 77. This large non-sentence-final clause is marked on the right edge by a dipping IP boundary tone (HLH%), whose final H target is not reached. [Ro15]

The undershot dipping IP boundary tone can be optionally transcribed either simply as HLH%, or with the final H target in parentheses, giving HL(H)%, to indicate to the reader that the tone is a phonetically reduced version of the dipping IP boundary tone (HLH%) instead of a variant of the falling IP boundary tone (HL%) primarily associated with non-negative yes/no questions.

By controlling word length, sentence length, clause type, and sentence type using punctuation marks and sentence-level clitics, three pitch accents (*i.e.* L*, H*, L*+H), two AP boundary tones (*i.e.* La, Ha), four ip boundary tones (*i.e.* L-, H-, LH-, HL-), and five IP boundary tones (*i.e.* L%, H%, LH%, HL%, HLH%) were found in Bangladeshi Standard Bengali. I now move on to discuss the interaction of these different tone types in §10.1.3.

10.1.3 Tonal interaction

Thus far, I have described the pitch accent and boundary tones of Bangladeshi Standard Bengali as independent tones. Of course, in practice, many pitch accents and boundary tones interact with one another, either because they occur in the same location or because they are found in close proximity to each other. In addition to concurrent tonal overriding, which simply deletes tones associated to smaller prosodic units when cooccurring with tones associated with larger prosodic units, the tonal interactions seen here include optional upstep of low pitch accents (¡L*) preceding H-initial IP boundary tones (*i.e.* H%, HL%, HLH%) and optional downstep of high ip boundary tones (!H-) following the rising pitch accent (L*+H). While both of these phenomena are given labels in the B-ToBI transcription system, it is not entirely clear if their presence or absence is phonologically relevant.

Upstep of low pitch accents (¡L*)

Due to the strong influence of following IP boundary tones, IP-final low pitch accents (L*) can be realized as upstepped (¡L*)[65] when preceding an IP boundary tone beginning with a H component: high (H%), falling (HL%), or dipping (HLH%). Upstepped low pitch accents (¡L*) are characterized by an F_0 valley during the phonologically stressed syllable, lower than the surrounding pitch (local F_0 minimum),

[65] The inverted exclamation mark (¡) is used as a diacritic for upstep in B-ToBI instead of the carat (^) used in GToBI (Grice *et al.* 2005). This is done to distinguish upstep from other types of phonetic raising (*e.g.* ^L*+H), which are marked by the carat (^) in B-ToBI.

but substantially higher than the previous AP's low pitch accent (L*). Figure 78 is an example of the upstepped low pitch accent (¡L*) preceding a high IP boundary tone (H%).

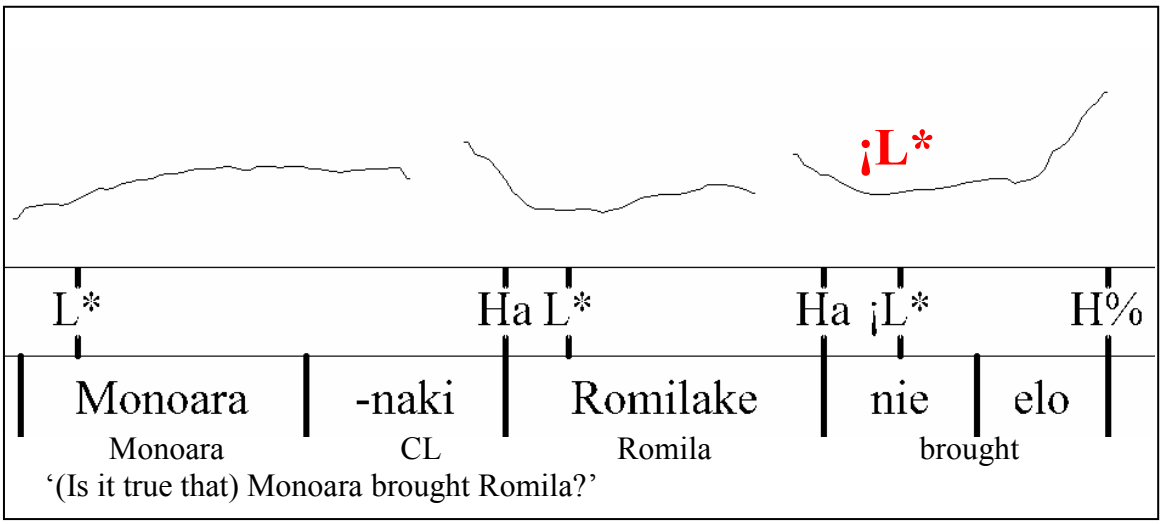


Figure 78. In this negative yes/no question, the low pitch accent (L*) of the IP-final AP [nie elo] 'brought' is upstepped (¡L*) in anticipation of the high IP boundary tone (H%). [Ba07]

The falling IP boundary tone (HL%), associated primarily with non-negative yes/no questions, can cause upstep on the preceding low pitch accent (L*), as in Figure 79 below.

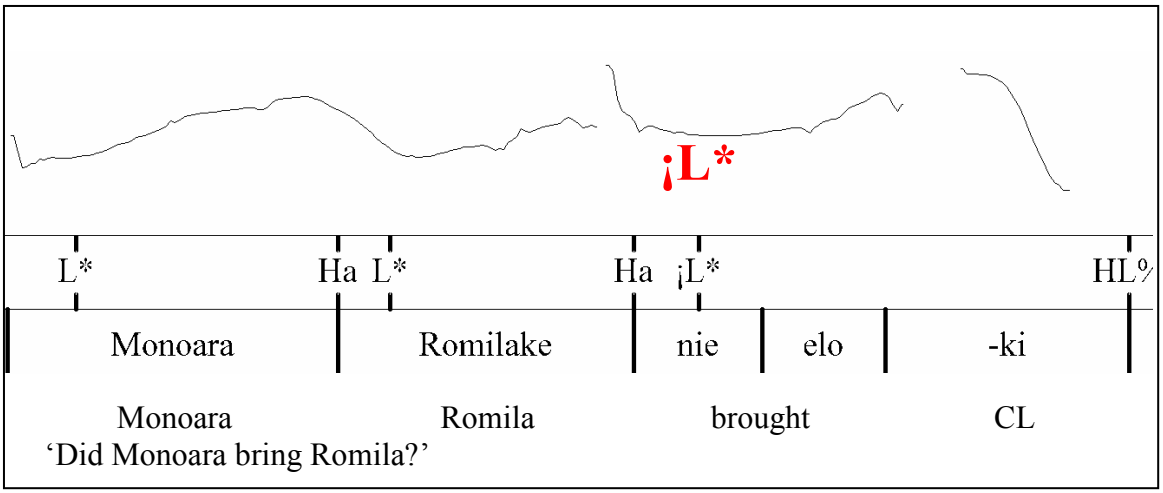


Figure 79. The low pitch accent (L*) of the IP-final AP [nie elo-ki] 'brought-CL' is upstepped (¡L*) in anticipation of the falling IP boundary tone (HL%). [Na04]

Upstep is not always restricted to the IP-final AP; in a few cases, the IP-penultimate AP can also bear an upstepped low pitch accent (¡L*), as in Figure 80 below.

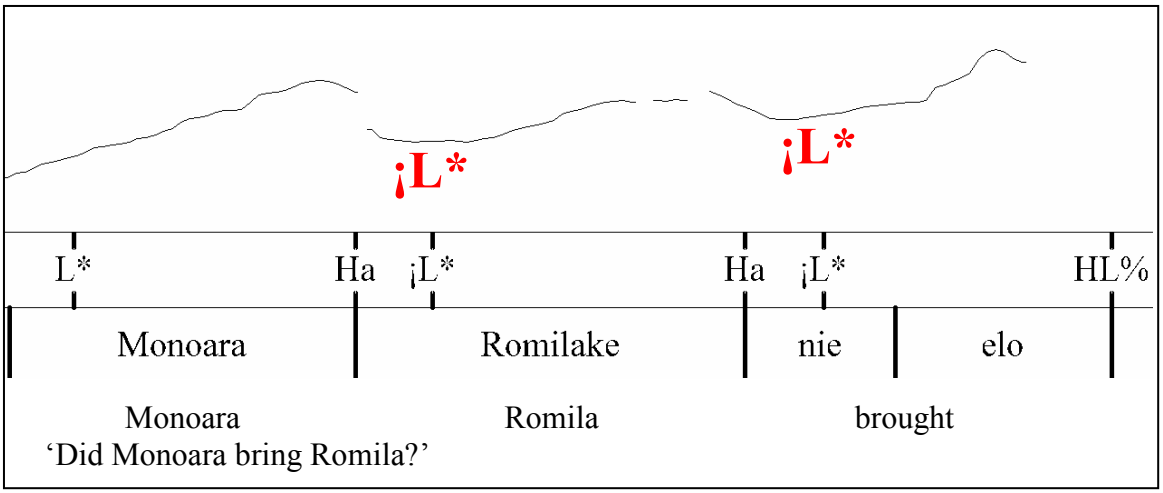


Figure 80. The low pitch accents of both the IP-final AP [nie elo] 'brought' and the IP-penultimate AP [ɹomilake] 'Romila-ACC' are upstepped (¡L*) in anticipation of the falling IP boundary tone (HL%). The pitch track during the L portion of the falling IP boundary tone (HL%) is broken due to creaky phonation. [Sh02]

As the tonal targets for the dipping IP boundary tone (HLH%) tend to be less extreme than other boundary tones, the upstep of a low pitch accent (L*) preceding it is also less extreme than other upstepped low pitch accents, as in Figure 81 below.

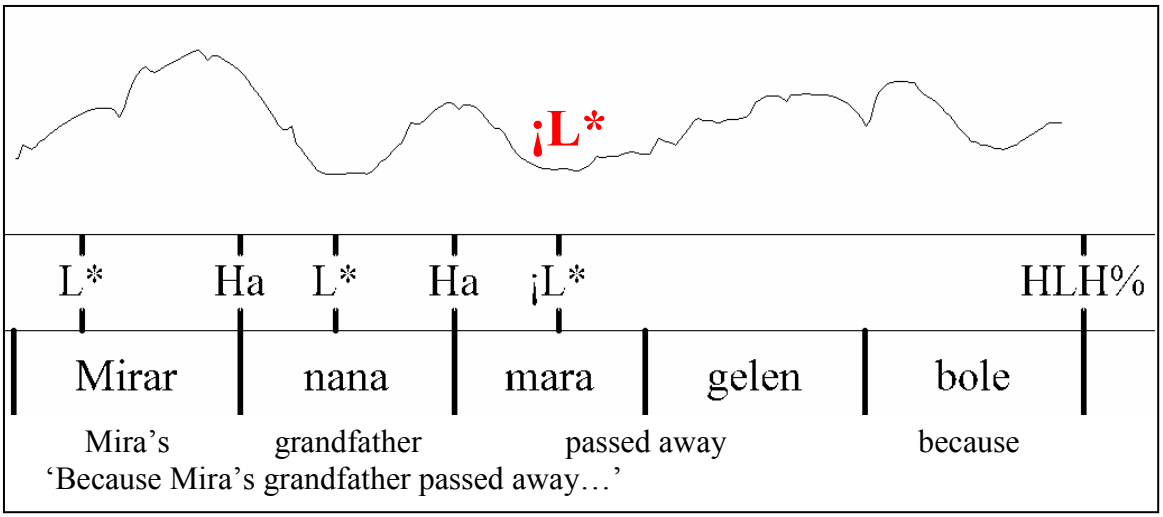


Figure 81. The low pitch accent (L*) borne on the IP-final AP [maɹa gelen bole] 'because…died' is slightly upstepped in anticipation of the dipping IP boundary tone (HLH%). [Fa34]

Unlike the downstep of the high pitch accent (H*) and Ha downtrend (§10.1.1), the upstep of the low pitch accent (¡L*) is not regular. It is not uncommon to see low pitch accents (L*) following regular downtrend even in contexts where upstep would be legal. Due to this unpredictability, upstepped low pitch accents (¡L*) should be labeled with the inverted exclamation mark (¡), unlike high pitch accents (H*) and high AP boundary tones (Ha), which always follow downtrend and thus need no special labeling. However, as I have not found a pattern as to when a low pitch accent preceding an H-initial IP boundary tone (H%, HL%, HLH%) will be upstepped, it is unclear whether there is a phonological distinction between upstepped and non-upstepped low pitch accents (*cf.* violation of H* and Ha downtrend as a marker of focus).

Downstep of high ip boundary tones (!H-)

The high ip boundary tone (H-), normally realized as a sharp rise in pitch on the ip-final syllable, is often downstepped (!H-) following a rising pitch accent (L*+H) in non-focused constituents, *i.e.* in very small ips. The downstepped high ip boundary tone (!H-) is realized as a (high-)mid plateau or slight fall in pitch on the ip-final syllable, just above the typical range for high AP boundary tones (Ha). The downstepped high ip boundary tone (!H-) can be distinguished from non-downstepped versions in two ways, as evident in Figure 82: (1) its pitch maximum can reach, but does not exceed, that of the preceding trailing high portion (+H) of the rising pitch accent (L*+H), and (2) there is no sharp rise in pitch in the ip-final syllable.

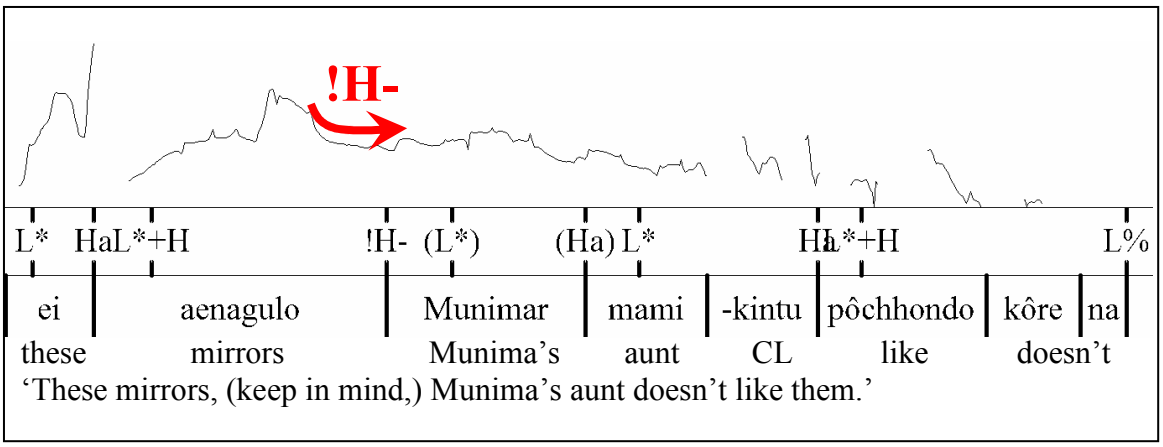


Figure 82. The high ip boundary tone (H-) borne at the right edge of the topicalized object [ej aʔnagulo] 'these mirrors' is downstepped due to the preceding rising pitch accent (L*+H). The boundary tone is realized as a mid plateau. Note that the pitch excursions following the demonstrative [ej] 'these' are due to the insertion of a glottal stop before the vowel-initial [aʔnagulo] 'the mirrors'. It is not clear why the smooth rise (L*…Ha) on [munimaɹ] 'Munima's' is phonetically undershot. [Do55]

In addition to a mid or high-mid plateau, the downstepped high ip boundary tone (!H-) can be realized with slightly falling pitch, also in the (high-)mid region of the speaker's pitch range, as illustrated in Figure 83 and Figure 84 below.

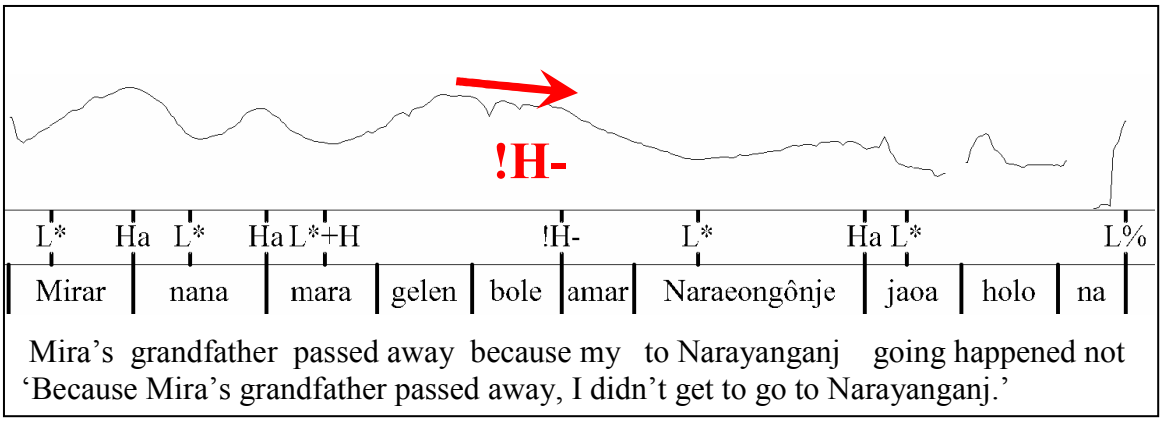


Figure 83. The high ip boundary tone (H-) borne at the right edge of the because clause [miɹar nana maɹa gelen bole] 'because Mira's grandfather passed away' is downstepped due to the preceding rising pitch accent (L*+H). The boundary tone is realized as smoothly falling high-mid pitch. Note that the irregular pitch track during the final AP [dʑaa̰a holo na] 'going didn't happen' is due to creaky phonation. [Na34]

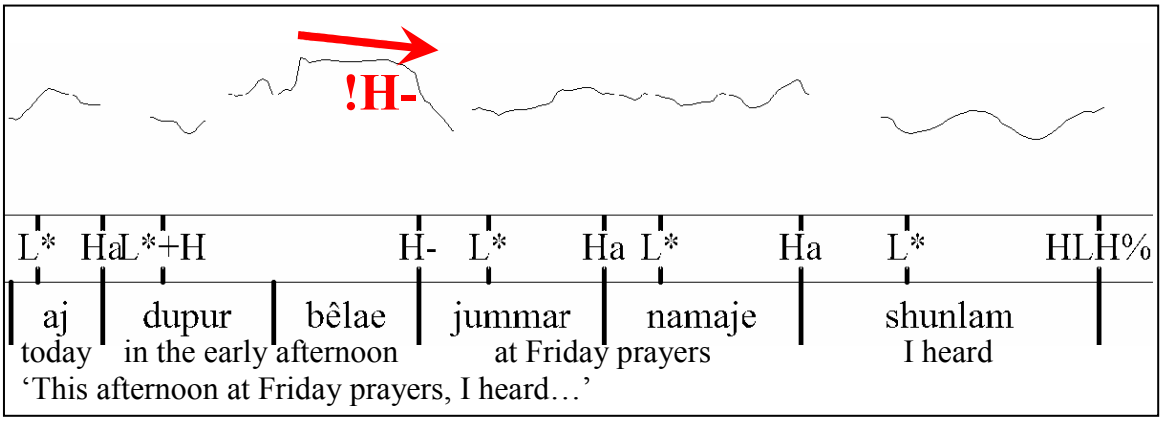


Figure 84. The high ip boundary tone (H-) borne at the right edge of the adverbials [adʑ d̪upuɹ bɛlaʂ̩] 'this afternoon' is downstepped due to the preceding rising pitch accent (L*+H). The boundary tone is realized as smoothly falling high-mid pitch. [Jo49]

The sentence in Figure 84 above was produced by another speaker with two downstepped high ip boundary tones (!H-), illustrated below in Figure 85.

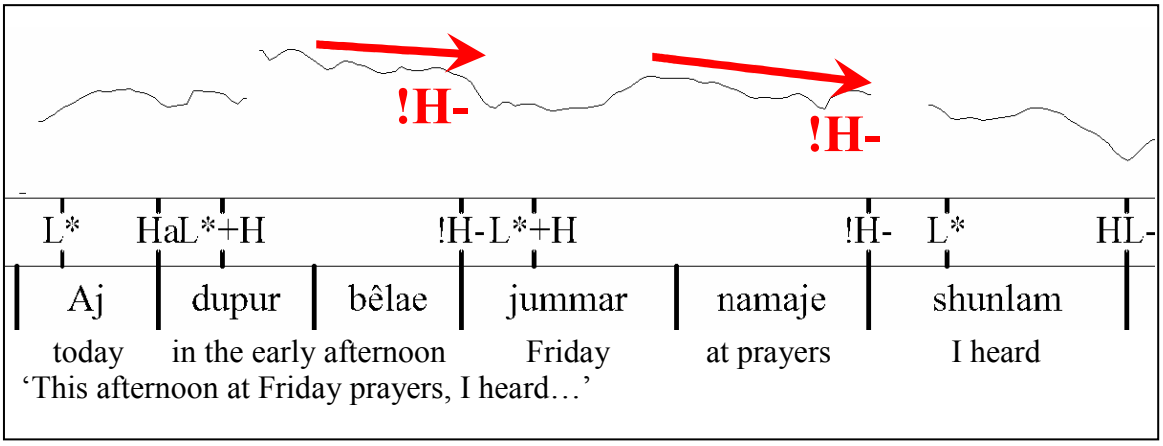


Figure 85. The high ip boundary tones (H-) borne at the right edge of the adverbials [adʑ d̪upuɹ bɛlaʂ̩] 'this afternoon' and [dʑumːaɹ namadʑe] 'at Friday prayers' are downstepped due to the preceding rising pitch accents (L*+H). The boundary tones are realized as smoothly falling high-mid pitch. [Ro49]

Downstep of the high ip boundary tone (H-) is not entirely predictable by examining the surrounding tones. Although downstep can only occur following a rising pitch accent (L*+H), there are examples of non-downstepped high ip boundary tones (H-) following the rising pitch accent (L*+H) in focused phrases. Examples of these are given in Chapter Four. Because of this extra complication in distribution, downstep is considered a component of the Bangladeshi Standard Bengali prosodic structure that should be transcribed in the B-ToBI system.

Because of the fact that both downstepped and non-downstepped versions of the high ip boundary tone (H-) can be found following the rising pitch accent (L*+H), one may consider the two to be contrastive units, as minimal pairs can be found when differences in focus are ignored. Indeed, there is no undefeatable evidence that the downstepped high ip boundary tone (!H-) must be considered a variant of the high ip boundary tone (H-); it could just as well be considered a variant of another ip boundary tone, or it could be an entirely separate ip boundary tone altogether. If it is a variant of an ip boundary tone, it could be considered an upstepped low ip boundary tone (¡L-), or a flattened rising or falling ip boundary tone (either ¡LH-, H¡L-, L!H-, or !HL-). This hypothesis is undesirable only in that it would not explain the high frequency of what is considered here the downstepped high ip boundary tone (!H-) with respect to the very low frequencies of the non-high ip boundary tones (*i.e.* L-, HL-, LH-). If this is an entirely separate tone, it could be labeled a mid ip boundary tone (M-) to represent its mid-level,

largely flat pitch. This hypothesis would run into theoretical questions, including whether it is valid to posit a third tonal target beyond the usual H and L.[66] As the label of this particular tone is not crucial to the model, I tentatively label it a downstepped high ip boundary tone (!H-) and leave the question of its "true identity" open for further study.

10.1.4 Boundary tone comparison

With so many boundary tones differentiated solely on their association to different prosodic units (*e.g.* Ha *vs.* H- *vs.* H%), it is essential that we find independent acoustic distinctions between tone pairs and triplets, before we can claim that they are phonologically distinct for Bangladeshi Standard Bengali speakers. Based on properties of the pitch contour, I note clear distinctions between high boundary tones (*i.e.* Ha *vs.* H- *vs.* H%). I also tentatively point out differences between the pitch contours of rising boundary tones (*i.e.* LH- *vs.* LH%), and those of falling boundary tones (*i.e.* HL- *vs.* HL% *vs.* HL(H)%), while noting that additional data will be required for confirmation of these tentative comparisons. Comparisons of low boundary tones (*i.e.* La, L-, L%) could not be made at this time, due to the lack of minimal pairs or triplets of the tones.

Comparing H boundary tones (Ha, H-, H%)

The current model proposes three types of high boundary tones in the language (Ha, H-, H%), corresponding to the three layers of tonally-marked phrases, instead of the two high boundary tones proposed in previous models. If there is a true three-way distinction

[66] See Beckman *et al.* (2002) for a description of Sp-ToBI, a model of Spanish prosody that includes a mid IP boundary tone (M%) in addition to high (H%) and low (L%) IP boundary tones. In lexical tone languages, proposals involving multiple tonal target levels are common.

among these three H boundary tones, it is predicted that they are distinguishable by native speakers using some acoustic means. I show that the three H boundary tones are acoustically distinguishable in terms of relative pitch (as well as final lengthening; see §10.2).

The high ip boundary tone (H-) reaches a higher pitch than the high AP boundary tone (Ha), as illustrated in Figure 86. By comparing the differences in pitch between the F_0 minimum corresponding to the low pitch accent (L*) and the F_0 maximum corresponding to the high boundary tone (Ha or H-) of identical words when AP-final and ip-final (measured within speaker),[67] it was found that the pitch of the high ip boundary tone (H-) is higher than that of the high AP boundary tone (Ha) [paired $t(5) = 10.90$, $p < 0.05$]. Depending on the speaker, the word measured was either a proper name subject (*i.e.* [monoaɹa] 'Monoara') produced in sentence-initial position, or one of two proper name objects (*i.e.* [ɹomilake] 'Romila-ACC', or [ninake] 'Nina-ACC') produced in sentence-medial position.

[67] It was possible to differentiate AP-final and ip-final examples by looking for signs of ip boundaries, including final syllable lengthening (see §10.2) and the pitch elbow associated with the ip boundary tone constraint.

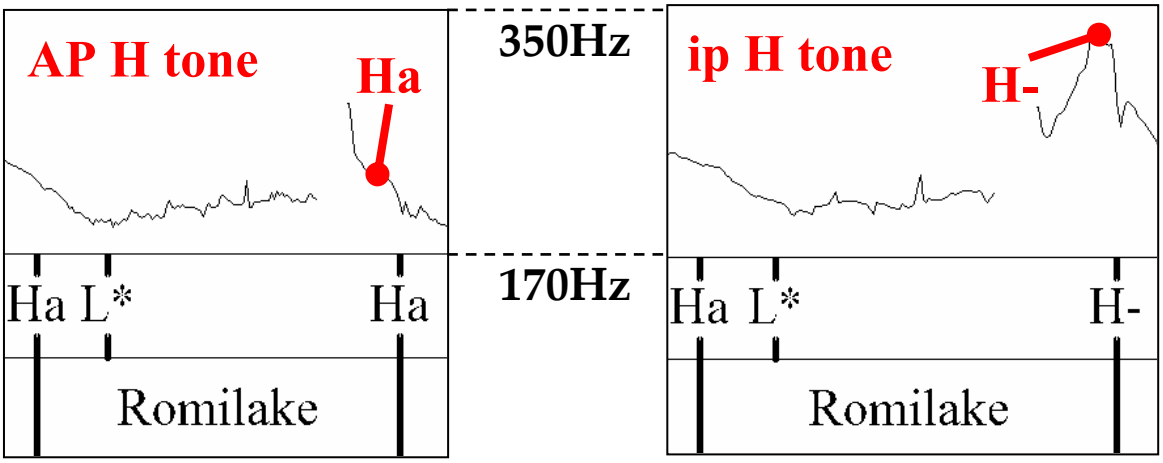


Figure 86. High boundary tones corresponding to the AP (Ha) and ip (H-) levels of phrasing, produced on identical words in syntactically-identical position. [BM02, BM01]

Similarly, the high IP boundary tone (H%) reaches a higher pitch than the high ip boundary tone (H-), as illustrated in Figure 87. By comparing the pitch difference between the F_0 minimum of the low pitch accent (L*) and the F_0 maximum of the high boundary tone (H- or H%) in structurally-equivalent[68] words when ip-final and IP-final, it was found that the pitch of a high IP boundary tone (H%) rises more than that of the high ip boundary tone (H-) [paired $t(5) = 3.59$, $p < 0.05$]. The ip examples from the AP-ip comparison were measured against the IP-final verb [nie elo-na] 'did not bring?', which is produced as four syllables [ni.e.lo.na], in interrogative sentences triggering the high IP boundary tone (H%).

[68] Unlike the measurements made for Ha *vs.* H-, identical words could not be compared for H- *vs.* H%, as ip boundaries and IP boundaries do not occur in syntactically identical positions. AP boundaries and ip boundaries, however, show more variation and often occur in syntactically identical positions.

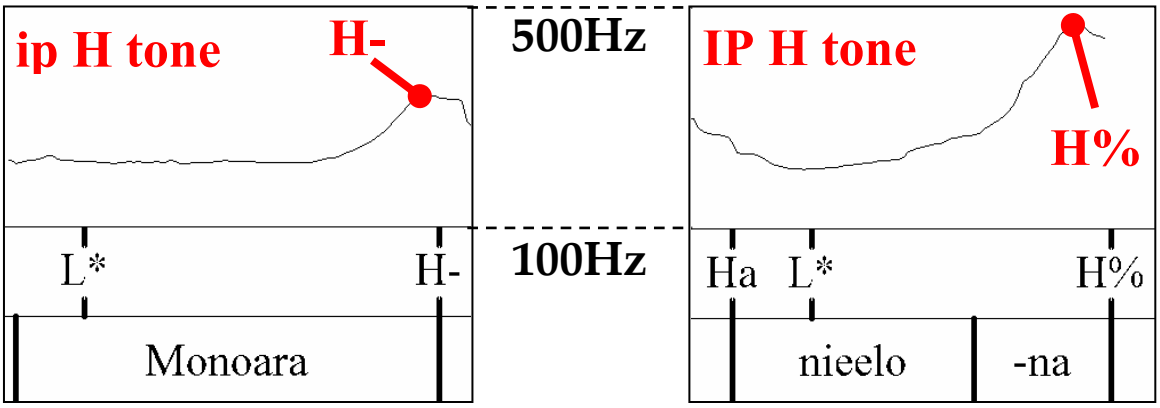


Figure 87. Comparison of the high boundary tones corresponding to the ip (H-) and IP (H%) levels of phrasing, produced on structurally-equivalent words. [Do01, Do06]

The numerical values corresponding to the difference in pitch between low pitch accents (L*) and the three high boundary tones (Ha, H-, H%) are given in Table 18 below.

Speaker	L*…Ha	L*…H-	L*…H%
BM	[ɹomilake] 38Hz	[ɹomilake] 131Hz	[nie elo-na] 223Hz
Sh	[monoaɹa] 60Hz	[monoaɹa] 119Hz	[nie elo][69] 182Hz
Do	[monoaɹa] 60Hz	[monoaɹa] 109Hz	[nie elo-na] 246Hz
Ba	[monoaɹa] 8Hz	[monoaɹa] 75Hz	[nie elo-na] 123Hz
Jh	[monoaɹa] 76Hz	[monoaɹa] 158Hz	[nie elo-na] 328Hz
Pi	[ninake] 22Hz	[ninake] 98Hz	[nie elo-na] 108Hz

Table 18. Pitch differences between low pitch accents (L*) and following high boundary tones of different phrase levels (Ha, H-, H%), during identical (Ha *vs.* H-) or similar (H- *vs.* H%) words.

[69] This speaker's production of the example ending in [nie elo-na] was not used as the IP-final low pitch accent (L*) was upstepped (¡L*), thus reducing the pitch difference between the two targets.

Comparing LH boundary tones (LH-, LH%)

In order to elicit minimal pairs of the rising ip boundary tone (LH-) and the rising IP boundary tone (LH%), the string [mone ɹakʰt̪e paɹe ni] ‘couldn’t remember’ was placed both at the end of a relative clause [ɹumu dʑej namgulo mone ɹakʰt̪e paɹe ni] ‘the names that Rumu couldn’t remember’ and at the end of various wh-questions of the basic shape [ɹumu (wh) mone ɹakʰt̪e paɹe ni?] ‘(wh) could Rumu not remember?’, where “(wh)” represents wh-phrases such as [kaɹ namgulo] ‘whose names’, [nepaleɹ ki dʑiniʃ] ‘what thing of Nepal’, *etc*. Relative clauses typically constitute their own ip, and can thus bear one of the four ip boundary tones (*i.e.* L-, H-, LH-, HL-). Wh-questions can bear one of three IP boundary tones (*i.e.* L%, H%, LH%), depending on the situation (*i.e.* extra focus, echo question, default situation). Given that both sentence types had the option of bearing a rising tone, it was expected that at least some speakers would in fact produce rising tones on both sentence types. Indeed, 11 examples of relative clauses bearing the rising ip boundary tone (LH-) and 30 examples of wh-questions bearing the rising IP boundary tone (LH%) were collected, distributed across 16 speakers. Unfortunately, most of the speakers produced only one member of the intended pair—either the rising ip boundary tone (LH-) or the rising IP boundary tone (LH%), but not both—and produced other boundary tones (*e.g.* L%, H-) on the other member of the intended pair. Thus, pairwise comparisons could not be made across a sufficient number of speakers to reach statistical significance for any acoustic measurements.

Only two speakers produced minimal pairs of the rising boundary tones (*i.e.* LH-, LH%); their recordings are provided below for a brief comparison of what may be distinguishing features of the two rising tones. Compare the first speaker's rising IP boundary tone (LH%) in Figure 88 to her rising ip boundary tone (LH-) in Figure 89. Notice that the rise from the L target to the H target of the IP boundary tone (106Hz) is much greater than that of the ip boundary tone (14Hz), even though both rises take place across the single syllable [ni].

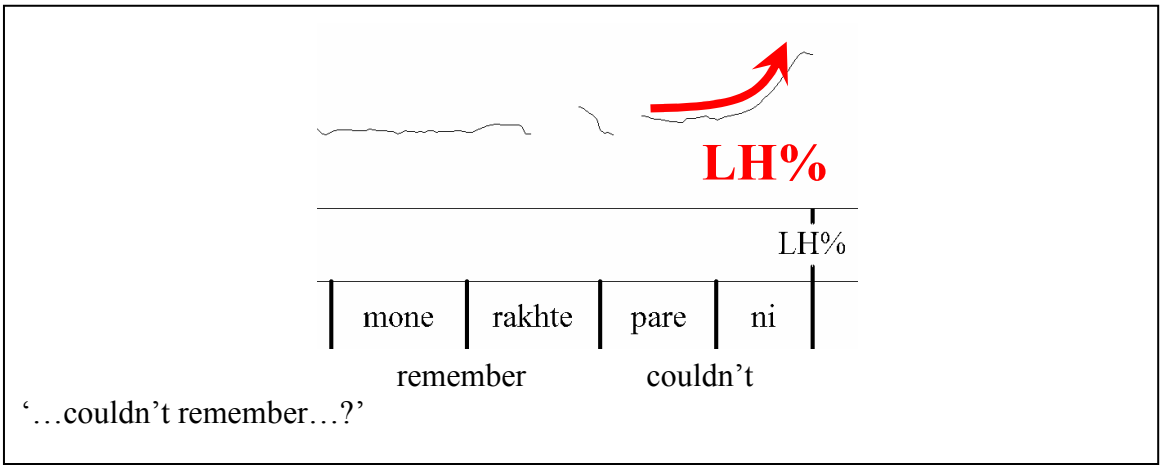


Figure 88. The pitch during the rising IP boundary tone (LH%) borne at the right edge of this wh-question rises from 239Hz to 345Hz within the final syllable [ni]. [Do43]

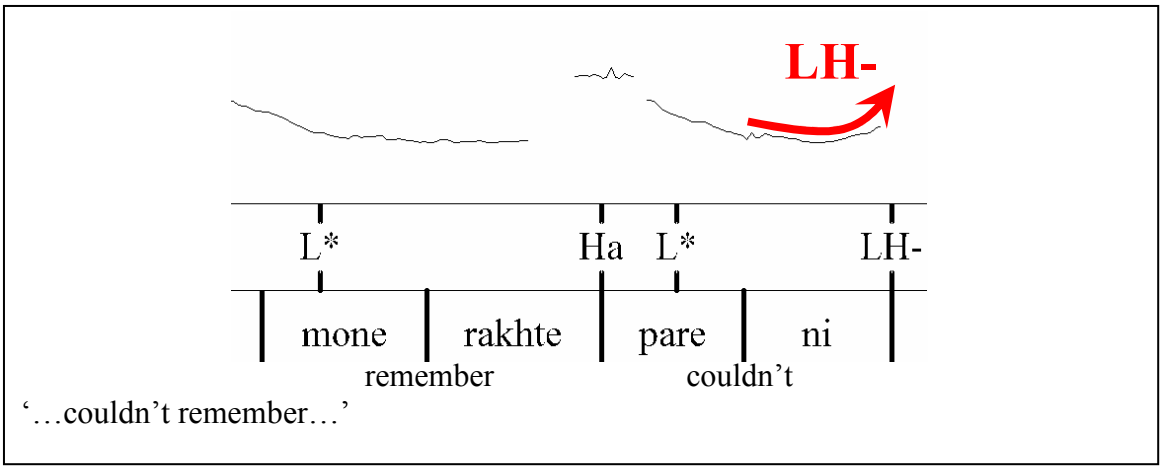
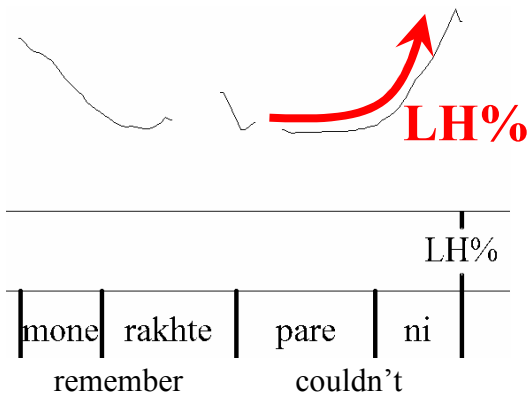


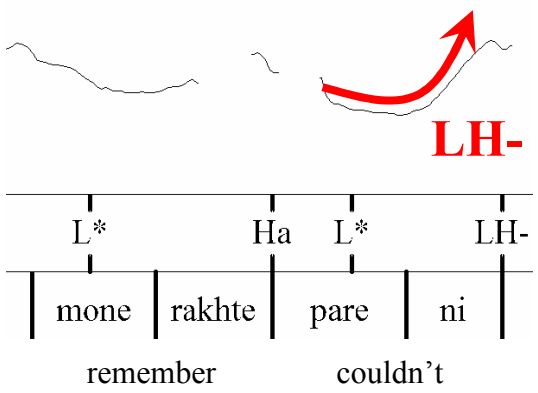
Figure 89. The pitch during the rising ip boundary tone (LH-) borne at the right edge of this relative clause rises from 180Hz to 194Hz within the final syllable [ni]. [Do52]

The second speaker also differentiates the extreme rise of the IP boundary tone (LH%) and the more shallow rise of the ip boundary tone (LH-), although she does not make as large a difference as the first speaker; the rise from the L target to the H target in the second speaker's IP boundary tone (LH%) is 177Hz (Figure 90), while that of the ip boundary tone (LH-) is 105Hz (Figure 91).



'…couldn't remember…?'

Figure 90. The pitch during the rising IP boundary tone (LH%) borne at the right edge of this wh-question rises from 239Hz to 345Hz within the final syllable [ni]. [Jo44]



'…couldn't remember…'

Figure 91. The pitch during the rising ip boundary tone (LH-) borne at the right edge of this relative clause rises from 213Hz to 323Hz within the final syllable [ni]. [Jo52]

Unfortunately, due to a lack of sufficient pairwise comparisons of rising ip and IP boundary tones (LH-, LH%), it is not clear whether the differences in relative pitch height seen in the two speakers' examples presented above are representative of the general population.

Comparing HL boundary tones (HL-, HL%, HL(H)%)

The falling ip boundary tone (HL-), falling IP boundary tone (HL%), and undershot dipping IP boundary tone (HL(H)%) all share a similar tonal makeup: an H target followed by an L target. However, due to the low frequency of both the falling ip boundary tone (HL-) and the undershot dipping IP boundary tone (HL(H)%), statistically significant measurements unfortunately cannot be made to compare them to each other or to the falling IP boundary tone (HL%). Instead, gross comparison of the shapes of each contour is provided below.

One clear difference between the undershot dipping IP boundary tone (HL(H)%) and the falling IP boundary tone (HL%) is the timing of the H target. The H target of the undershot dipping IP boundary tone (HL(H)%) is realized on the IP-penultimate syllable—separated from the L target by a syllable boundary, just as in non-undershot forms—as in Figure 92 below (repeated from Figure 77 above).

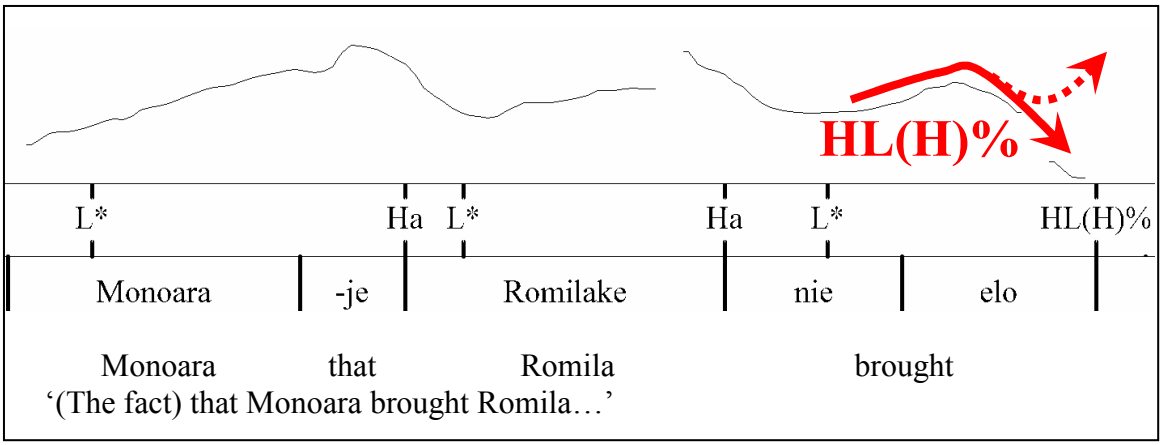


Figure 92. This large non-sentence-final clause is marked on the right edge by a dipping IP boundary tone (HLH%), whose final H target is not reached. Its pitch maximum (149Hz) is located during the vowel [e]. [Ro15]

The H target of the falling IP boundary tone (HL%), in contrast, is realized on the IP-final syllable along with the L target, as in Figure 93 below, produced by the same speaker as in Figure 92 above. Both H and L targets are reached within one syllable.

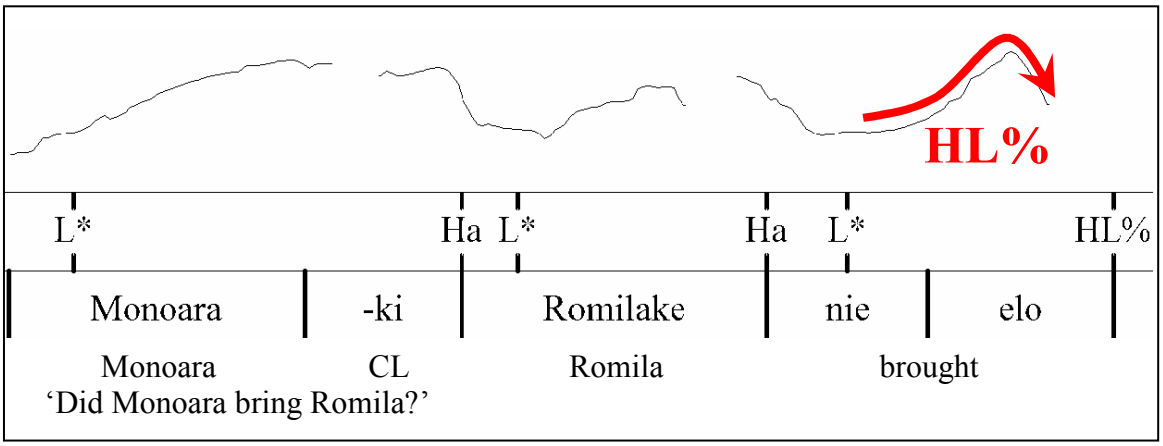


Figure 93. This yes/no question is marked on the right edge by a falling IP boundary tone (HL%). Its pitch maximum (162Hz) is located during the vowel [o]. [Ro03]

The falling IP boundary tone (HL%) can also be distinguished from the falling ip boundary tone (HL-) due to the effects of the ip boundary tone locality constraint. Note how the H target of the falling IP boundary tone (HL%) causes the preceding pitch to rise (rising from 197Hz on [i] to 223Hz at the end of [e]) from the IP-final pitch accent all the way to the IP-final syllable in Figure 94 below.

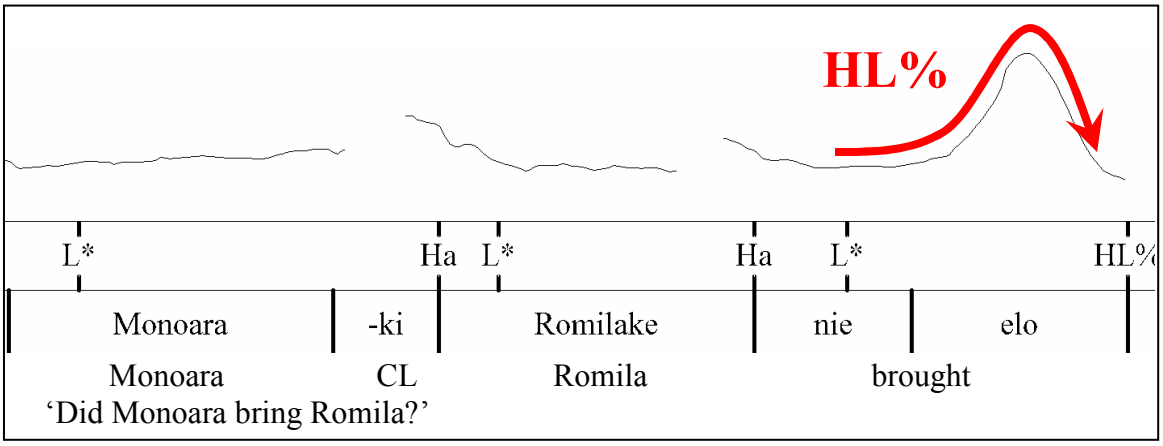


Figure 94. The yes/no question [monoaɹa-ki ɹomilake nie elo] 'Did Monoara bring Romila?' bears a falling IP boundary tone (HL%) at its right edge. The pitch following the IP-final pitch accent rises steadily from 197Hz to 223Hz (the end of [e]), and then jumps to 398Hz during the onset of the final syllable. [Do03]

Compare the upward pitch interpolation in Figure 94 above to the situation in Figure 95 below (repeated from Figure 52 above), where the falling ip boundary tone (HL-) has no ability to raise the pitch between the ip-final pitch accent and the ip-final syllable.

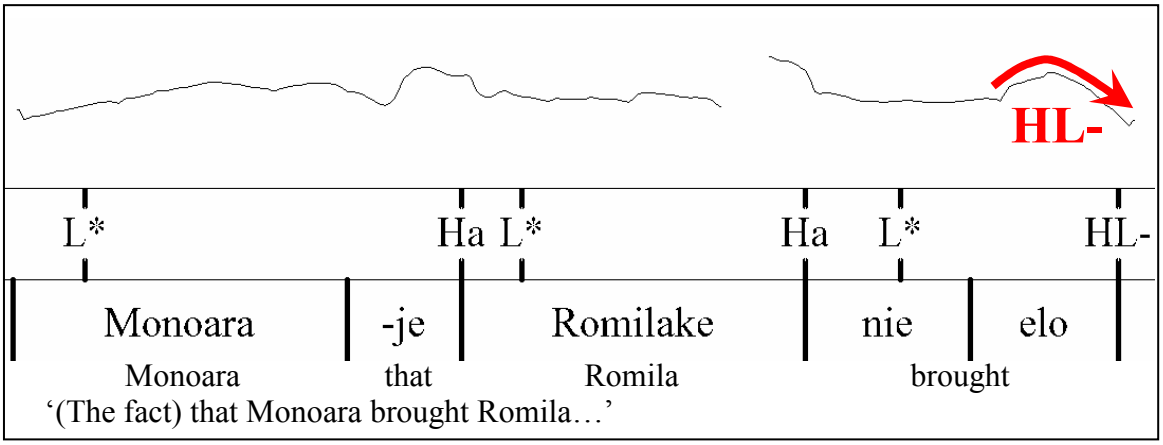


Figure 95. This dependent clause bears a falling ip boundary tone (HL-) at its right edge. The pitch following the IP-final pitch accent remains largely flat (between 218-221Hz) until the onset of the final syllable, where it jumps to 256Hz. [Do15]

The last two examples also show that the relative pitch of the falling IP boundary tone (HL%) is also higher than that of the falling ip boundary tone (HL-); while the H target realized during the IP-final syllable in Figure 94 reaches 398Hz (preceding L* level: 197Hz), the H target realized during the ip-final syllable in Figure 95 reaches only 256Hz (preceding L* level: 218Hz).

By generalizing from a small set of minimal pairs, it appears that the three HL boundary tones (*i.e.* HL-, HL%, HL(H)%) can be distinguished by three characteristics of their H target: timing, ability to raise preceding pitch, and relative pitch height. Given the lack of sufficient data, however, it is unclear how widespread these distinguishing features of the various HL tones truly are.

10.1.5 Summary

Table 19 below summarizes the full inventory of postlexical tones in the current intonational phonological model of Bangladeshi Standard Bengali and the accompanying B-ToBI transcription system.

Association	**Target**
Pitch accents	H*, L*, L*+H
AP boundary tones	Ha, (La)
ip boundary tones	H-, LH-, (HL-), (L-)
IP boundary tones	H%, L%, LH%, HL%, HLH%

Table 19. Full inventory of postlexical tones in the B-ToBI transcription of the current intonational phonological model of Bangladeshi Standard Bengali. Tones that occur infrequently in the corpus and/or occur primarily in recordings of other dialects are enclosed in parentheses.

In addition to the largely predictable tonal phenomena (*i.e.* Ha downtrend, H* downtrend, concurrent boundary tone overriding), the current model describes several tonal variations whose distributions are not fully predictable, and which consequently are transcribed with special diacritics in the B-ToBI system. Table 20Table 19 below summarizes these particular variants of tones.

Tonal change	Environment
Raising of L*+H L*+H → ^L*+H	On wh-words? On voiceless-initial words?
Downstep of H- H- → !H-	Boundary of a small (*i.e.* non-clausal) ip
Upstep of L* L* → ¡L*	Preceding a H-initial IP tone (*i.e.* H%, HL%, HLH%)
Undershoot of HLH% HLH% → HL(H)%	Unclear

Table 20. Variants of postlexical tones and the environments in which they occur.

In Table 21 below, tonal sequences observed in the current study of Bangladeshi Standard Bengali and introduced in this chapter are compared with those described in previous prosodic models of other variants of Bengali. Additional sequences introduced in Chapter Four (*e.g.* fH*…L%, L*+H…HL%, L*+H…H%) are not included.

Kolkata Standard Bengali			Eastern Bengali	Bangladeshi Standard Bengali
Hayes & Lahiri (1991), Truckenbrodt (2003)	**Lahiri & Fitzpatrick-Cole (1999)**	**Selkirk (2006)**	**Michaels & Nelson (2004)**	**Current study**
L*…H_P	L*…H_P	T*…[H]	L*…H-	L*…Ha
				L*…H-
L*…H_I	L*…H_I	T*…$[H]_{INT}$	*not described*	L*…H%
(L+H*…L_I)[70]	(L+H*…L_I)	*not described*	L*… L%	L*…L%
L*…H_IL_I	L*…H_IL_I	T*…$[HL]_{QUES}$	L*…HL%	L*…HL%
(L+H*…L_IH_I)[70]	(L+H*…L_I)	*not described*	*not described*	L*…HL-
not described	*not described*	*not described*	*not described*	H*…La
H*…L_I	H*…L_I	T*…$[L]_{DECL}$	*not described*	H*…L%
L*…H_P…L_I	L*…H*…L_I	T*…$[H]_{FOC}$ …$[L]_{DECL}$	L*+H…L%	L*…fHa…L%[71]
	L*…H_P…L_I			L*+H…L%
L*…H_P…L_IH_I	L*…H_P…L_IH_I	T*…$[H]_{FOC}$ …$[LH]_{CONT}$	L*+H…H%	L*…HLH%
				L*+H…LH%[71]

Table 21. Comparing the representation of selected tonal sequences across different models of Bengali intonation. In places where the particular model makes no distinction between two sequences in another model, the cells are merged.

10.2 Phrasing and break indices

In adopting the ToBI framework, the transcription of tone must go hand in hand with the transcription of perceived phrasal disjuncture size, or break index. After each word

[70] Although sequences of L tones are not attested in Hayes & Lahiri's (1991) model of Kolkata Standard Bengali, the sequences I describe as L*…L% and L*…LH% in Bangladeshi Standard Bengali may correspond to the downstepped high pitch accent (L+H*) followed by the low or rising (L_I, L_IH_I) I-phrase boundary tones of Hayes & Lahiri (1991).

[71] This tonal sequence is introduced in Chapter Four.

transcribed in the Word Tier of a ToBI transcription, there must be a corresponding numerical break index in the Break Index Tier. Larger numbers denote larger perceived breaks—which can be affected by final lengthening, the existence and duration of pause, changes in voice quality (*e.g.* final creakiness), segmental alternations, and other suprasegmental phenomena—and larger perceived breaks should denote the disjunctures between higher phrases in the prosodic hierarchy. I begin with a brief description of break indices in other ToBI systems, and then summarize the various cues to disjuncture size in §10.2.1. The handling of ambiguous disjuncture size is discussed in §10.2.2. A summary of phrasing and break indices is provided in §10.2.3.

Languages described as having only one tonally-marked prosodic unit tend to require fewer break indices in their prosodic transcriptions. For example, only three break indices are used in the C_ToBI model of Hong Kong Cantonese (Wong *et al.* 2005), as shown in Table 22 below.

Break index	**Disjuncture represented**
0	syllable boundary
1	foot boundary
2	IP boundary

Table 22. Break indices used in the C_ToBI transcription of Hong Kong Cantonese.

For languages with two higher level prosodic units marked by tone, an additional break index is provided for the ip *vs.* IP distinction, as in the GRToBI transcription of Standard Athens Greek (Arvaniti & Baltazani 2000), reproduced below in Table 23.

Break index	**Disjuncture represented**
0	word-clitic boundary
1	word boundary
2	ip boundary
3	IP boundary

Table 23. Break indices used in the GRToBI transcription of Standard Athens Greek.

Mainstream American English and Catalan are like Standard Athens Greek in that they are described as having two prosodic units marked by tone (*i.e.* ip, IP). However, the systems used to transcribe the prosody of Mainstream American English (MAE_ToBI; Beckman & Ayers Elam 1994, 1997) and Catalan (CatToBI; Prieto *et al.* 2007) include an additional break index; the number 2 is specifically reserved for cases in which the perceived break size is unclear (*i.e.* the "mismatch" label)—often when the tonal and non-tonal cues suggest different prosodic units. This type of system is presented in Table 24 below. (A subset of this system is used in the GToBI transcription of German; see Grice *et al.* 2005.)

Break index	Disjuncture represented
0	word-clitic boundary
1	word boundary
2	mismatch
3	ip boundary
4	IP boundary

Table 24. Break indices used in the MAE_ToBI transcription of Mainstream American English and the CatToBI transcription of Catalan.

Languages whose prosodic systems include accentual phrases (*i.e.* AP languages) use break index 2 not for mismatches between perceived disjuncture sizes, but for the AP level. Mismatches are often labeled with diacritics instead (*e.g.* m, -). The K-ToBI system for Standard Seoul Korean (Jun 1999) and the J_ToBI system for Standard Tokyo Japanese (Venditti 1997), for example, use the break indices shown in Table 25 below.

Break index	Disjuncture represented
0	word-clitic boundary
1	word boundary
2	AP boundary
3	IP boundary

Table 25. Break indices used in the transcriptions of AP languages such as Standard Seoul Korean (K-ToBI) and Standard Tokyo Japanese (J_ToBI).

While Japanese and Korean are typically described as not possessing an ip *vs.* IP distinction,[72] the current model of Bangladeshi Standard Bengali posits this distinction along with an AP level. Thus, the B-ToBI transcription system for Bengali uses break index 2 for AP level breaks as in J_ToBI and K-ToBI, while also using break indices 3 and 4 for ip and IP level breaks, respectively, as in MAE_ToBI, GToBI, and CatToBI. The B-ToBI system is shown in Table 26 below.

Break index	**Disjuncture represented**
0	word-clitic boundary
1	word boundary
2	AP boundary
3	ip boundary
4	IP boundary

Table 26. Break indices used in the B-ToBI transcription of Bangladeshi Standard Bengali (and, to some extent, of other Bengali dialects).

As in other ToBI-style transcription systems, the break indices of B-ToBI are transcribed on the third tier below the pitch track, as illustrated below in Figure 96.

[72] More recent work presents evidence of an ip, a third tonally-marked prosodic unit, in Korean (Jun 2007).

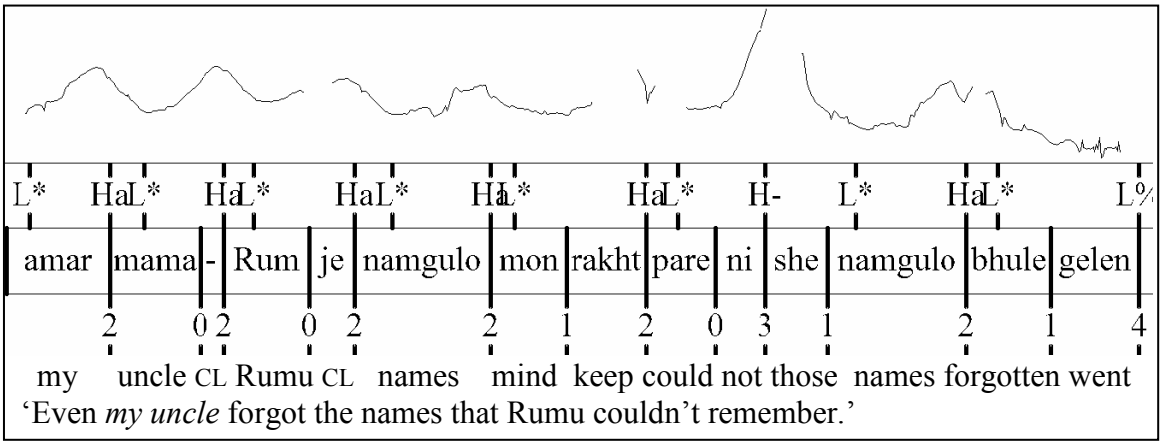


Figure 96. Break indices 0, 1, 2, 3, and 4 are all found in this sentence, transcribed on the break indices tier—the third tier under the pitch track. [Na51]

In Figure 96, all five possible break indices (*i.e.* 0, 1, 2, 3, 4) are found. Level 0, which designates a disjuncture perceived to separate a clitic from its host, is seen between [mama] 'mother's brother' and the emphatic clitic -[ɔ̯] 'also', 'even'. The breaks preceding the morphemes [dʑej] (relative clause marker) and [ni] (negation of perfect verbs) are also labeled 0. The disjuncture between [mone] 'mind-LOC' and [ɹakʰt̪e] 'keep-INF' and the disjuncture between [bʱule] 'forget-PERF' and [gelen] 'go-PAST-HON' are labeled with break index 1, identifying the disjunctures as word boundaries within a single AP. Each of these disjunctures occurs between the two halves of a complex verb ([mone ɹakʰt̪e] 'to remember'; [bʱule gelen] 'forgot-HON'). In addition, the disjuncture between the demonstrative [ʃej] 'that' and its noun [namgulo] 'name-DEF-PL' is labeled with break index 1. Most of the disjunctures in the sentence are marked with break index 2, representing perceived AP boundaries. The disjuncture between the relative clause

[ɹumu dʑej namgulo mone ɹakʰt̪e paɹe ni] ‘the names Rumu couldn’t remember’ and the correlative [ʃej namgulo] ‘those names’[73] is marked with break index 3, representing a perceived ip boundary. Finally, the break between the final word [gelen] ‘went-HON’ and the end of the sentence is marked with break index 4, representing a perceived IP boundary.

10.2.1 Cues to disjuncture size

As prosodic disjuncture could be cued by numerous phonetic cues (*e.g.* pause, duration, voice quality, segmental allophony, tone), it is necessary to examine which cues are most salient for Bengali listeners. Previous studies regarding the perceptibility of break size in other languages include Carlson *et al.* (2005)—which examined the cues to break size in Swedish as perceived by Swedes, Americans, and Chinese—and three major studies on the perception of prosodic breaks in English: Kreiman (1982), Wightman *et al.* (1992), and de Pijper & Sanderman (1994). The following subsections briefly describe the potential cues Bengali listeners could use when parsing an utterance into prosodic phrases.

Tone is presumably among the most salient cues for prosodic disjuncture. The lack of a tonal target can help identify a disjuncture below 2 (*i.e.* the level of the AP, the smallest tonally-marked prosodic unit). Additional tonal properties can help distinguish between disjunctures 2, 3, and 4, as shown previously in §10.1.4. Minimal triplets such as

[73] Bengali uses the correlative construction, and thus a noun being relativized appears both in the relative and correlative clause.

the high boundary tones at the AP (Ha), ip (H-), and IP (H%) levels demonstrate that even the same tonal target (*i.e.* H) at different levels of prosodic structure can be differentiated by relative pitch difference. As these distinctions are largely tone-specific—equivalent measurements cannot be taken across target types (*e.g.* Ha *vs.* L- *vs.* HL%). This section will focus on measurements and alternations that are not tone-specific, and can be compared regardless of the exact type of boundary tone present at the boundary in question. These include durational measurements (*i.e.* final syllable duration, following pause) and segmental alternations (*i.e.* vowel coalescence, onset lenition).

Final syllable duration

One cross-linguistically common property of the ends of prosodic units is the lengthening of the final syllable or segment (see Wightman *et al.* 1992; Jun 2005). By comparing the durations of syllables when occurring adjacent to various boundary tones (*i.e.* T%, T-, Ta), it is clear that IP-final syllables are longer than ip-final syllables, which are longer than AP-final syllables, suggesting that Bengali has IP- and ip-final lengthening. For example, compare the duration of the syllable [ke] in [ninake] 'Nina-ACC', when AP-final and when ip-final in Figure 97 below. Note how in addition to the tonal distinctions between the two examples (*i.e.* Ha *vs.* H-), the relative (and, in this case, also the absolute) duration of the syllable [ke] is longer in ip-final position (540ms, 44.1%) than in AP-final position (403ms, 29.8%).

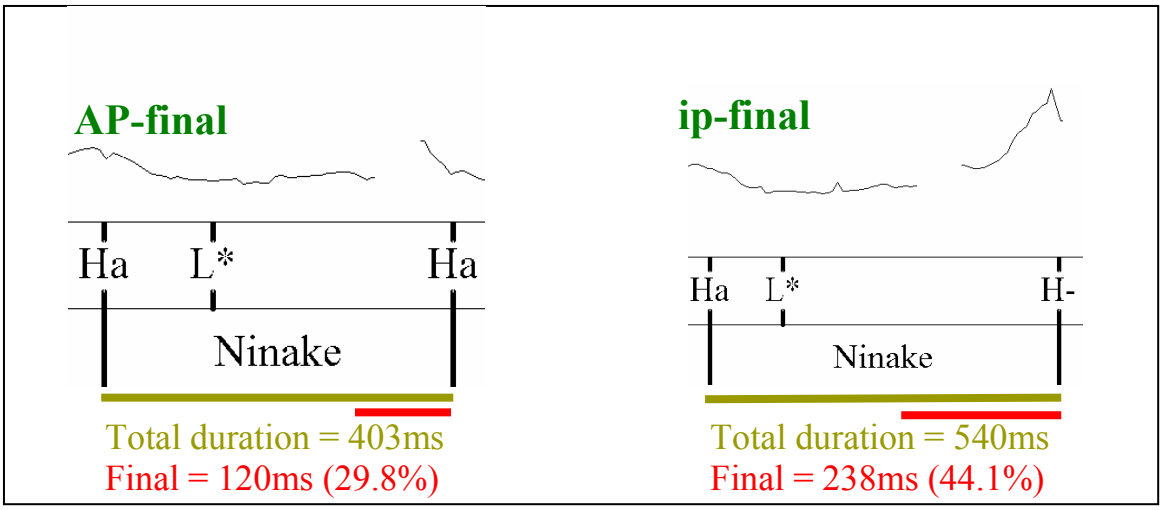


Figure 97. Measuring within speaker, the duration of the final syllable [ke] relative to the whole word [ninake] 'Nina-ACC' is longer ip-finally (right) than AP-finally (left) in otherwise identical sentences [paired t(8) = 3.05, p = .02].

The relative durations of the final syllables of ten pairs of identical (*i.e.* same word, same speaker, same syntactic position) APs and ips are listed in Table 27 below.

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

Table 27. Durational differences between the final syllables of ten non-ip-final APs and their corresponding ip-final APs, measured as a percentage of the total word duration, measured within speaker.

By comparing the relative durations of AP- and ip-final syllables listed in Table 27, it was found that ip-final syllables are significantly longer than AP-final syllables, as shown in Figure 98 below [paired $t(8) = 3.05$, $p < .05$].

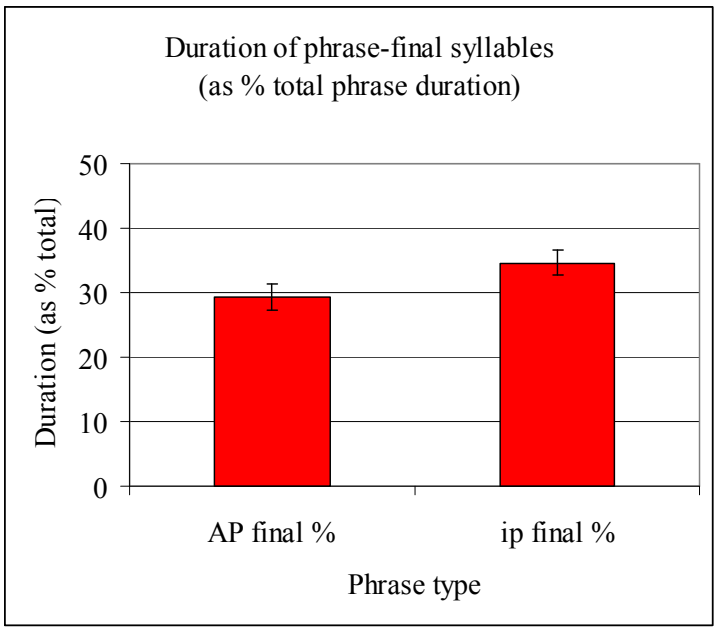


Figure 98. Relative duration of AP-final syllables compared to that of ip-final syllables. Error bars indicate standard error.

Despite the lengthening seen at the ip level, no evidence was found for AP-final lengthening. Although the AP level is ranked higher than the word level in the prosodic structure of Bengali, this hierarchical distinction is primarily relevant in terms of tonal phenomena. The AP is not significantly distinguishable from the word in terms of final lengthening or the duration of the pause following the boundary (as neither the word nor the AP is followed by a pause). Supporting evidence was found in examining a string of two words (*i.e.* permitting the production of two APs) that form a tight semantic constituent with relatively short length (*i.e.* permitting lumping into one AP) across speakers: [lina mamike] 'Aunt Lina-ACC'. Using the tonal pattern as a diagnostic of

phrasing, it was possible to determine whether the phrase [lina mamike] 'Aunt Lina-ACC' was produced with one (*i.e.* L*…Ha) or two (*i.e.* L*…Ha…L*…Ha) APs. Then, by comparing the duration of the syllable [na] as a percentage of the entire word [lina] across the two phrasing possibilities—(a) two words composing a single AP and (b) two words composing two separate APs—produced by different speakers, it was found that there is no significant difference between the two phrasing possibilities [$t(10) = 0.04$, $p = 0.97$]. The absolute and relative durations of the strings [lina mamike], [lina], and [na] are provided below in Table 28.

1 AP (L*…Ha)					2 APs (L*…Ha)(L*…Ha)			
Speaker	**[lina mamike]**	**[lina]**	**[na] (% of [lina])**		**Speaker**	**[(lina) (mamike)]**	**[lina]**	**[na] (% of [lina])**
Az	754	311	144		Ba	798	358	128
			46%					36%
Bo	689	293	112		BM	1014	457	217
			38%					47%
Do	780	374	180		Jh	1030	481	241
			48%					50%
Fa	647	324	155		Re	704	306	122
			47%					40%
Fe	845	313	163		Ro	812	382	176
			52%					46%
Jo	602	260	108		Sh	938	437	197
			42%					45%
Pi	614	277	97					
			35%					
Sf	769	319	134					
			42%					
To	660	289	125					
			43%					
Tu	837	383	173					
			45%					

Table 28. Absolute and relative durations (in ms and %, respectively) of the strings [lina mamike], [lina], and [na] in the phrase [lina mamike] 'Aunt Lina-ACC', produced as a single AP (left), or as two separate APs (right). The relative duration of [na] in the word [lina] is not significantly different across the 1-AP and 2-AP phrasings [$t(10) = 0.04$, $p = 0.97$], suggesting that there is no AP-level lengthening.

Given that the relative duration of word-final syllables is not significantly different when AP-final and AP-medial, it is clear that final lengthening is not a reliable cue for distinguishing break indices 1 (word level) and 2 (AP level). This is not surprising,

considering that lengthening is often only used to mark larger boundaries in other languages (Jun 1995).

Following pause: a cue for break indices 3 and 4

With the exception of slow, extra-careful speech, pauses between words are indicative of an ip or IP break, corresponding to break indices 3 and 4, respectively. For example, compare the following recordings of the subject [monoaɹa] 'Monoara' preceding the direct object [ɹomilake] 'Romila-ACC'. In the first instance (left), the subject [monoaɹa] 'Monoara' is parsed as a separate ip (as can be seen in the use of the high ip boundary tone H-), and thus it can be followed by a pause. However, in the second instance (right), [monoaɹa] and [ɹomilake] are not separated by an ip break (as can be seen in the use of the high AP boundary tone Ha), and thus no pause can be produced.

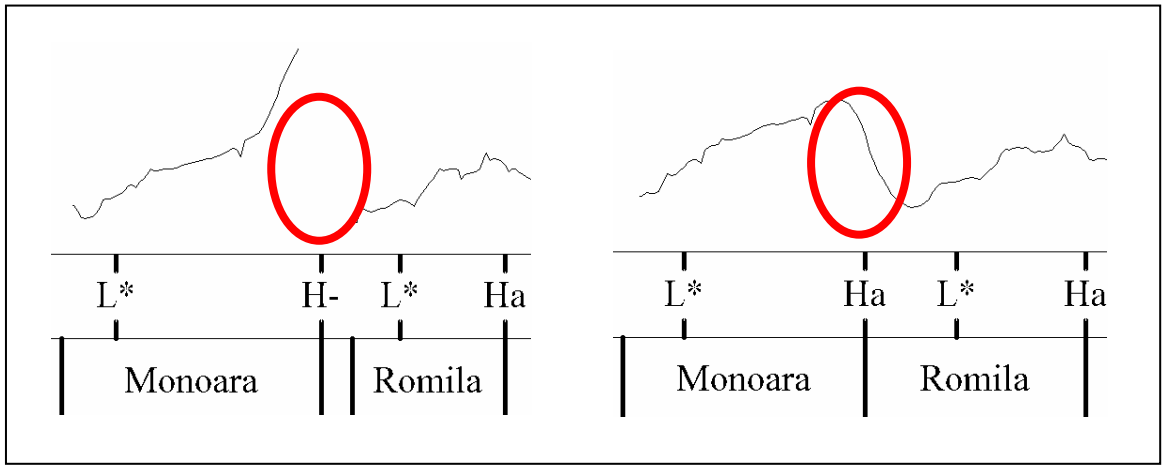


Figure 99. While the subject [monoaɹa] 'Monoara' in the speaker's first production is followed by a pause (evident in the lack of a pitch track between the two words), the same word in the second production is not—and cannot—be followed by a pause (evident in the continuity of the pitch track between the two words), as the break following [monoaɹa] is not an ip or IP break. [Ba28], [Ba26]

Thus, although pauses are not always obligatory, the existence of a pause can indicate a disjuncture larger than break index 2 (AP level).

Vowel coalescence: a cue for break index 0

As introduced in Chapter Two, vowels lose their vocalic status when following other vowels (NB: only particular vowel combinations undergo this alternation), surfacing as off-glides of diphthongs. This is always true of suffixed vowels (*e.g.* /hɔ-e/ → [hɔe̯] 'it happens', /d̪ɛ-o/ → [d̪ɛo̯] 'you give'), and also true of encliticized vowels when immediately following the primary stressed syllable (*e.g.* /kʰala-o/ → [kʰalao̯] 'also aunt',

/gɛlo-i/ → [gɛloj] ‘nevertheless went’). This is never true across word boundaries (*e.g.* /koɹia o dʑapan/ → [koɹia o dʑapan] ‘Korea and Japan’, *[koɹiao̯ dʑapan], /aʃa utɕit̪/ → [aʃa utɕit̪] ‘should come’, *[aʃawtɕit̪]). Thus, the coalescence of two vowels into a vowel-offglide diphthong can be seen as a cue for a break smaller than a word, *i.e.* break index 0. In Figure 96 above, the noun /mama/ ‘uncle’ ends in a vowel [a] that is adjacent to the enclitic -/o/ ‘also’. The string is realized as a diphthong [ao̯], and thus the boundary between the two morphemes /mama/ and -/o/ is labeled with break index 0.

Onset lenition: a cue for break indices 0 and 1

Intervocalic stops and nasals are often lenited into their corresponding fricatives or approximants. Thus, /p, b, bʱ, m/ lenite to [ɸ, β, βʱ, ṽ], /k, kʰ, g, gʱ/ lenite to [x, xʰ, ɣ, ɣʱ], and so on. In a few cases, intervocalic voiceless stops are produced with voicing (*e.g.* /p/ → [b]); this voicing of underlyingly voiceless stops can be considered another type of lenition. Both types of lenition (*i.e.* loss of stop closure, voicing) are seen even word-initially; however, these processes are blocked when the consonant is the first in a tonally-marked domain (*i.e.* AP-initial, ip-initial, IP-initial). Compare the following two examples of the NP [lina mamike] ‘Aunt Lina-ACC’ in Figure 100. In the first recording (left), the speaker parses the two words [lina] and [mamike] together into a single AP. Thus, since the first /m/ in [mamike] is intervocalic and not initial in a tonally-marked domain, it can lenite to [ṽ]. However, in the second recording (right), another speaker parses the two words [lina] and [mamike] into two separate APs. Thus, since the first /m/

in [mamike] is AP-initial, it cannot undergo lenition to [ṽ]. This resistance to lenition can be considered a form of initial strengthening, seen in various languages (Fougeron & Keating 1997, Jun 1998, Fougeron 1999, Cho & Keating 2001, Keating *et al.* 2003).

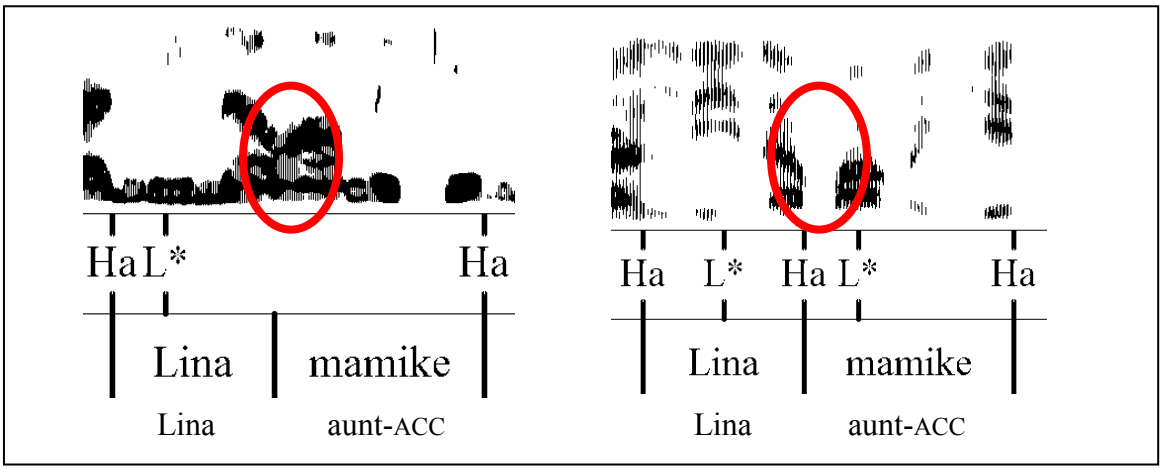


Figure 100. While the word-initial /m/ in the first speaker's production of [lina mamike] 'aunt Lina-ACC' shows clear signs of lenition (*i.e.* evidence of strong formant structure during the consonant), the word-initial /m/ in the second speaker's production of the same phrase is not lenited (as evident in the overall lack of acoustic energy during the consonant), and in fact cannot be lenited due to its AP-initial position. [To24], [Re24]

Although lenition is not obligatory in any context, it is indicative of a boundary smaller than that of an AP. Thus, the existence of a lenited stop or nasal can be a cue for break indices 0 or 1. This can be particularly helpful in cases where the tonal cues are not sufficient to determine whether a word is domain-initial in an AP (break index 2).

I now discuss cases of mismatch between the various cues for boundary tones, and the appropriate transcription of such ambiguous cases.

10.2.2 Ambiguous disjuncture size

In fast, running speech, it is particularly easy to differentiate ip breaks from corresponding IP breaks, by examining the shape of the contour (as affected by the ip boundary tone locality contour), the relative pitch height (discussed in §10.1.4), and the syntactic distribution (*e.g.* the high ip boundary tone H- typically occurs after small, often topicalized, phrases, while the high IP boundary tone H% typically occurs after certain interrogatives).[74] However, when speakers produce larger breaks due to slower overall speech rate, unfamiliarity with the topic of conversation, or other hesitation, the boundary tones before certain large breaks are ambiguous in their identification. In terms of distribution (*i.e.* setting off a topic or other small phrase instead of marking an interrogative or declarative) and tonal interaction (*i.e.* not causing upstep of the preceding low pitch accent, L*), they pattern with ip boundary tones, while phonetically (*i.e.* in terms of pitch height, final lengthening, and following pause), they can sometimes resemble IP breaks and corresponding boundary tones. Furthermore, the contour shape often does not follow the ip boundary tone locality constraint, while in many instances it does. For example, the high boundary tone at the end of the ip [monoaɹa] 'Monoara' in Figure 101 below resembles a high ip boundary tone (H-) in that it obeys the ip boundary tone locality constraint, and because it occurs at the right edge of a topicalized subject. However, the pitch height (184Hz, even higher than the preceding H% not shown here) and large following pause (450ms) resemble characteristics of the high IP boundary tone (H%). Thus, the boundary tone is labeled H-%, and its category is left ambiguous.

[74] However, see Woodbury (1987) for a discussion of ambiguous disjuncture size in running speech.

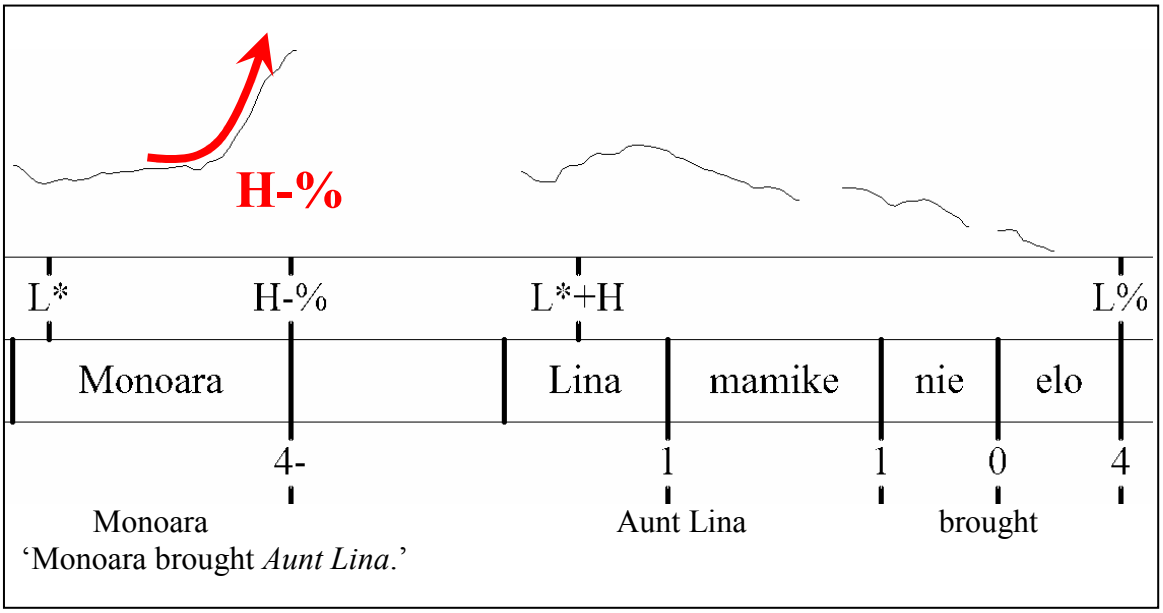


Figure 101. The phrase [monoaɹa] 'Monoara' is marked on its right edge with a high boundary tone whose category is ambiguous. The tone is thus labeled H-%, and its break index can be labeled 3+ or 4- depending on which cues are perceived as stronger by the transcriber. [Ba25]

In terms of break index, the mismatch between the various cues leaves the perceived disjuncture size ambiguous between 3 (based on contour shape and syntactic distribution) and 4 (based on relative pitch height and following pause), and thus the break indices 3+ or 4- can be used, depending on which cues are perceived more strongly; if the disjuncture is primarily perceived as an ip break, but with some properties of an IP break, it should be labeled 3+ (*i.e.* 3, possibly higher), while it should be labeled 4- (*i.e.* 4, possibly lower) if the disjuncture is primarily perceived as an IP break, but with some properties of an ip break. This convention is similar to the "mismatch" labeling in other ToBI transcriptions. For example, the break following the subject [monoaɹa] 'Monoara'

in Figure 101 above is labeled 4-, signifying that the disjuncture was perceived as an IP boundary with some characteristics of an ip boundary.

Similarly, the falling boundary tone at the end of the phrase [monoaɹa ɹomilake nie elo-dʐe] ‘(the fact) that Monoara brought Romila’ in Figure 102 below resembles the falling IP boundary tone (HL%) in terms of lengthening of the IP-final syllable [dʐe], upstep of the final pitch accent (¡L*), and the violation of the ip boundary tone locality constraint. However, the lack of a following pause, as well as its location at the end of a dependent clause (associated with ip boundary tones) instead of a non-negative yes/no question (associated with the falling IP boundary tone HL%) more closely resemble the falling ip boundary tone (HL-). Therefore, the boundary tone is labeled HL-%, and its category is left ambiguous. Furthermore, as I perceived the disjuncture to be more characteristic of an IP boundary than of an ip boundary, the break index after -[dʐe] is 4-.

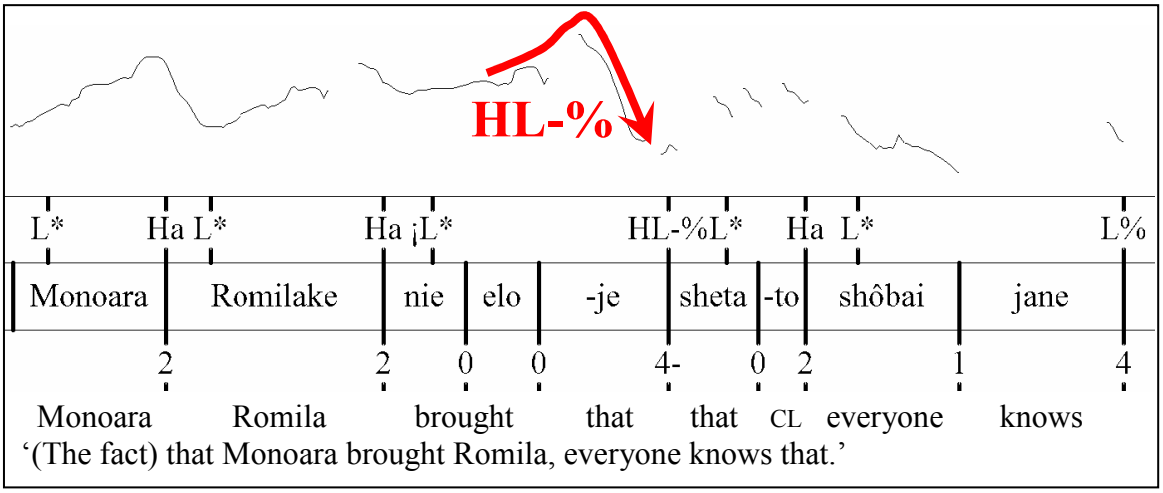


Figure 102. The phrase [monoaɹa ɹomilake nie elo-dʑe] '(the fact) that Monoara brought Romila' is marked on its right edge with a falling boundary tone whose category is ambiguous between the ip and IP levels. The tone is thus labeled HL-%, and its break index can be labeled 3+ or 4- depending on which cues are perceived as stronger by the transcriber. In this instance, the break index given is 4-. [Fa16]

Similar cases of ambiguous category of tones are found between the low ip and IP boundary tones (*i.e.* L- and L%) and the rising ip and IP boundary tones (*i.e.* LH- and LH%). Given that examples of these ambiguous tones are relatively infrequent in the corpus of data collected for the current study, I do not attempt to categorize them. Instead, I leave this area open to further investigation and interpretation. One interpretation of this apparent overlap in equivalent ip and IP boundary tones would be that there is no distinction between the two phrase types. Indeed, previous Intonational Phonological models of Bengali all agree on the existence of only two levels of tonally-marked prosodic units. Such an interpretation, however, would ignore the F_0, durational, and

distributional properties that differentiate minimal pairs between the equivalent boundary tones of the two units (*e.g.* H- *vs.* H%, HL- *vs.* HL%, *etc.*). Thus, instead of collapsing the two units, the current model proposes that ips and their corresponding boundary tones (T-) can optionally be promoted to IPs and their corresponding boundary tones (T%) in very careful speech or speech characterized by frequent hesitations and pauses. Since these ambiguous tones more closely resemble IP boundary tones in their phonetic realization, they can be considered equivalent to their corresponding prototypical IP boundary tones, and can be labeled as such (*i.e.* H%, L%, HL%, LH%). However, if the transcriber wishes to distinguish between prototypical IP boundary tones and ambiguous ip-IP boundary tones, either to characterize its distribution (*e.g.* marking a topic, as expected of H-, instead of marking an interrogative, as expected of H%) or its phonetic realization (*e.g.* adherence to the ip boundary tone locality constraint, non-extreme pitch levels, non-extreme final lengthening, *etc.*) the tones can be labeled with both the hyphen (-) associated with ip boundary tones and the percentage sign (%) associated with IP boundary tone (*i.e.* H-%, L-%, HL-%, LH-%). Regardless of the choice of transcription, the underlying characterization of these ambiguous tones is left undetermined for the time being.

Similarly, the high AP boundary tone can, in limited contexts, defy some of its defining characteristics. When reading lists, speakers can either assign the smooth rise AP tonal pattern (L*…Ha) or a combination of low pitch accent and high ip boundary tone (L*…H-) to each of the non-final members of a list. When speaking slowly (*e.g.* due

to careful reading), however, some of the H boundary tones can be ambiguous in their properties, resembling high AP boundary tones (Ha) in some respects and high ip boundary tones (H-) in others. For example, in Figure 103 below, only one list member (*i.e.* [oɾia] 'Oriya') bears a largely unremarkable high ip boundary tone (H-), characterized by flat mid-level pitch followed by a sharp rise in the ip-final syllable. The other non-final list members are more ambiguous. Two list members (*i.e.* [hin̪di] 'Hindi' and [malaʒ̞alam][75] 'Malayalam') bear smoothly rising pitch, characteristic of the high AP boundary tone (Ha), but are followed by a pause, characteristic of an ip break. Two other APs (*i.e.* [baŋla] 'Bengali' and [t̪amil] 'Tamil') have sharp rises on their final syllables followed by a short pause, characteristic of the high ip boundary tone (H-), although they also bear smoothly rising pitch leading up to the ip-final syllable, characteristic of the high AP boundary tone (Ha). Due to this ambiguous combination of cues, the tone is transcribed Ha-, representing both the "a" diacritic for AP boundary tones and the "-" diacritic for ip boundary tones. The corresponding break index is 3-, signifying that the disjuncture is perceived primarily as an ip level break, but with some properties of an AP level break.

[75] The Bengali name for the Malayalam language is commonly spelled and pronounced in a variety of ways; in order of decreasing frequency on a Google search, the spellings মালায়ালাম [malaʒ̞alam], মালায়ালম [malaʒ̞alom], মলয়ালম [mɔlɔʒ̞alom], and মালয়ালম [malɔʒ̞alom] were most common. Bengali Wikipedia also lists মালয়লম [malɔʒ̞ɔlom]. Additional pronunciations (*e.g.* মলায়লম [mɔlaʒ̞ɔlom], মালায়লম [malaʒ̞ɔlom]) were found in the current study's data corpus.

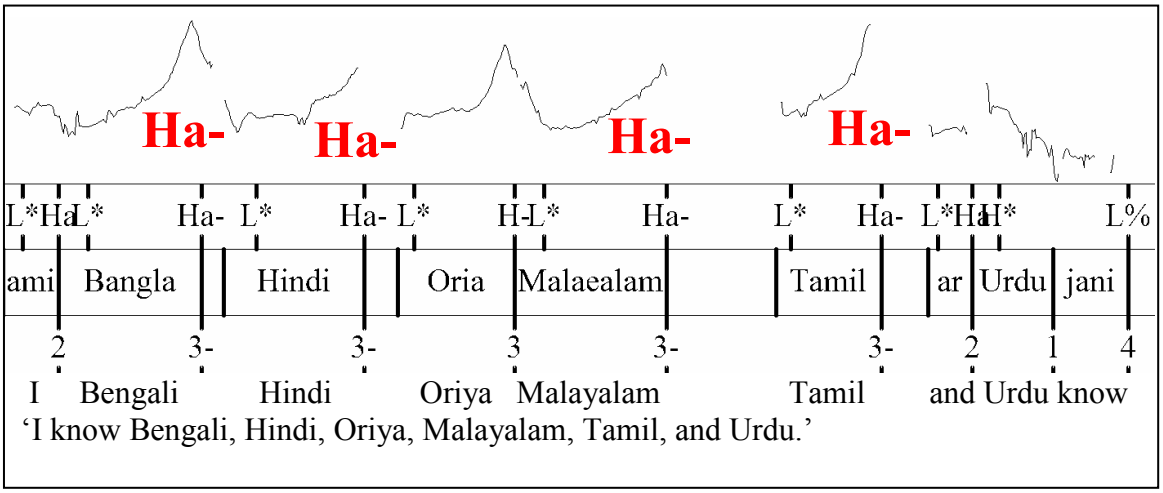


Figure 103. Four of the language names listed in this sentence bear H boundary tones, whose category is ambiguous between the AP and ip levels. The tones are thus labeled Ha-, and the corresponding break indices can be labeled 2+ or 3- depending on which cues are perceived as stronger by the transcriber. [To30]

Another speaker produced the same sentence with similarly ambiguous H boundary tones at the edges of four of the non-final list members, shown below in Figure 104. While all of the H boundary tones resemble high AP boundary tones (Ha) in their smoothly rising pitch, four of them do not obey Ha downtrend, thus resembling high ip boundary tones (H-). Furthermore, two of the APs (*i.e.* [oɾia] 'Oriya'[76] and [mɔlaɡ̊ɔlom] 'Malayalam') are followed by pauses, characteristic of higher prosodic boundaries. Thus, they are labeled Ha-, and their corresponding break indices are 2+, indicating that they are primarily perceived as AP breaks, although with properties of ip breaks as well.

[76] The Bengali name for the Oriya language varies in spelling and pronunciation. Both ওড়িয়া [oɾia] and উড়িয়া [uɾia] are found. In dialects that do not preserve the /ɹ/~/ɾ/ distinction (*e.g.* Eastern Bengali, Eastern forms of Bangladeshi Standard Bengali), these names are pronounced [oɹia] and [uɹia], respectively.

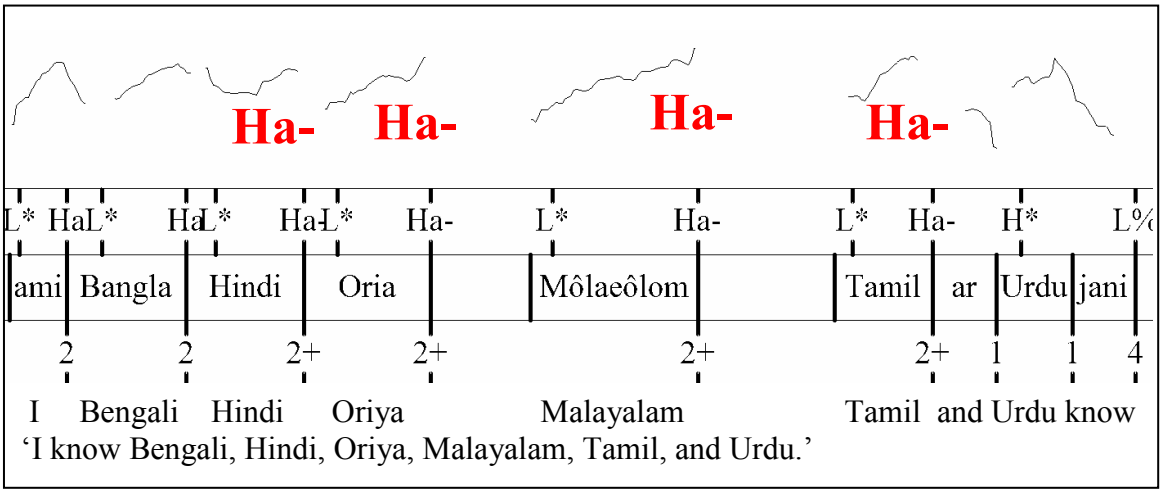


Figure 104. Four of the language names listed in this sentence bear H boundary tones whose category is ambiguous between the AP and ip levels. Their boundary tones are thus labeled Ha-, and the break indices can be labeled 2+ or 3- depending on which cues are perceived as stronger by the transcriber. In this case, they were labeled 2+. [Ro30]

As with the ambiguous ip-IP boundary tones, more data on this phenomenon—as well as input from additional transcribers—could shed light on whether to categorize these tones as AP boundary tones or ip boundary tones.

10.2.3 Summary

The full inventory of prosodic phrases in the Bangladeshi Standard Bengali inventory, including the word and the three tonally-marked units (*i.e.* AP, ip, IP) are listed in Table 29 below along with the metrical, distributional, tonal, durational, and segmental features of each unit.

Properties	Unit →	IP	ip	AP	Word
Metrical	**Head of domain**	none	none	pitch accent	primary stressed syllable
Distributional	**Syntax-sensitive?**	yes	mostly	no	yes
Tonal	**Boundary tone?**	yes	yes	yes	no
	Tone choice depends on	sentence type	unclear	pitch accent type	—
	Pitch level	extreme	moderate	subtle	—
	Tonal domain	distant	local	distant	—
Durational	**Final lengthening**	extreme	moderate	none	unclear
	Following pause	obligatory	optional	none	none
Segmental					nasal/oral stop lenition
Break index		**4**	**3**	**2**	**1**

Table 29. Distributional, tonal, durational, and segmental properties of IPs, ips, APs, and words in the Bangladeshi Standard Bengali prosodic hierarchy.

CHAPTER FOUR
FOCUS

Focus is a feature that highlights the part or parts of a sentence that are considered to be particularly salient in the information structure (Ladd 1980, Kálmán & van Leusen 1993, Selkirk 1995, É.Kiss 1998, Beaver & Clark 2008). Answers to wh-questions, corrections or contrasts with previous statements, information requiring special emphasis, or otherwise unexpected information can be marked with this underlyingly semantic-pragmatic feature, while other parts of the sentence can be considered to have neutral or no focus. Applying focus to a constituent can affect word order, the assignment and arrangement of prosodic structure, and the phonetic realization of the prosody (Büring 1997, Büring & Jun 2006). In this sense, focus serves as an interface between pragmatics, semantics, syntax, phonology, and phonetics. The goal of this chapter is to characterize the prosodic expression of focus in Bangladeshi Standard Bengali, based on multiple speakers' data collected in Experiment I, and described in Chapter Three. For the purposes of this chapter, I will use the terms *neutral focus* and *no focus* interchangeably. Furthermore, *wh-answer focus* and *corrective focus* can be taken to be prosodically equivalent, despite their possible semantic differences. *Encliticized focus* and *surprise focus* are terms coined in the current study for two additional types of focus, whose prosodic properties differ from the *corrective* and *wh-answer* type.

Previous studies of Bengali prosody agree that focus is realized in the prosody as rising pitch on the focused constituent, and that post-focal words do not display strong

changes in the pitch contour, instead typically bearing smoothly falling pitch. Considering that non-focused constituents are also characterized as bearing rising pitch, the three existing Intonational Phonological models of Kolkata Bengali (introduced in Chapter Two) describe focus realization using either the same or very similar structure as the default rising pattern applied to non-focused constituents: a low pitch accent (L*) followed by a high tone associated to a prosodic right-boundary (Hayes & Lahiri 1991: L*…$[H]_P$, Lahiri & Fitzpatrick-Cole 1999 L*…$[H]_P$, Selkirk 2006: L*…$[H]_{FOC}$). Lahiri & Fitzpatrick-Cole (1999) add another pattern for constituents attached to focus enclitics: L*…H*. The Michaels & Nelson (2004) study of Eastern Bengali finds the rising pitch on focused constituents to be fundamentally distinct from the rising pitch on non-focused constituents, attributing the rising pitch on focused constituents to a bitonal pitch accent (L*+H) instead of to the interpolation of pitch from the default low pitch accent (L*) and high ip boundary tone (H-) associated with non-focused constituents. Experimental data from the current study, indicate that Bangladeshi Standard Bengali in fact employs focus realization patterns similar to Lahiri & Fitzpatrick-Cole's (1999) L*…H* and Michaels & Nelson's (2004) L*+H. In fact, the current study's corpus of data reveals three distinct focus realization patterns, each of which uses a non-default high tone: the focused smooth rise (L*…fHa), the sharp rise (L*+H), and the focused high pitch accent (H*). I will show how all three of these can be analyzed as default AP tonal patterns (as described in Chapter Three) with the addition of an underlying focus high tone (fH), which can only be realized when fused with or adjoined to a default AP-level tone (*i.e.* pitch accent or AP boundary tone).

In this chapter, examples of the three focus tonal patterns are analyzed and annotated with the concepts and labels introduced in Chapter Three. Experimental data with acoustic calculations are used along with theoretical rationale as support for the particular analysis presented in the current study. I begin by describing the three tonal patterns associated with focused constituents (§11), including both theoretical and empirical support for a bitonal pitch accent as opposed to a boundary-marking tone in the most basic focused phrase types, and then continue with a description of the phrasing patterns of focused constituents as well as of post-focal material (§12).

11 TONAL PATTERNS

While previous studies describe either boundary-aligned H tones or bitonal pitch accents in focused phrases, the current study finds in addition to these patterns a third pattern associated with focus. Each of the three focus tonal patterns contains a surface realization of the underlying high focus tone (fH), which must fuse with or adjoin to other tones. In §11.1, I describe the focused smooth rise tonal pattern (L*…fHa), the result of fusion of the focus high tone (fH) with the high AP boundary tone (Ha) seen in default non-focused phrases. Unlike the default smooth rise (L*…Ha), however, the high AP boundary tone (Ha) associated with focused phrases is realized with a greatly raised pitch (fHa), due to fusion with the focus high tone (fH). This pattern is most similar to Lahiri & Fitzpatrick-Cole's (1999) L*…H* pattern. The focus realization pattern most frequently observed in the current study, however, uses the rising pitch accent (L*+H) or sharp rise

AP tonal pattern, as first described by Michaels & Nelson (2004), which is the surface realization of the adjunction of the low pitch accent (L*) and the focus high tone (fH). I provide examples of this focus tonal pattern along with comparisons to default non-focused constituents in §11.2. Lastly, the focus high tone (fH) can fuse with the high pitch accent (H*)—raising it to fH*—as a marker of focus, as discussed in §11.3. The underlying structures resulting in these three patterns are schematized in Figure 105 below.

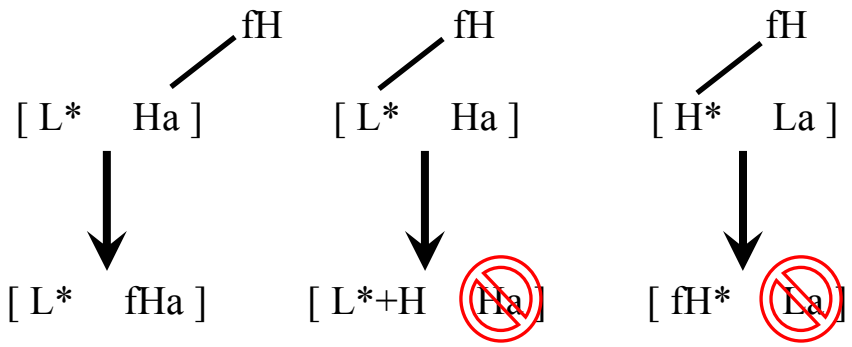


Figure 105. Schematic illustration of the three possible docking points of the underlying focus high tone (fH). In the first example, the high AP boundary tone (Ha) of the smooth rise (L*…Ha) serves as the docking point, becoming a focused high AP boundary tone (fHa). In the second example, the low pitch accent (L*) of the smooth rise (L*…Ha) serves as the docking point, becoming a bitonal pitch accent (L*+H). In the third example, the high pitch accent (H*) of the smooth fall (H*…La) serves as the docking point, becoming a focused high pitch accent (fH*). In the second and third cases, crossed circles represent post-focal deletion of AP-level tones; see §12.2 for a discussion of this phenomenon.

The three focus tonal patterns are summarized in §11.4.

11.1 Focused smooth rise (L*…Ha → L*…fHa)

Arguably the most intuitive way to transform a non-focused AP into a focused AP is to raise the pitch maximum during the focused string and lower the pitch maxima thereafter. This can serve to not only highlight the importance of the AP intended to be focused, but can also indicate that all other (more accurately, all following) information is of lesser importance, without having to introduce entirely new tonal patterns. The raising of the pitch maximum of the focused string's high AP boundary tone is analyzed here as the result of the fusion of the boundary tone with an underlying focus high tone (fH), giving the focus tonal pattern known as focused smooth rise (L*…fHa). This is composed of a low pitch accent (L*), much like the default pitch accent of non-focused phrases, and a high AP boundary tone (Ha) fused with the focus high tone (fH), creating fHa—labeled with the "f" diacritic to indicate that the raised pitch value is due to the constituent's focused status. The optional compression and/or deletion of tones following the focused phrase is described more thoroughly in §12.2.

The focused smooth rise tonal pattern (L*…fHa) is used exclusively on words whose focused status is made segmentally explicit through the use of focus enclitics (*i.e.* -[i]~[j] 'only', 'indeed' or -[o]~[ɔ̰] 'also', 'even'). As first introduced in Lahiri & Fitzpatrick-Cole (1999), focus enclitics attach directly to the right edge of the word under focus. Thus, while the string [monoaɹa ɹomilake nie elo] can be translated as 'Monoara brought Romila', the string [monoaɹaj ɹomilake nie elo] makes explicit that it should be translated '(Only) *Monoara* brought Romila' through the use of the focus enclitic -[j] 'only', 'indeed' at the right edge of [monoaɹa] 'Monoara'. Similarly, by using the focus enclitic -

[ɔ̯] ‘also’, ‘even’, the string [monoaɹaɔ̯ ɹomilake nie elo] could be translated ‘Even *Monoara* brought Romila’. An example of fH fusion with the high AP boundary tone in a focus-encliticized AP accompanied by post-focal tone deletion is provided in Figure 19 below.

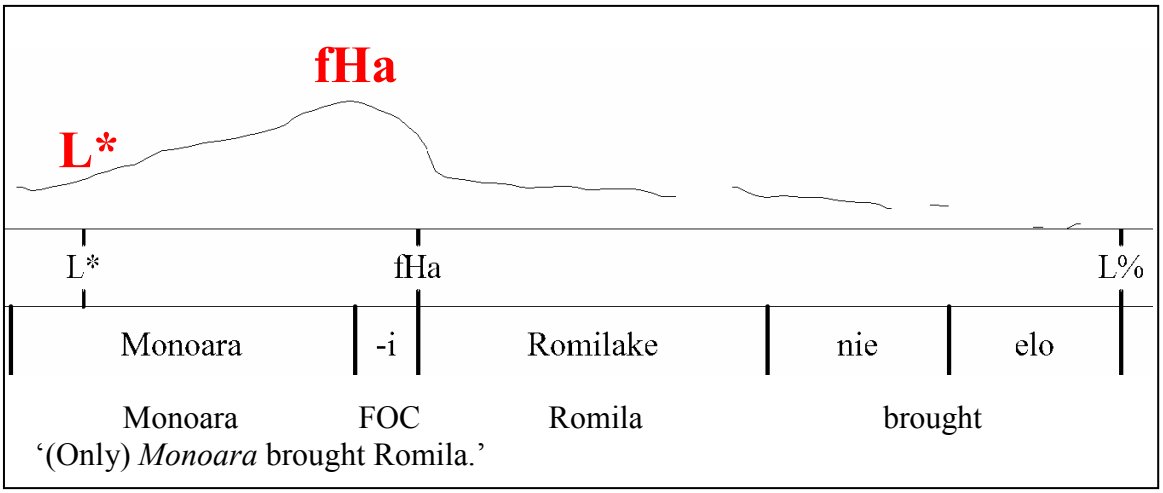


Figure 106. The focus-encliticized AP [monoaɹaj] ‘(only) *Monoara*’ bears a focused smooth rise tonal pattern, composed of a low pitch accent (L*) and extra-high AP boundary tone (fHa), and is followed by post-focal tone deletion (see §12.2). [Az11]

When the focus-encliticized AP is the first AP in the sentence, it may not be very clear that it bears a focused smooth rise (L*…fHa). When it occurs sentence-medially, however, the raised high AP boundary tone (fHa) is hard to miss. Note in Figure 107 below how the AP boundary tone on the focus-encliticized AP [ɹomilakej] ‘(only) *Romila*-ACC’ defies Ha downtrend (*i.e.* all else equal, the pitch of Ha will be lower than that of the preceding Ha) by reaching an equal or higher pitch value than the preceding AP boundary tone on [monoaɹa] ‘Monoara’.

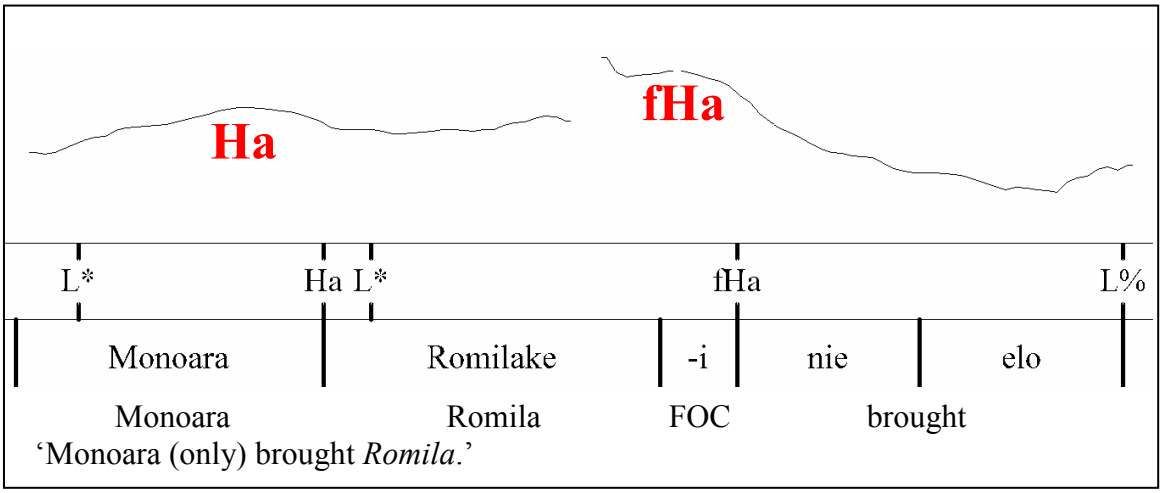


Figure 107. The focus-encliticized AP [ɹomilakej] '(only) *Romila*-ACC' bears a focused smooth rise tonal pattern, composed of a low pitch accent (L*) and extra-high AP boundary tone (fHa), and is followed by post-focal tone deletion (see §12.2)**.** [Re12]

Considering the regularity with which Ha downtrend applies across APs of equivalent length, the violation of the downtrend pattern is presumably the most salient cues for the raised high AP boundary tone (fHa). Compare the following two pitch tracks: the first (Figure 108) illustrates the default non-focused pattern, with two APs bearing smooth rise patterns (L*…Ha) whose high AP boundary tones (Ha) follow Ha downtrend, while the second (Figure 109) illustrates the focus-encliticized pattern, with the focused smooth rise (L*…fHa) defying Ha downtrend by reaching a higher pitch value than that of the high AP boundary tone (Ha) preceding it.

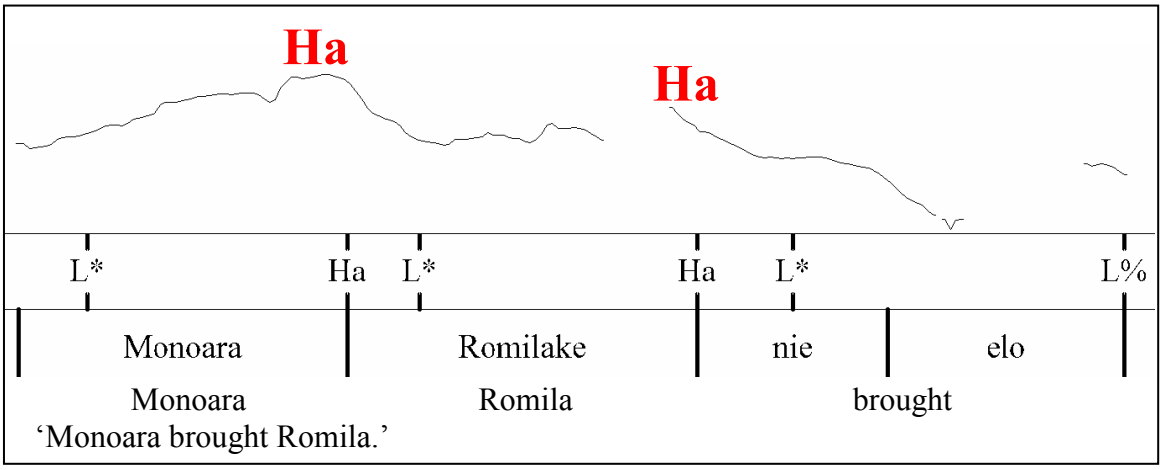


Figure 108. Without focus, Ha downtrend requires that high AP boundary tones (Ha) progressively lower in pitch. Here, the F_0 maximum of [ɹomilake] 'Romila-ACC' is 245Hz, while that of [monoaɹa] 'Monoara' is 299Hz. [Fa01]

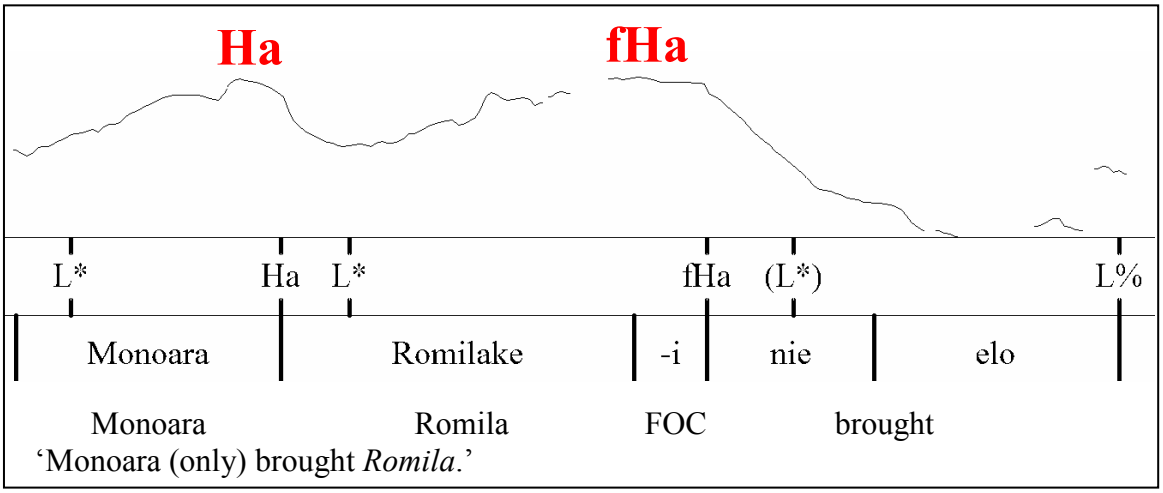


Figure 109. Under encliticized-focus, the pitch range of the high AP boundary tone (Ha) following [ɹomilakej] '(only) *Romila*-ACC' is raised, violating Ha downtrend. Its F_0 maximum is 297Hz, while that of [monoaɹa] 'Monoara' is 295Hz. It is unclear whether the following AP [nie elo] 'brought' bears a low pitch accent (L*). [Fa12]

As mentioned before, the focused smooth rise tonal pattern (L*…fHa) is restricted to APs that bear one of the two focus enclitics (*i.e.* -[i]~[j] 'only', 'indeed', -[o]~[o̯] 'also',

‘even’) on the right edge. What is noteworthy about this distributional fact is that it is not described in previous studies of Bengali prosody. In Hayes & Lahiri’s (1991) study of Kolkata Bengali, the analog of the focused smooth rise contour (*i.e.* their L*…H_P) occurs on focused elements that do not bear focus enclitics; the study does not include descriptions of focus enclitics. Lahiri & Fitzpatrick-Cole’s (1999) study of focus in Kolkata Bengali, however, explicitly discusses focus enclitics, and is the first to differentiate focused elements that bear focus enclitics from those that do not. In their model, all focused elements bear a low pitch accent (L*) on the stressed syllable, but the alignment of the high tone is sensitive to the existence of enclitics. Focused elements without enclitics bear a high boundary tone (H_P in their model) at the right edge of the focus domain (resembling the current study’s focused smooth rise L*…fHa) while focused elements with enclitics bear a high pitch accent (H*) on the focus enclitic (L*…H*). See Chapter Two for a more detailed summary of their findings.

Note that due to the fact that both focus enclitics and AP boundary tones are defined as typically associating with the right edges of words or small groups of words, the raised high AP boundary tone (fHa) is always aligned to the focus enclitic. This association is not just incidental; when the focus enclitic occurs in a location that would not be expected to bear an AP boundary tone, its presence in fact demands that an AP boundary tone be aligned with it nonetheless. For example, the compound verb [meɹe felːo] ‘killed-3RD’ (*lit.* beaten dropped) is normally parsed as one AP, as in Figure 110 below.

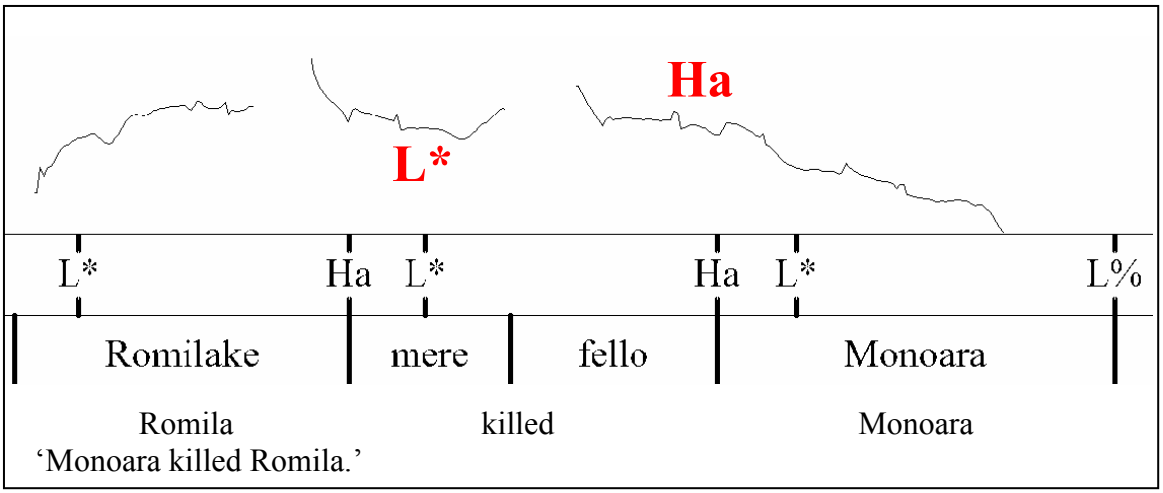


Figure 110. The complex verb [meɹe felːo] 'killed' in this OVS sentence is parsed as one AP, bearing a smooth rise contour (L*…Ha). [FoFSB1]

The complex verb is also parsed as one AP when a focus enclitic attaches to its right edge; for example, [meɹe felːoj] '(indeed) *killed*' is parsed as one AP bearing the focused smooth rise pattern (L*…fHa) in Figure 111 below.

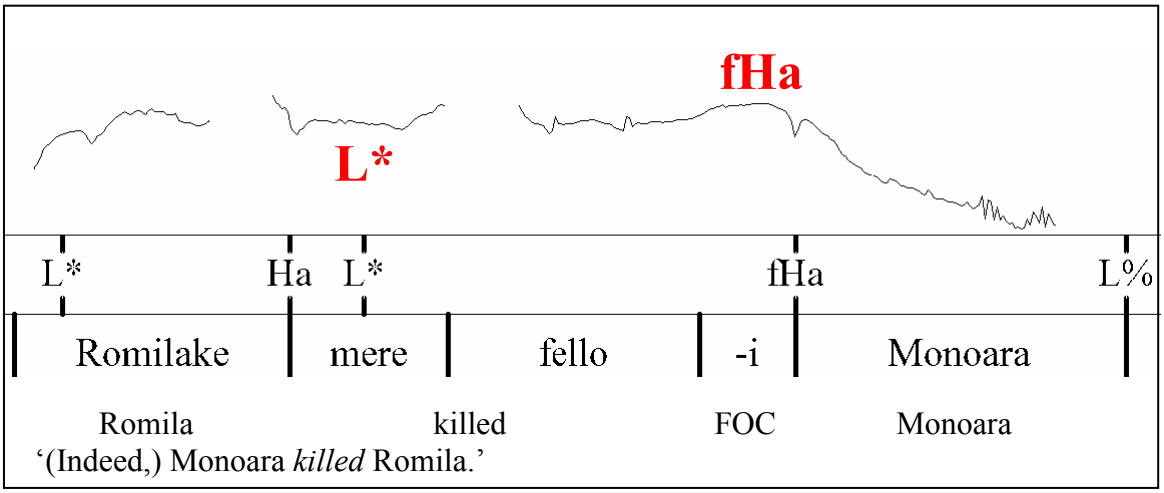


Figure 111. The focus-encliticized complex verb [meɹe felːoj] 'killed' in this OVS sentence is parsed as one AP, bearing a focused smooth rise contour (L*…fHa). [FoFSB5]

One peculiar trait of focus enclitics is the ability to attach to the end of virtually any orthographic word.[77] In fact, the more natural place to attach a focus enclitic to a complex verb (*e.g.* [meɹe felːo] 'killed') is on the right edge of the content word (*e.g.* [meɹe] 'beaten'), instead of on the right edge of the auxiliary verb (*e.g.* [felːo] *lit.* 'dropped', used to signify actions taken to completion). Such examples (*e.g.* [meɹej felːo] 'just *killed*') can have a slightly different shade of meaning than examples in which the focus enclitic attaches to the auxiliary (*e.g.* [meɹe felːoj] 'indeed *killed'*, 'killed nonetheless'). In such cases, one might expect the focused high AP boundary tone (fHa) to be disentangled from the focus enclitic. As discovered in Lahiri & Fitzpatrick-Cole (1999), however, the high tone appears to remain attached the focus enclitic, as in Figure 112 below.

[77] Indeed, Lahiri & Fitzpatrick-Cole (1999) even show examples where the focus enclitic attaches to the end of a phonological word, within an orthographic word (*e.g.* [meɹe-o-tɕʰe] 'has also *beaten'*, from [meɹetɕʰe] 'has beaten'; [bʱeŋe-o-tɕʰi] 'I have also *broken'*, from [bʱeŋetɕʰi] 'I have broken'). These examples were unfortunately not considered grammatical by many speakers in the current study, suggesting a dialect difference.

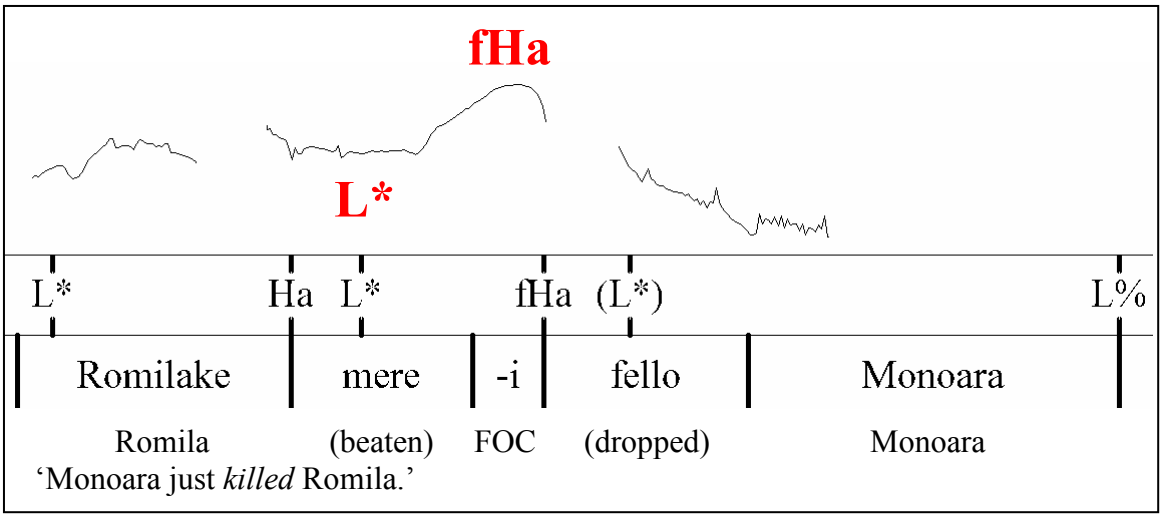


Figure 112. The focus-encliticized complex verb [meɹej felːo] 'just *killed'* in this OVS sentence bears a focused high AP boundary tone (fHa) between the two orthographic words, each of which bears a low pitch accent (L*). [FoFSB4]

Examples like the sentence shown in Figure 112 above seem to suggest that complex verbs such as [meɹe felːo] 'killed' are parsed into two APs when a focus enclitic intervenes between the words. But this is not the only way to interpret the phenomenon; the fact that the focused high tone and focus enclitic seem to be in sync has been interpreted in a number of ways in the past, from Selkirk's (2006) proposal that the high tone is in fact a tonal morpheme that attaches to focus enclitics wherever they are, to Lahiri & Fitzpatrick-Cole's (1999) proposal that focus enclitics underlyingly bear lexically-specified pitch accents that they carry around wherever they attach. Just as Selkirk's (2006) and Lahiri & Fitzpatrick-Cole's (1999) proposals come about naturally due to the construction of their respective models[78], the current study chooses the AP

[78] Lahiri & Fitzpatrick-Cole's (1999) reasoning for positing a lexically-specified high pitch accent (H*) is that the analog of the current model's high AP boundary tone (Ha) is a phonological phrase boundary tone (H_P), which can only occur at the edges of phonological phrases; the boundary between two parts of a

boundary tone label for this high tone based on a number of factors. Firstly, the current intonational phonological model assumes that pitch accents are attached to stressed syllables; the syllables hosting the focus enclitics, however, are not phonologically stressed (see Chapter Two, §11.1)[79]—a fact that Lahiri & Fitzpatrick-Cole (1999) also find troubling. Secondly, data from both Bangladeshi Standard Bengali and Eastern Bengali suggest that material immediately following the high tone must bear phonological stress, even supporting a pitch accent at times. Note in Figure 112 above that the word [felːo] bears a weak low pitch accent (L*) when a focus enclitic precedes it, as opposed to the case in Figure 111 where the word [felːo] does not bear a pitch accent and it is not preceded by a focus enclitic. It appears that an auxiliary verb's ability to bear a pitch accent dependent on whether or not a focus enclitic attracts a boundary tone before it. Of course, the reader may wonder how one can be sure that the word [felːo] in fact bears a low pitch accent (L*), when I have already described the pitch accent as weak. To confirm my proposal, I provide examples from Eastern Bengali.

The Bangladeshi Standard complex verb [meɹe felːo] 'killed' can be translated into Eastern Bengali in two ways: [majɹa falajlo] and [majɹalajlo]. The first version (*i.e.* [majɹa falajlo]) is made up of two full words, thus containing two stressed syllables total. This version can therefore maximally bear two pitch accents (*i.e.* on [maj] and [fa]). The

complex verb like [meɹe felːo], however, is judged on segmental grounds to not qualify as a phonological phrase boundary.

[79] They are often, however, realized with what sounds like phonetic stress. Unfortunately, as I did not control sentences with focus enclitics for factors such as vowel quality, I cannot make accurate phonetic assessments of this putative phonetic stress.

contracted version (*i.e.* [majɹalajlo]), however, acts as one word, bearing no more than one pitch accent (*i.e.* on [maj]). The two forms are largely interchangeable, unless one wants to insert a focus enclitic between the two parts of the complex verb; in this situation, only the [majɹa falajlo] version is grammatical, giving [majɹaj falajlo] '(just) *killed*' and [majɹaa̯ falajlo] 'even *killed*'. The [majɹalajlo] version cannot be encliticized in this way (*i.e.* *[majɹajlajlo], *[majɹaa̯lajlo]). Considering that the main difference between the two forms is the existence of an additional stressed syllable (*i.e.* [ˈfa]), this suggests that placing a focus enclitic after the verb [majɹa] requires that the following word contain a stressed syllable. Indeed, examples like Figure 113 below show that the auxiliary verb needs to contain a stressed syllable, and also bears a (weak) pitch accent[80]—not expected unless the first part of the complex verb (*i.e.* [majɹa]) serves as a complete AP on its own.

[80] The low pitch accent (L*) is weakened due to post-focal tone deletion (see §12.2).

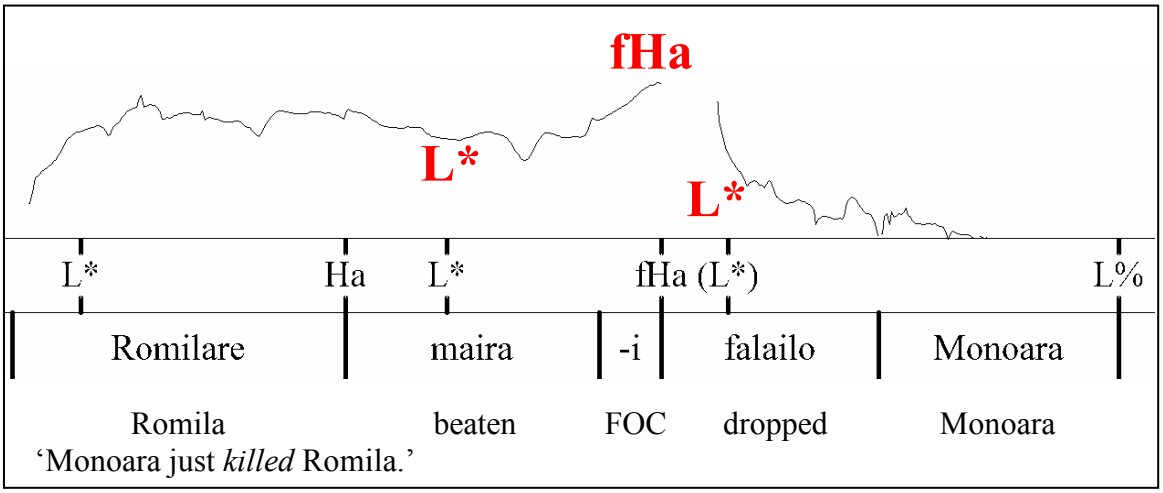


Figure 113. The focus-encliticized verb [majɹaj falajlo] 'just *killed'* in this OVS sentence is parsed as two APs, the first bearing a focused smooth rise contour (L*…fHa) and the second bearing a compressed low pitch accent (L*). This example was produced in Eastern Bengali. [FoFGB4]

Since its main defining feature is the raised high AP boundary tone (fHa), the focused smooth rise tonal pattern (L*…fHa) may at first not seem like a realistic focus realization option for APs in ip-final position, where concurrent boundary tone overriding would likely cause the ip boundary tone (T-) to delete the AP boundary tone (fHa). This would leave behind only the low pitch accent (L*) of the tonal pattern, leaving focused and non-focused APs identical in ip-final position (*i.e.* L*…T-). To resolve this problem, the focused high AP boundary tone (fHa) can either adjoin to the higher prosodic unit's boundary tone or shift away from it. These patterns are described in §12.3.1.

As previously mentioned, the focused smooth rise (L*…fHa) in focus-encliticized constituents can often only be distinguished from the default smooth rise (L*…Ha) by

examining the surrounding phrases: the raised high AP boundary tone (fHa) in its defiance of Ha downtrend, and post-focal words in their tendency to be stripped of their AP-level tones (*i.e.* pitch accents and AP boundary tones). When in ip-final position, the focused smooth rise (L*…fHa) can also be identified by its violation of concurrent boundary tone overriding. Essentially, the focused smooth rise pattern (L*…fHa) can be considered a variant of the default smooth rise pattern (L*…Ha) with the AP boundary tone fused with the focus high tone (fH).

11.2 Sharp rise (L*…Ha → L*+H)

Frequently observed on non-encliticized focused constituents in the experimental data is the sharp rise pattern, composed of a single pitch accent with two tonal targets—a rising pitch accent (L*+H). Unlike other pitch accents, this pitch accent does not appear to project an AP boundary tone (Ta), regardless of whether or not it is ip-final.[81] This is due to the fact that the sharp rise (L*+H) is in fact derived by adjoining the focus high tone (fH) to the low pitch accent (L*) of the smooth rise pattern (L*…Ha), and subsequent post-focal deletion of the high AP boundary tone (Ha). (See §12.2 for a detailed discussion of this post-focal phenomenon.) The sharp rise (L*+H) is primarily used on words bearing corrective focus or wh-answer focus.

[81] Depending on the analysis of post-focal tone deletion, rising pitch accents (L*+H) may in fact be technically ip-final. See §12.2 for a discussion of post-focal tone deletion.

As bitonal pitch accents are otherwise unattested in most previous models of Bengali prosody (attested only in Michaels & Nelson's 2004 study of Eastern Bengali)[82], the goal of this section is to accurately describe the bitonal pitch accent (L*+H) and identify the features that distinguish it from the non-focused smooth rise pattern (L*…H), including gross differences in the contour shape (§11.2.1), the interruption of Ha downtrend (§11.2.2), and the variability in the location of pitch maxima (§11.2.3). These differences are summarized in §11.2.4.

11.2.1 Differences in contour shape

To explore the differences between the pitch contours of non-focused and focused constituents, the subjects of Experiment I (described in Chapter Two) were asked to read 14 sentences of the frame [monoaɹa _____ nie elo.] 'Monoara brought _____.' Seven of the sentences were controlled to elicit neutral focus, by leaving out any clitics, punctuation, or context sentences that could lead the subject into assuming that he or she should apply focus to any of the words. These seven sentences differed only in the direct object, which was controlled for syllable count. Each of the seven sentences was matched with its corrective focus variant, thus totaling 14 sentences. Corrective focus on the direct object was elicited by incorporating the sentence into a mini-dialogue. The subject was asked to pretend that he or she was serving as both Speaker A and Speaker B, and read aloud Speaker A's confirmation question [monoaɹa ɹumuke nie elo, t̪aj na?] 'Monoara

[82] While the downstepped high tone marking (L+H*) "finality", first introduced in Hayes & Lahiri (1991), is bitonal in terms of its formal notation, it does not represent a contour tone, distinguishing it from the rising pitch accent (L*+H) introduced in Michaels & Nelson (2004) and adopted in the current analysis.

brought Rumu, right?', followed by [na, na. monoaɹa _____ nie elo.] 'No, no. Monoara brought _____.' Thus, the first sentence in each mini-dialogue was to be interpreted as Speaker A making an incorrect statement, and Speaker B would then be expected to produce the sentence with corrective focus on the direct object. In this way, examples of the same word in non-focused and corrective focused environments could be elicited from the same speaker in an otherwise identical syntactic position. An example using the word [ɹomilake] 'Romila-ACC' from the sentences [monoaɹa ɹomilake nie elo] 'Monoara brought Romila' (non-focused) and [na, na. monoaɹa *ɹomilake* nie elo.] 'No, no. Monoara brought *Romila*' (corrective focused) is provided below in Figure 114.

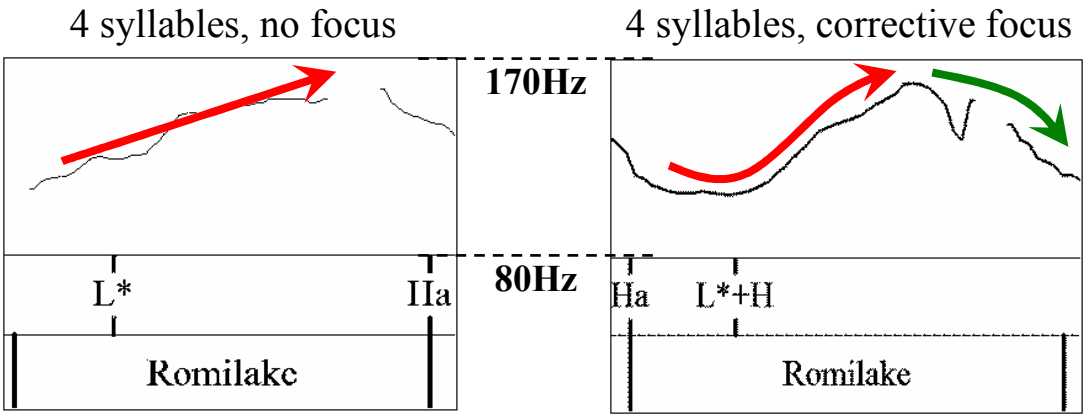


Figure 114. The word [ɹomilake] 'Romila-ACC' produced by the same speaker in the carrier [monoaɹa _____ nie elo] 'Monoara brought _____', with the first instance bearing no focus (F_0 max: 145Hz on [e]), and the second instance bearing corrective focus (F_0 max: 156Hz during [la]). [Ro01], [Ro23][83]

As can be seen in the recordings in Figure 114, both non-focused and corrective focused instances of the words [ɹomilake] 'Romila-ACC' bear rising pitch contours. As

[83] The falling pitch during the later part of the focused word [ɹomilake] 'Romila-ACC' is interpolation towards a following low IP boundary tone (L%).

introduced in Chapter Three, the first recording bears a smooth rise AP tonal pattern (L*…Ha) and the second bears a sharp rise (L*+H). Both recordings begin with low pitch, and as they are produced by the same speaker in the same syntactic position performing the same experimental task, we can also note that the two productions of each word begin with approximately the same F_0 level. The major difference between the two productions, however, begins immediately after the F_0 minimum; the pitch in the non-focused word on the left rises gently towards the right edge of the AP, while the pitch in the focused word on the right rises quite sharply during the initial syllable, and plateaus before the word's right edge, even allowing for a significant length of falling pitch at the end of the word.

Two possible hypotheses could arise from this observation: (1) corrective focused constituents, although underlyingly made up of the same tones as non-focused constituents, display a sharper rise from the low pitch accent (L*) towards an early realization of the high AP boundary tone (Ha), or (2) the difference in pitch contour shape is due to a difference in pitch accent type. Michaels & Nelson (2004) identify and describe the different rising contour patterns of focused and non-focused constituents, and conclude that focused constituents bear a rising pitch accent (L*+H) instead of the default (*i.e.* non-focused) low pitch accent (L*) in Eastern Bengali. The current study adopts the analysis proposed in Michaels & Nelson (2004), having weighed additional evidence provided in subsequent sections; as a preview, I reproduce Figure 114 below in Figure 115, labeled with the low pitch accent on the non-focused version of the word

[ɹomilake] 'Romila-ACC', and with Michaels & Nelson's (2004) bitonal pitch accent (L*+H) on the focused version.

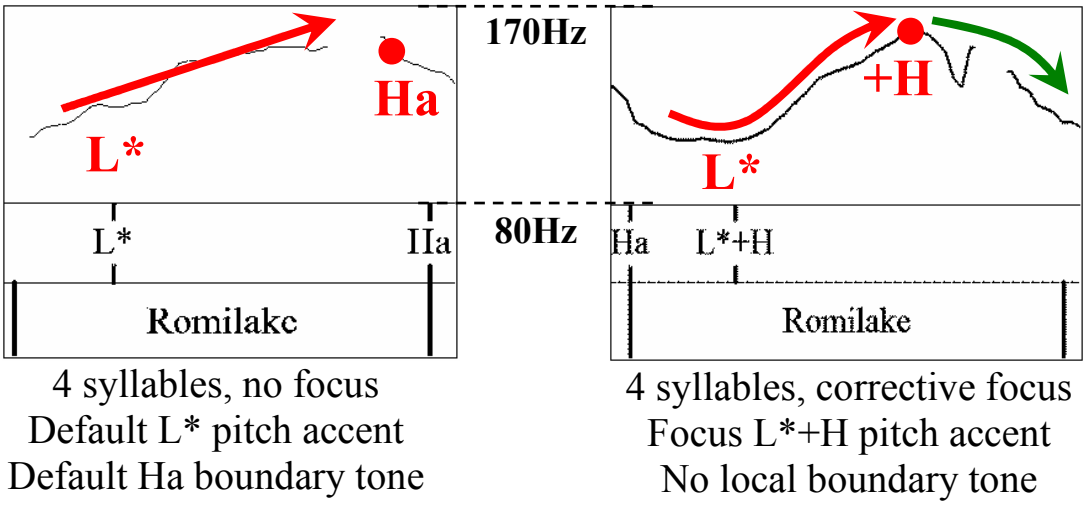


Figure 115. The word [ɹomilake] 'Romila-ACC' produced by the same speaker in the same syntactic position, with the first instance bearing neutral focus, and the second instance bearing corrective focus. [Ro01], [Ro23]

11.2.2 Interruption of Ha downtrend

Like the focused smooth rise tonal pattern (L*…fHa), the sharp rise focus tonal pattern (L*+H) can be distinguished from default non-focused APs bearing the smooth rise pattern (L*…Ha) in its defiance of Ha downtrend. The F_0 maximum on a focused constituent bearing the sharp rise (L*+H) exceeds the pitch of the preceding high AP boundary tone (Ha), thus serving to highlight the focused constituent as the most important AP in the ip. Compare the non-focused Ha downtrend pattern in Figure 116 with the sharp rise pattern of corrective focus in Figure 117 below.

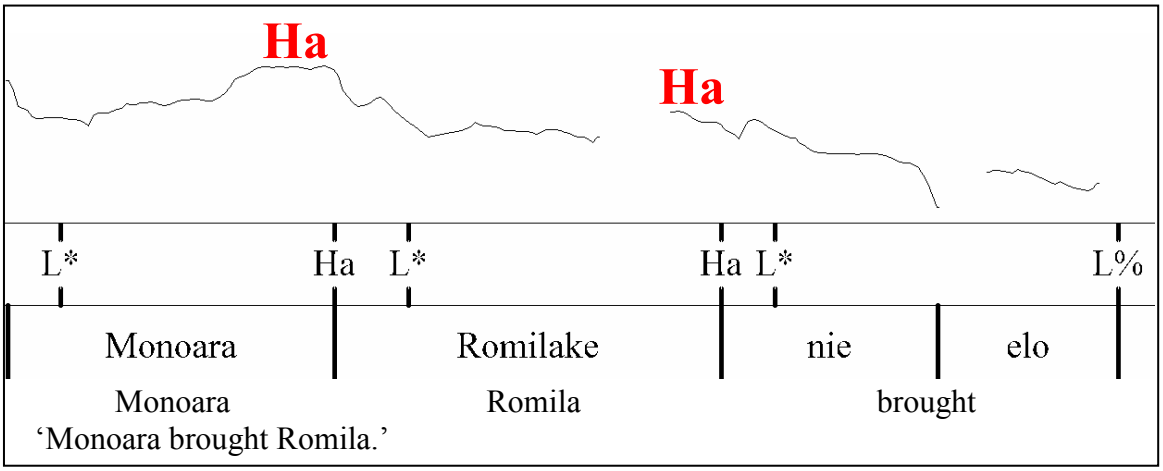


Figure 116. The pitch of the high AP boundary tone (Ha) of the second smooth rise AP tonal pattern (L*…Ha) reaches a lower pitch than the first AP, following Ha downtrend. The F_0 values of the high AP boundary tones (Ha) from left to right are 245Hz and 203Hz. [Da01]

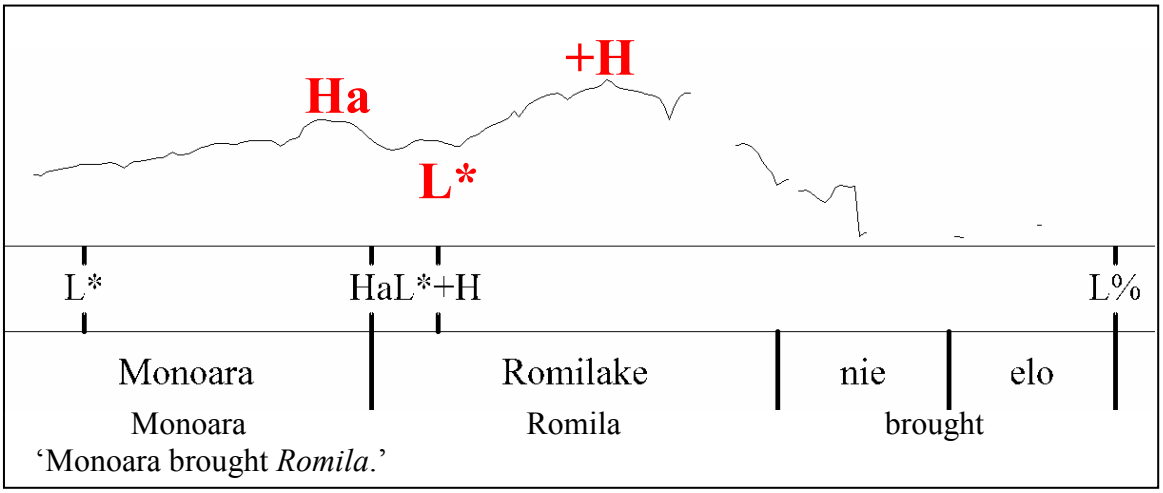


Figure 117. The rising pitch accent's (L*+H) H target during the corrective focused AP [ɹomilake] 'Romila-ACC' reaches a higher pitch than the previous high AP boundary tone (Ha), unlike the expected situation if the same AP were to bear a high AP boundary tone (Ha). The F_0 values of the H tones from left to right are 214Hz and 250Hz.[84] [Da23]

[84] The reader may notice that the high AP boundary tone (Ha) of [monoaɹa] 'Monoara' is lower in Figure 117 (214Hz) than in Figure 116 (245Hz). This is likely due to the overall lower pitch produced in Figure 117; the low pitch accent (L*) of [monoaɹa] is also higher in Figure 116 (197Hz) than in Figure 117

While the focused constituent's sharp rise (L*+H) has been shown to be distinguishable from the non-focused constituent's smooth rise (L*…Ha) in the relative height of the H tone, this does not distinguish the sharp rise (L*+H) from the focused smooth rise (L*…fHa), which also defies the downtrend normally seen across APs of equivalent size. To differentiate the sharp rise (L*+H) from the two kinds of smooth rise (*i.e.* both default and focused), I move on to differences in the variability of pitch maximum location.

11.2.3 Variability in pitch maximum location

Sharp rises (L*+H) can be distinguished from smooth rises (L*…Ha), and by extension, focused smooth rises (L*…fHa) in the location of the pitch maximum. While default non-focused constituents bearing a smooth rise (L*…Ha) and focus-encliticized constituents bearing a focused smooth rise (L*…fHa) consistently project their pitch maxima on the final syllable, focused constituents bearing sharp rises (L*+H) show far more variability in the location of the pitch maximum. For the purposes of this section, I am collapsing the smooth rise (L*…Ha) and focused smooth rise (L*…fHa) into one category, as the feature being measured here (*i.e.* the location of the pitch maximum) does not vary between the two types of smooth rises.

(173Hz). It is unclear if this overall lowering of pitch on [monoaɹa] is related to its immediately pre-focal position.

To compose the smooth rise AP tonal pattern (L*…Ha), default non-focused APs only bear two tones: the low pitch accent (L*) and the high AP boundary tone (Ha). The high AP boundary tone (Ha) is always simultaneously the rightmost point in the AP in terms of duration (or very close to it) and the highest point in terms of pitch, regardless of factors such as word length. Note in Figure 118 below how the disyllabic non-focused AP [make] 'mother-ACC' bears its F_0 maximum on the final syllable [ke], as does the non-focused AP [lina mamike] 'Aunt Lina-ACC', with five syllables. The number of syllables does not affect the fact that the highest pitch in a default non-focused AP is found at its right edge.

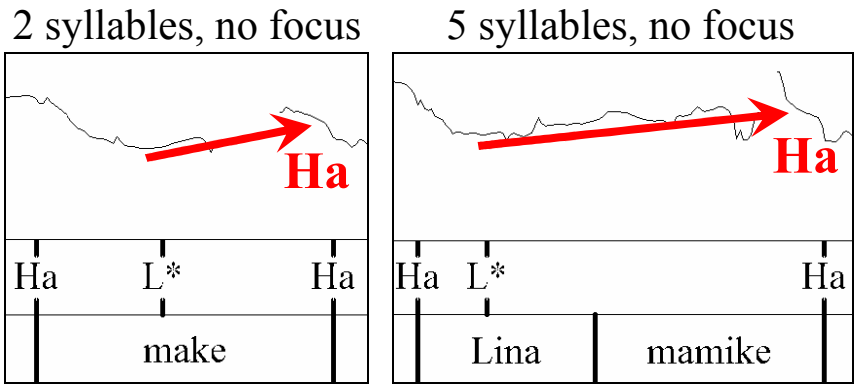


Figure 118. Non-focused constituents [make] 'mother-ACC' and [lina mamike] 'Aunt Lina-ACC' illustrate how the F_0 maximum is consistently realized at the right edge of non-focused constituents, regardless of the number of syllables. Both constituents serve as the object in the frame [monoaɹa _____ nie elo] 'Monoara brought _____'. [To19], [To24]

In focused APs bearing sharp rises (L*+H), however, the F_0 maximum can occur at several different points; it can be word-final, resembling the high AP boundary tone (Ha) of a non-focused constituent (Figure 119, left), but more often, the F_0 maximum occurs

word-medially (Figure 119, right). Naturally, when the F_0 maximum is word-medial, the pitch falls towards the right edge of the word.

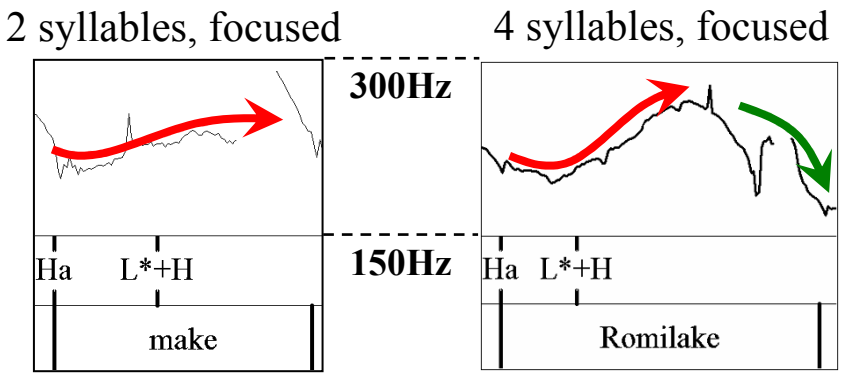


Figure 119. Focused constituents [make] 'mother-ACC' and [ɹomilake] 'Romila-ACC' produced by the same speaker in the carrier phrase [monoaɹa _____ nie elo] 'Monoara brought _____'. [BM20], [BM23]

In a few cases, extremely short focused words bear a rising pitch accent (L*+H) whose pitch maximum is realized on the following word, due to the insufficient duration of the focused word itself, as in Figure 120 below.

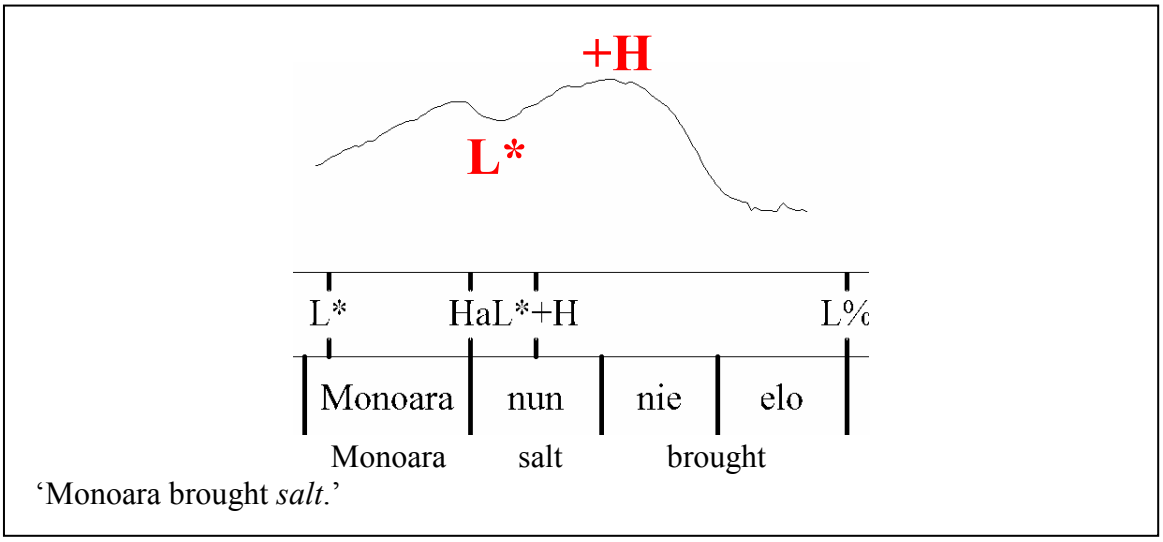
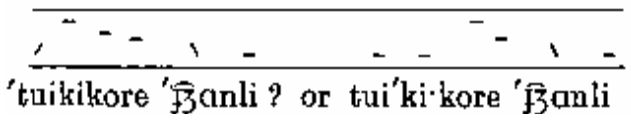


Figure 120. The pitch maximum for the rising pitch accent (L*+H) on the corrective focused word [nun] 'salt' is in fact realized during the initial syllable of the following word [nie] 'taken' due to the short duration of the focused word. The lack of a pitch accent on the complex verb [nie elo] 'brought' is due to post-focal tonal deletion, described in §12.2. [Na18]

The same phenomenon is described in Chatterji's (1921) description of Kolkata Bengali. The subject [t̪uj] 'you' in the first version of [t̪uj ki koɹe dʑanli] 'How did *you* know?' is under focus, but is too short to bear the pitch maximum of the rising pitch accent (L*+H), leaving the pitch maximum realized during the following word. Chatterji's example (§63, p. 21) is reproduced below in (5) with pitch tracks added. Note how the pitch maximum associated with focused [t̪uj] 'you' appears on [ki] 'what' in the first example.



(33) [ˈt̪uj ki koɹe ˈdʑanli] [t̪uj ˈkiː koɹe ˈdʑanli]

you how knew you how knew

‘How did *you* know?’[85] ‘How did you know?’

The focused constituent, bearing a bitonal pitch accent (L*+H), is not required to bear the highest pitch at its right edge; instead, it simply bears its F_0 maximum at some point after the F_0 minimum borne by the pitch-accented syllable. This gives much more freedom for other factors to determine exactly where the F_0 maximum will be realized, instead of artificially restricting it to the edge of the constituent. The variability in the location of the F_0 maximum in focused constituents can be measured directly as a function of relative duration. While both the bitonal pitch accent (L*+H) and smooth rise (L*…Ha) patterns bear rising pitch, the exact timing of the F_0 maximum would be predicted to be different depending on the analysis. If the F_0 maximum in a focused constituent is consistently borne at or near the constituent’s right edge, regardless of the length of the word, it is likely to be a right-anchored tone (*i.e.* boundary tone). However, if the F_0 maximum remains within a fixed distance of the pitch accented syllable (with “fixed distance” either referring to the percentage of the total word duration or the number of syllables intervening between the pitch minimum and maximum), with

[85] The italics to show emphasis and the colored pitch tracks are my addition.

increasing durations of falling pitch between the F_0 maximum and constituent's right edge in increasingly longer words, it is more likely a left-anchored tone (*i.e.* pitch accent). By controlling the length and focus feature of a word, smooth rises (L*…Ha) and sharp rises (L*+H) are predicted to behave differently, illustrated below in Figure 121 and Figure 122.

	[Short constituent]	[Significantly longer constituent]
F_0 rise:	**shorter**	**longer**
F_0 fall:	**~same**	**~same**
% F_0 rise:	**~same**	**~same**
% F_0 fall:	**~same**	**~same**

Figure 121. Predicted properties of smooth rises (L*…Ha).

	[Short constituent]	[Significantly longer constituent]
F_0 rise:	**~same**	**~same**
F_0 fall:	**shorter**	**longer**
% F_0 rise:	**larger %**	**smaller %**
% F_0 fall:	**smaller %**	**larger %**

Figure 122. Predicted properties of sharp rises (L*+H).

The predictions illustrated in Figure 121 and Figure 122 above were tested against data collected in the current study. Using an "incorrect" background sentence [monoaɹa ɹumuke nie elo, ʈaj na?] 'Monoara brought Rumu, right?' and the frame [na, na. monoaɹa _____ nie elo.] 'No, no. Monoara brought _____' to elicit corrective focus on the direct

object, the size of the focused constituent in number of syllables was controlled. Corrective-focused direct objects ranged from two (*e.g.* [make] 'mother-ACC') to five (*e.g.* [lina mamike] 'Aunt Lina-ACC') syllables in length. These sentences were compared with the corresponding non-focused sentences. The data were selected from the eight speakers who produced all eight sentences fluently (*i.e.* four corrective-focused sentences and their four corresponding non-focused versions), without disfluent prosodic breaks. The duration of pitch rise from the F_0 minimum to the F_0 maximum, as a percentage of total word duration, was calculated as % F_0 rise. This was compared to % F_0 fall—the duration of pitch fall from the F_0 maximum to the right edge of the word, as a percentage of total word duration. Table 30 below displays the mean duration of % F_0 rise and % F_0 fall (averaged across all eight speakers) in non-focused words of increasing length.

Number of syllables	**Word**	**Mean duration (as % of total duration)**	
		% F_0 rise L* → Ha	**% F_0 fall Ha → R-edge**
2	[make] 'mother-ACC'	54.3%	10.9%
3	[ninake] 'Nina-ACC'	62.8%	12.1%
4	[ɹomilake] 'Romila-ACC'	71.4%	10.2%
5	[lina mamike] 'Aunt Lina-ACC'	70.2%	12.7%

Table 30. Means of the duration of rising and falling pitch in non-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.* L*).

The data summarized in Table 30 are represented as a line graph in Figure 123, illustrating the effect of word length (measured as the number of syllables) on the durations of % F_0 rise and % F_0 fall. Note that the percentage of word duration that bears rising pitch (*i.e.* % F_0 rise) subtly increases and the percentage of word duration that bears falling pitch near the right edge (*i.e.* % F_0 fall) stays relatively constant as the length of the word increases. This largely matches the predictions of smooth rises (L*…Ha) presented in Figure 121, confirming that the pitch maximum is phonetically realized within the final syllable and thus underlyingly anchored to the right edge of the non-focused word.

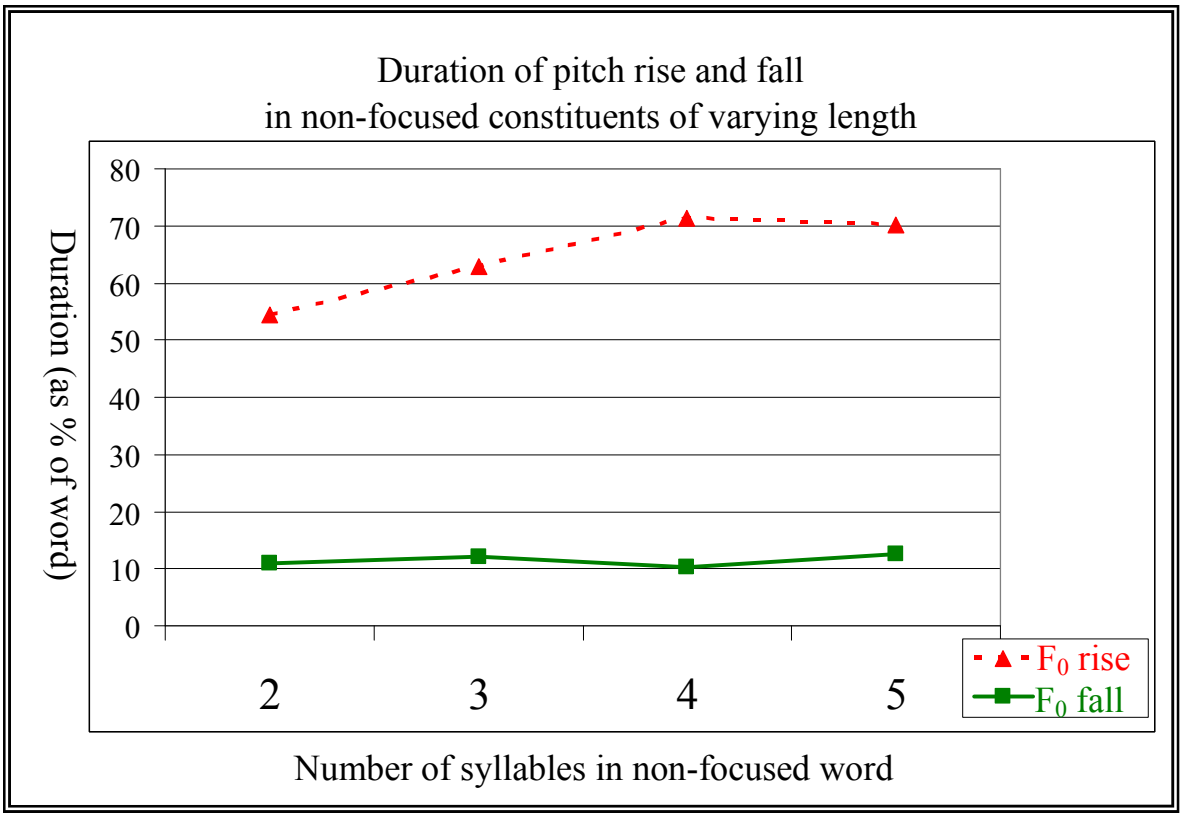


Figure 123. Durations of pitch rise (as a percentage of total word duration) and pitch fall (as a percentage of total word duration) across non-focused constituents of varying length (measured as the number of syllables).

The same measurements (*i.e.* % F_0 rise, % F_0 fall) were made for the corrective-focused words corresponding to the non-focused words in Table 30 above. The corrective-focused measurements are given in Table 31 below.

Number of syllables	Word	Mean duration (as % of total duration)	
		% F_0 rise L* → H	% F_0 fall H → R-edge
2	[make] 'mother-ACC'	58.9%	12.8%
3	[ninake] 'Nina-ACC'	45.2%	34.0%
4	[ɹomilake] 'Romila-ACC'	43.8%	41.0%
5	[lina mamike] 'Aunt Lina-ACC'	35.0%	54.2%

Table 31. Means of the duration of rising and falling pitch in corrective 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 data summarized in Table 31 are represented as a line graph below in Figure 124, illustrating the effect of corrective-focused word length (measured as the number of syllables) on the durations of % F_0 rise and % F_0 fall. Note the very different pattern seen in corrective-focused words in contrast to the pattern seen in non-focused words; as the length of the corrective-focused word increases, % F_0 rise *decreases* and % F_0 *increases*, confirming the predictions for sharp rises (L*+H) introduced in Figure 122 above.

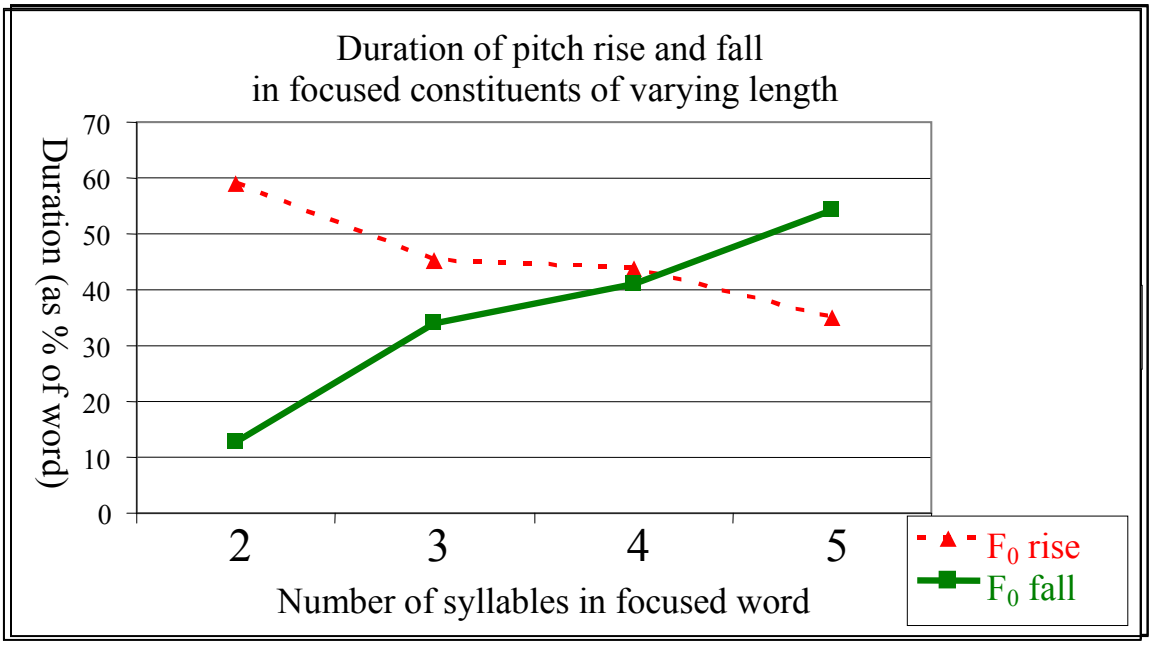


Figure 124. Durations of pitch rise (as a percentage of total word duration) and pitch fall (as a percentage of total word duration) across corrective focused constituents of varying length (measured as the number of syllables).

The data clearly show that as the length of a corrective focused word (*i.e.* an AP bearing a sharp rise, L*+H) increases, the percentage of F_0 rise decreases, and the percentage of F_0 fall increases. While disyllabic words such as [make] 'mother-ACC' bear rising pitch for more than half of the entire word's duration under focus, and bear minimal pitch fall at the word's right edge, longer corrective focus domains such as [lina mamike] 'Aunt Lina-ACC' bear rising pitch for less than 40% of the total word duration, and bear falling pitch through more than 50% of the constituent from the F_0 maximum to the corrective focus domain's right edge. The rise in pitch in a corrective focused word occurs within a fixed distance from the pitch accented syllable—most often in the syllable immediately following the pitch accented syllable, and typically no farther than

the onset of the syllable following that—which would be unexpected unless the H tone were associated with the pitch accent as opposed to a tone on the constituent's right boundary. The pitch maximum in corrective-focused words is consistently realized during the second syllable from the left—or (if applicable) between the second and third syllable—suggesting left anchoring rather than right anchoring.

In addition, as the corrective focused word increases in length, its F_0 maximum occurs farther from the domain's right edge. Unlike non-focused constituents, pitch starts to fall well before the right edge, which could not be easily explained if the H tone were associated to the right edge. This observation supports the distinction between smooth rises (L*…Ha), as seen on default non-focused constituents, and sharp rises (L*+H), as seen on corrective focused constituents, confirming the findings of Michaels & Nelson (2004). Figure 125 below shows the focused words [make] 'mother-ACC' and [ɹomilake] 'Romila-ACC' originally presented in Figure 119, labeled with the bitonal pitch accent (L*+H) and annotated with durational properties to illustrate the realization pattern across corrective focus constituents of different syllable counts.

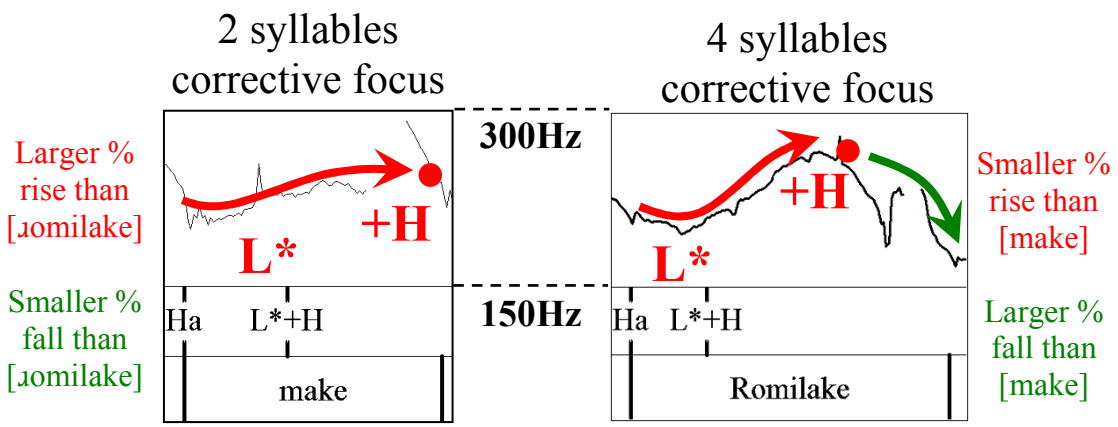


Figure 125. Corrective focused constituents [make] 'mother-ACC' and [ɹomilake] 'Romila-ACC' produced by the same speaker in the same syntactic position. [BM20], [BM23]

11.2.4 Wh-answers

Sharp rises (L*+H) are also seen in the answers to wh-questions. The sentence [ɹumu nepaleɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni] (Rumu Nepal-GEN queen-GEN gardener-DEF.PL.GEN name-DEF.PL mind-GEN keep-INF can-3 PERF.NEG) 'Rumu couldn't remember the names of the gardeners of the queen of Nepal' was produced by 25 speakers both under neutral focus (*i.e.* no special focus) and under wh-answer focus, elicited by preceding the statement with various wh-questions (*e.g.* [ɹumu kon d̪eʃeɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni] 'Which country's queen's gardeners' names could Rumu not remember?'). Subjects' responses were judged by a consultant to ensure that each sentence matched with the preceding wh-question in terms of focus realization. This eliminated the data from four subjects, as their responses were judged as pragmatically incongruous with the preceding wh-question (an English equivalent would be 'When are you going to school?' followed by #'I'm going to *school* at nine.' *vs.* 'I'm going to school

at *nine*.'). The resulting contours on the wh-answer focused words very much resemble those of corrective focused words, in their early F_0 maxima and violation of downtrend, indicating that they too bear sharp rises (L*+H). Examples of wh-answer focus are provided in below.

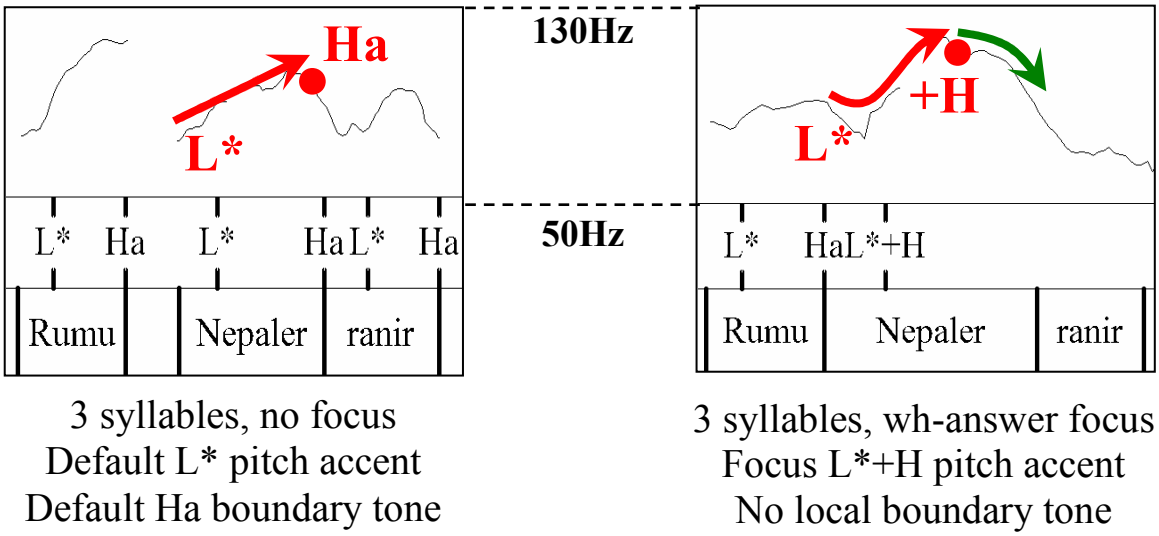


Figure 126. The word [nepaleɹ] 'Nepal-GEN' produced by the same speaker in same syntactic position, with the first instance bearing neutral focus (F_0 max: 101Hz on [e]), and the second instance bearing wh-answer focus (F_0 max: 122Hz on [a]). [Sf50], [Sf38]

Unfortunately, as wh-answers were not controlled for individual word length (although they were controlled for focus domain size; see §12.1.2), the same measurements made on corrective focused words (§11.2.2, §11.2.3) could not be made on wh-answer focused words.

11.2.5 Summary

By observing that the F_0 maximum of a corrective/wh-answer focused constituent does not obey the Ha downtrend seen in non-focused constituents, we can infer that it either bears a special high AP boundary tone (fHa) or that its H target belongs to a different category altogether. Since focus encliticized constituents already make use of the Ha downtrend-violating focused high AP boundary tone (fHa), one may assume that the pitch maximum of wh-answer/corrective focused constituents is also associated to the AP boundary. However, the observation that the pitch maximum during corrective/wh-answer focused constituents is often not realized at a word boundary suggests that the H tone is in fact not associated to a boundary tone at all. In fact, the F_0 maximum of a wh-answer/corrective focused constituent strays further from the word's right boundary as the word increases in length—and instead remains within the first or second syllable following the main stress—suggesting that the H tone is associated to the pitch accent and not to the AP boundary, supporting the bitonal pitch accent (L*+H) hypothesis proposed in Michaels & Nelson (2004).

The bitonal pitch accent (L*+H) itself can be broken down into the low pitch accent (L*) and the underlying focus high tone (fH). The focus high tone (fH) adjoins to the low pitch accent (L*) of the smooth rise AP tonal pattern (L*…Ha). The high AP boundary tone is deleted as it occurs to the right of the realization of the focus high tone (see §12.2). These two processes (*i.e.* adjunction of the focus high tone to the pitch accent, deletion of the AP boundary tone) work together to ensure that the constituent's pitch maximum is

borne near the pitch accent, and not necessarily at the right edge, unlike non-focused and encliticized focused constituents.

The three major phonetic differences across the three types of constituents seen thus far (*i.e.* no focus, focus-encliticized, corrective/wh-answer focus), along with the difference in docking point for the focus high tone (fH) are summarized in Table 32 below.

	Non-focused constituents	**Focus enclitized constituents**	**Corrective/wh-answer focused constituents**
Docking point of fH	*none*	Ha	L*
Pitch contour	Smooth rise		Sharp rise, plateau, fall
Location of F_0 maximum	Consistently AP-final		Not associated with AP-boundary
Downtrend	F_0 maxima get progressively lower	F_0 maximum higher than previous	

Table 32. A summary of three basic differences across non-focused, focus-encliticized, and corrective/wh-answer focused constituents.

11.3 Focused high pitch accent (H*…La → fH*)

Just as the focus high tone (fH) can dock to the low pitch accent (L*) of the smooth rise AP tonal pattern (L*…Ha) when under corrective and wh-answer focus, it can also dock to the high pitch accent (H*) of the smooth fall AP tonal pattern when under surprise focus. Like the focused smooth rise tonal pattern (L*…fHa), it involves raising the pitch maximum of an existing high tone as a phonetic realization of the focus high

tone (fH). Like the sharp rise focus tonal pattern (L*+H), the focused high pitch accent (fH*) is composed of a single pitch accent, and the expected AP boundary tone (in this case, La) is deleted due to post-focal tone deletion.

As mentioned in Chapter Three, the high pitch accent (H*) can be used as the ip-final pitch accent in non-focused phrases in Eastern Bengali dialects, as in Figure 36 below, where the ip-final non-focused verb [d̪ekʰt̪ese] 'they are looking at' bears a high pitch accent (H*).

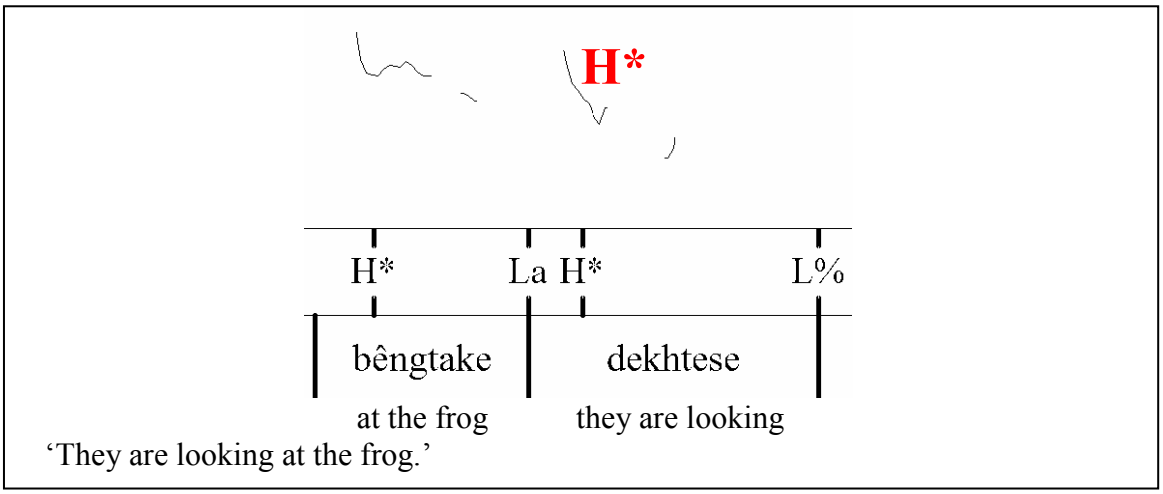


Figure 127. The ip-final non-focused AP [d̪ekʰt̪ese] 'they are looking' bears a high pitch accent (H*). The phrase was produced by a speaker from Comilla District, using his native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. [ReS9]

Note in Figure 36 how the pitch maximum on the ip-final high pitch accent (H*) of the AP [d̪ekʰt̪ese] 'they are looking' is lower than that of the preceding smooth fall pattern (H*…La) during [bɛŋtake] 'at the frog'. This downtrend in high pitch accents (H*)

appears to be similar to the downtrend in high AP boundary tones (Ha). Unfortunately, because long stretches of repeated high pitch accents (H*) are not commonly seen in the current study's corpus of data, it cannot be said with certainty that H* downtrend is as regular as Ha downtrend. Nevertheless, it appears that the interruption of H* downtrend is the phonetic realization of the docking of the focus high tone (fH) on an existing H target. This signals a type of focus specific to surprising or ironic information, in the same way that the interruption of Ha downtrend can signal encliticized, corrective, or wh-answer focus. This type of focus is referred to as surprise focus in the current study. Compare the pitch maxima of the two high pitch accents (H*) in Figure 43 below. Note how the ip-final surprise-focused verb [poɹe dʑatɕʰːe] 'are falling down' bears a much higher pitch value (273Hz) than the preceding APs ([t̪aɹ pɔɹe tɕʰeleta aɹ] 'then the boy and': 189Hz, [kukuɹta ɛkd̪ɔm] 'the dog totally': 185Hz). To represent the fusion of the focus high tone (fH) to the high pitch accent (H*)—identified by the violation of H* downtrend—the "f" diacritic used on the raised high AP boundary tone (fHa) is applied to the raised high pitch accent (fH*).

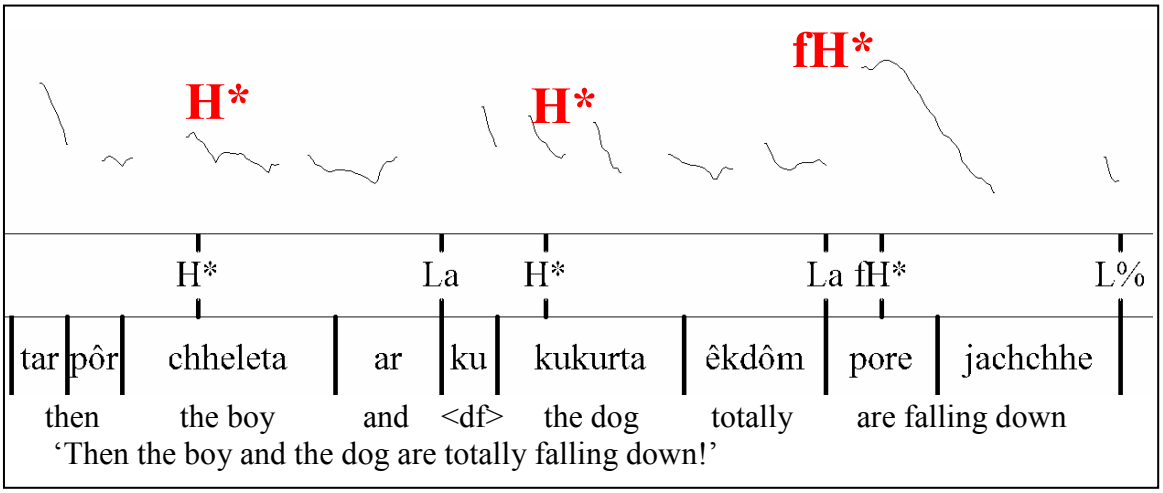


Figure 128. The final AP [poɹe dʑatɕʰːe] 'are falling down' bears a focused high pitch accent (fH*), signaling sudden or unexpected information. The phrase was produced by a speaker from Mymensingh District, using her native dialect (a hybrid of Bangladeshi Standard Bengali and Eastern Bengali) in a recording session of naturalistic speech. [JhS98]

Surprising or unexpected information may not seem like the most canonical focus type; it might be more accurate to label this as "new information". In the literature, this type of focus is often referred to as "broad focus" (see Frota 2000 §1.4.1 for a review). In the current study, this type of focus is structurally very similar to other types of focus (*i.e.* encliticized, wh-answer, corrective) as it involves the use of the focus high tone (fH) and post-focal tone deletion (see §12.2), both of which are associated with the other two types of focus realization (*i.e.* focused smooth rise L*…fHa and sharp rise L*+H). However, due to the low frequency of the focused high pitch accent (fH*), strong claims cannot be made as to its syntactic and prosodic distribution, or phonetic realization, without additional data.

It is equally unclear if the focused high pitch accent (fH*) is exclusively a focus realization option for Eastern dialects of Bengali, or if it can be posited for Bangladeshi Standard Bengali as well. As the sentences from Experiment I (scripted production of Bangladeshi Standard Bengali) were not constructed for elicitation of this focus realization type, only examples from Experiment II (naturalistic production of subjects' native dialects) could be found. Considering the relatively low frequency of high pitch accents (H*) and smooth fall AP tonal patterns (H*…La) in the scripted data of Bangladeshi Standard Bengali, it would not be surprising if the focused high pitch accent (fH*) were equally infrequent. Additional data will be necessary to determine its status in Bangladeshi Standard Bengali.

11.4 Summary

The current study reveals three types of focus realizations, distinguished by the docking point of the focus high tone (fH): focused smooth rise (L*…fHa) with the high AP boundary tone (Ha) as the docking point for the focus high tone (fH), sharp rise (L*+H) with the low pitch accent (L*) as the docking point, and focused high pitch accent (fH*), with the high pitch accent (H*) as the docking point. When the focus high tone (fH) docks to an H target (*i.e.* Ha or H*), it is realized as an increase in the pitch of the tone (*i.e.* identified by a lapse in downtrend). When it docks to an L target (*i.e.* L*) it adds a trailing H component to the pitch accent, creating a bitonal/rising pitch accent (L*+H). All three H tones in the focused patterns reach higher pitch levels than the preceding AP-level H tone (either a high pitch accent H* or a high AP boundary tone Ha).

Additionally, all three patterns also have the tendency to trigger tone deletion following the realization of the focus high tone (fH), as explained further in §12.2.

The three patterns are largely restricted in their distribution, in that the focused smooth rise (L*…fHa) is used on constituents whose focused status is elicited by the use of focus clitics (*i.e.* -[i]~[j] 'only', 'indeed' and -[o]~[ɔ̰] 'also', 'even') and the sharp rise (L*+H) is used on answers to wh-questions and corrections to previous statements. The focused high pitch accent (fH*) is used on particularly sudden, surprising, or unexpected information in Eastern dialects of Bengali, and possibly also in Bangladeshi Standard Bengali. The three major focused tonal patterns are summarized in Table 33 below, alongside their distribution (in terms of focus type) and their non-focused versions.

Focused	**Non-focused**
L*…fHa *encliticized*	L*…Ha
L*+H *corrective, wh-answer*	L*…Ha
fH* *surprise*	H*…La

Table 33. The three focus tonal patterns, alongside the type of focus associated with each pattern, and the non-focused counterpart of each pattern.

Having described their tonal properties, I now move on to discuss the phrasal properties of focused constituents in §12.

12 PHRASING

In addition to bearing one of the three focus realization patterns (*i.e.* L*…fHa, L*+H, fH*), focused constituents can often be distinguished by their phrasing patterns. Focused constituents show a greater tendency to group longer strings of syllables into a single prosodic unit within the focus domain when the domain contains exactly one noun phrase (NP), while complex-NP domains show two different kinds of realizations (*i.e.* representative and repeated). Post-focal APs are characterized by either phonetically reduced or deleted tones (*i.e.* post-focal tone deletion). Along with the distinctive tones associated with the three focus realization patterns, this special phrasing presumably helps focused constituents sound more prominent. In this section, I describe the non-tonal prosodic characteristics of the focus domain (§12.1) as well as of post-focal material (§12.2), using experimental data illustrating variation in focus domain size (§12.1.1), realization type (§12.1.2), and acoustic measurements of both focal and post-focal material (§12.2). I also describe the different boundary tone types that can be borne by IPs containing a focused constituent, and discuss the interaction of various pitch accents and boundary tones with regard to previous claims concerning the role of phonological constraints on tonal sequences in Bengali (§12.3).

12.1 Focus domain

In the current study, focused constituents ranged in size from as small as a single word to as large as a string of APs. Regardless of size, however, focus domains bear

particular phrasal characteristics and tendencies that distinguish them from non-focused constituents. I begin in §12.1.1 by describing the phrasing of material within focus domains composed of one NP, and continue in §12.1.2 by describing the variability in phrasing within focus domains composed of complex NPs.

12.1.1 Simple-NP domain: Grouping

A single Bengali word is not typically assigned more than one pitch accent.[86] Strings of multiple words, however, can be variably assigned one or more pitch accents, as mentioned in the beginning of Chapter Three. When a string of multiple words composes a simple NP, the individual words are free to group together bearing a single pitch accent, or stay in separate APs and bear one pitch accent each, maximally bearing as many pitch accents as there are words.[87] As multiple-word NPs increase in length, the likelihood that the speaker will assign it multiple pitch accents increases. Out of the nine speakers judged by my primary consultant to have correctly produced all six sentences involving multiple-word simple NPs, four chose to produce the five-syllable two-word non-focused string [lina mamike] 'Aunt Lina-ACC' with two pitch accents, one on the initial syllable of each word. The number of speakers producing two pitch accents increased to five as the number of syllables increased to six [ɹomila mamike] 'Aunt Romila-ACC' (also two words). When the number of syllables reached seven [ɹomila mamanike] 'Aunt Romila-

[86] It is unknown whether Bengali words can bear additional pitch accents on secondary stressed syllables, as can be done in languages such as English. No such examples were found in the current study.

[87] The term "word" is ambiguous; the relevant use of "word" here may refer to something more like a phonological word (see Lahiri & Fitzpatrick-Cole 1999). For the purposes of the current study, I simply default to the use of orthographic breaks to define word edges, although I acknowledge that this is not likely to be the definition of word used by the native speaker.

DIM-ACC', all but one speaker produced the string with two pitch accents. This effect is summarized in Table 34 below.

Speaker → / # Syllables	Da	Fa	Ba	Az	Na	BM	Sh	Re	To
Five	1	1	**2**	1	1	**2**	**2**	**2**	1
Six	**2**	1	**2**	1	**2**	**2**	**2**	1	1
Seven	**2**	1	**2**	**2**	**2**	**2**	**2**	**2**	**2**

Table 34. Number of pitch accents produced in non-focused simple NPs composed of two orthographic words each, arranged by speaker and number of syllables in the NP.

In Figure 129 and Figure 130 below, the non-focused simple NP [ɹomila mamike] 'Aunt Romila-ACC' is assigned one pitch accent (and thus phrased as one AP) by one speaker and two pitch accents (thus phrased as two APs) by another.

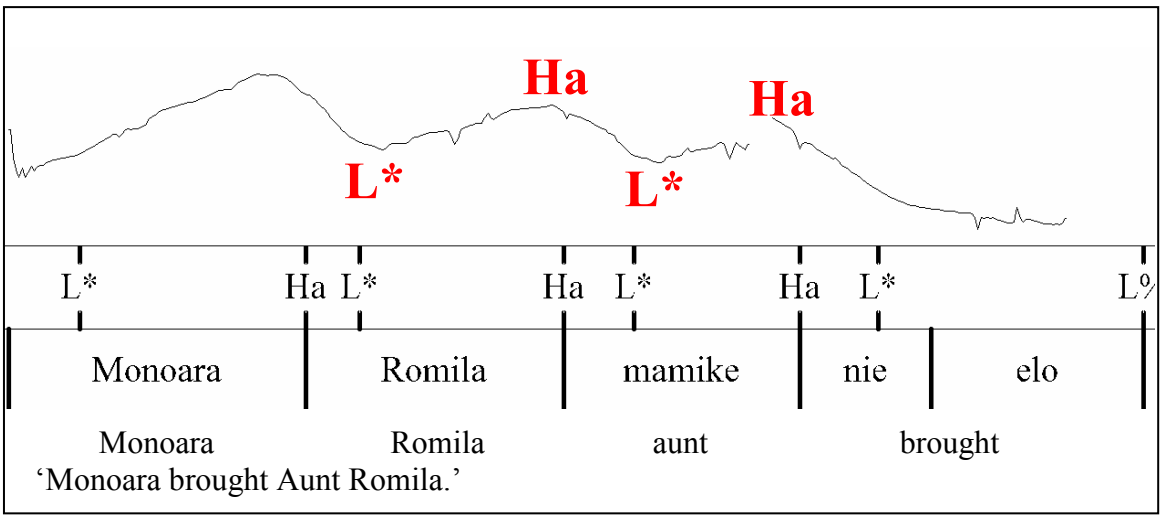


Figure 129. The non-focused simple NP [ɹomila mamike] 'Aunt Romila-ACC' is assigned two pitch accents by this speaker. [Na26]

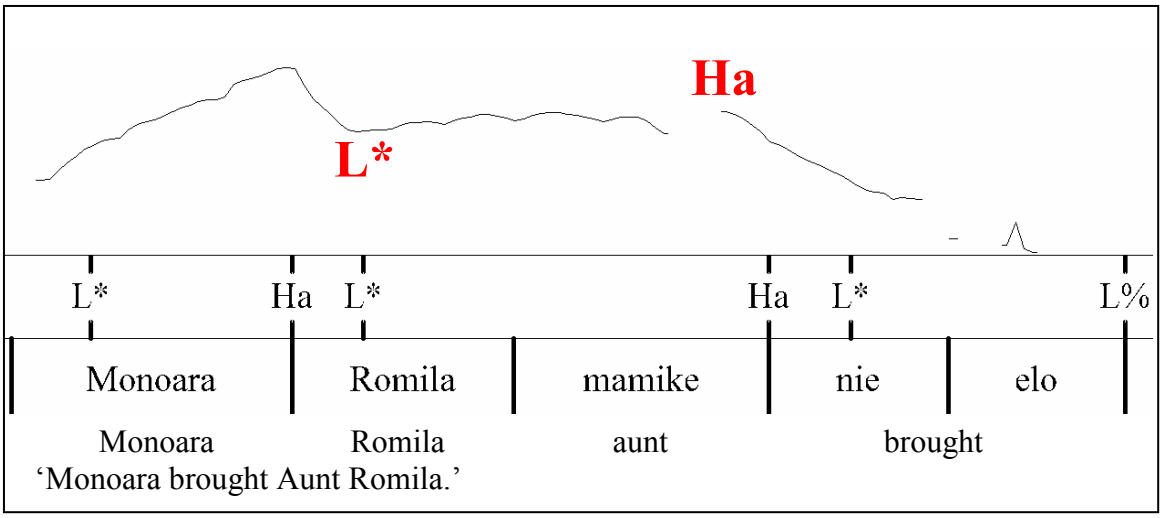


Figure 130. The non-focused simple NP [ɹomila mamike] 'Aunt Romila-ACC' is assigned only one pitch accent by this speaker. [Az27]

When the multiple-word simple NP is placed under corrective focus, however, speakers prefer not to assign the NP multiple pitch accents. Instead, only the leftmost word in the NP bears a pitch accent, leaving the second word unaccented. Speakers overwhelmingly preferred to assign only one pitch accent to corrective-focused NPs composed of two words containing five or six syllables total. Only one speaker (BM) produced two pitch accents on the five-syllable NP [lina mamike] 'Aunt Lina-ACC', and all speakers produced only one pitch accent on the six-syllable NP [ɹomila mamike] 'Aunt Romila-ACC'. Only when the syllable count was raised to seven in [ɹomila mamanike] 'Aunt Romila-DIM-ACC' did speakers prefer to assign two pitch accents. This effect is summarized in Table 35 below.

Speaker → / # Syllables	Da	Fa	Ba	Az	Na	BM	Sh	Re	To
Five	1	1	1	1	1	**2**	1	1	1
Six	1	1	1	1	1	1	1	1	1
Seven	1	1	**2**	**2**	1	**2**	**2**	**2**	1

Table 35. Number of pitch accents produced in corrective focused simple NPs composed of more than one word, arranged by speaker and number of syllables in the NP.

When the two speakers from Figure 129 and Figure 130 above produced the same two-word NP [ɹomila mamike] 'Aunt Romila-ACC' under corrective focus, they did not assign it two pitch accents. Instead, they assigned a pitch accent only to the leftmost word [ɹomila] 'Romila', as shown in Figure 131 and Figure 132 below. For clarity, a box is drawn around all multi-word focus domains in this chapter.

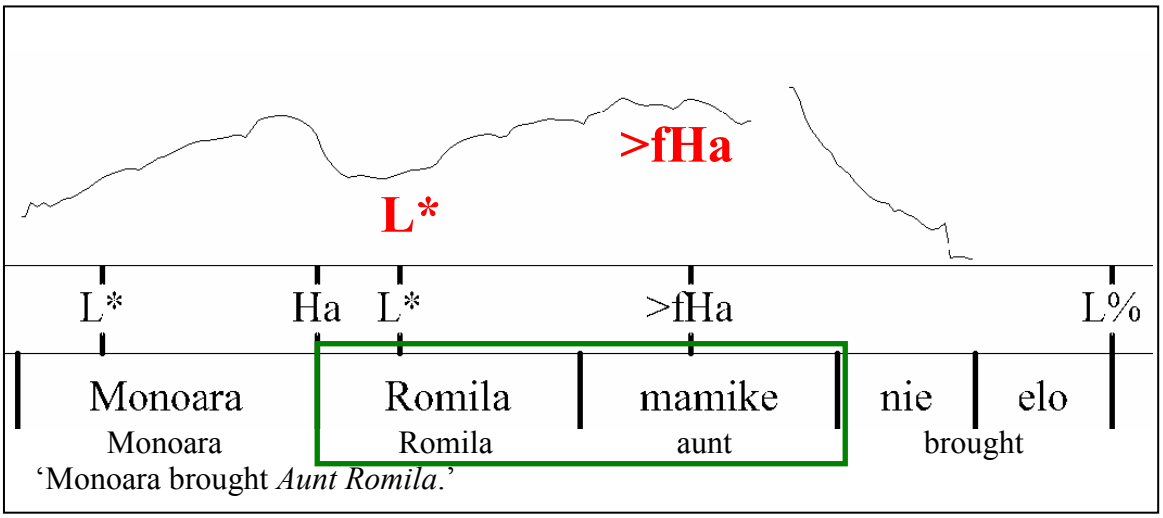


Figure 131. The corrective-focused simple NP [ɹomila mamike] 'Aunt Romila-ACC' is assigned only one pitch accent by this speaker. Although it is not perfectly clear why the speaker produced this NP with what appears to be a focused smooth rise (L*…fHa) with an early realization of the boundary tone (>fHa), it is clear that there is only one pitch accent in the focus domain. [Az27]

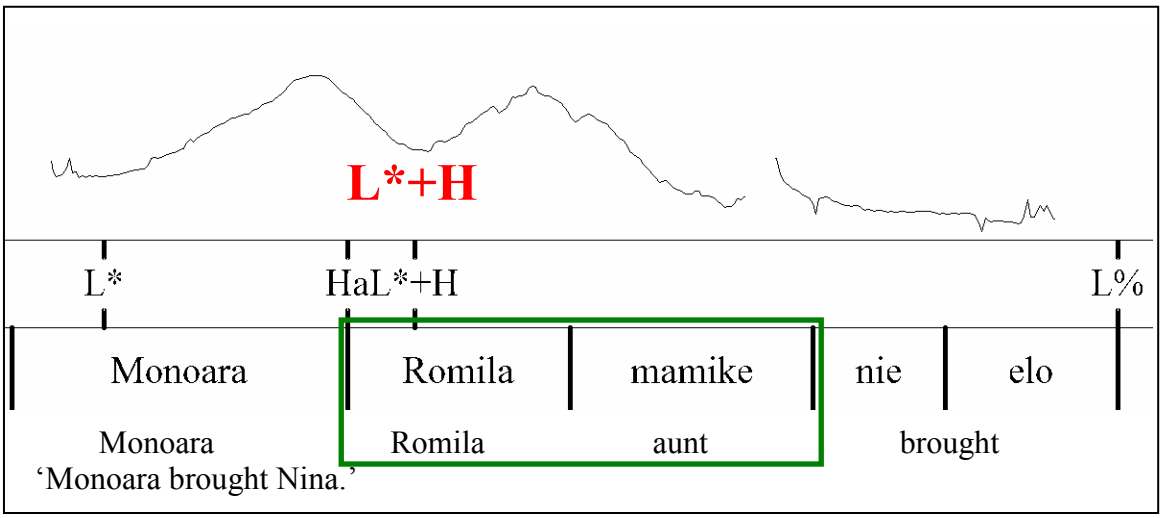
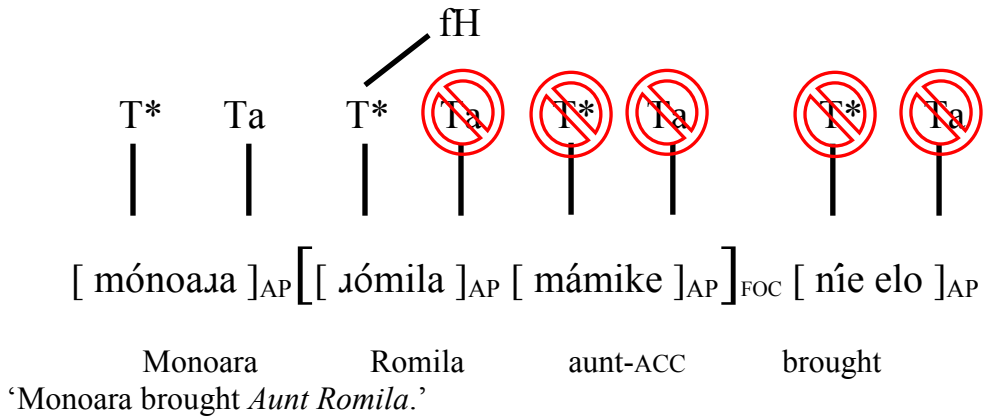


Figure 132. The corrective-focused simple NP [ɹomila mamike] 'Aunt Romila-ACC' is also assigned only one pitch accent by this speaker. [Na27]

The fact that multiple-word simple NPs prefer to be assigned only one pitch accent is presumably related to post-focal tone deletion, which compresses or entirely deletes pitch accents following the realization of the focus high tone (fH). In this sense, the two words in the simple NP under corrective focus could theoretically be assigned one or two pitch accents, the first of which serves as the docking site for the focus high tone (fH). If assigned two pitch accents (as would be expected in non-focused constituents), however, the second pitch accent would be weakened or deleted entirely as it would occur to the right of the realization of the focus high tone (fH), as schematized in Figure 7 below (crossed circles represent tone deletion). See §12.2 below for a more detailed discussion of post-focal tone deletion.



'Monoara brought *Aunt Romila*.'

Figure 133. Schematic illustration of post-focal tone deletion—compression or total deletion of pitch accents and corresponding boundary tones following the realization of the underlying focus high tone (fH), which is seen here adjoined to the low pitch accent (L*) to create the bitonal pitch accent (L*+H) borne on the first word in a multiple-word corrective-focused simple NP.

Examples like the one schematized in Figure 7 suggest that post-focal tone deletion is defined as the compression or deletion of pitch accents and AP boundary tones after the surface realization of the focus high tone (fH), instead of after the entire focus domain. If post-focal tone deletion began after the right edge of the focus domain, the pitch accent on [mamike] 'aunt-ACC' would not be deleted. Studies of Kolkata Standard Bengali—a dialect in which the right edge of the focus domain is also presumably the docking site of the focus high tone (fH)—could not differentiate whether tone deletion occurred to the right of the focus domain or to the right of the focus tone. Bangladeshi Standard Bengali—with its focused high pitch accent (fH*) and bitonal pitch accent (L*+H) realizations of focus—indicates that the string of deleted tones in fact begins at the realization of the focus tone. I will return to this point in the section on post-focal tone deletion (§12.2), after a discussion of complex-NP focus domains in §12.1.2.

12.1.2 Complex-NP domain: Repeated and representative realizations

To examine how speakers apply focus realization to constituents containing complex NPs, 25 subjects were asked to read question-answer pairs based on the sentence [ɹumu nepaleɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni] 'Rumu couldn't remember the names of the gardeners of the queen of Nepal.' Each wh-question elicited wh-answer focus on domains of different sizes; while the wh-question [ɹumu ki mone ɹakʰt̪e paɹe ni?] 'What could Rumu not remember?' elicited wh-answer focus on the complex-NP object [nepaleɹ ɹaniɹ maliḏeɹ namgulo] 'the names of the gardeners of the queen of Nepal', the wh-question [ɹumu kon ḏeʃeɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni?] 'Which country's queen's gardeners' names could Rumu not remember?' only elicited wh-answer focus on one word—[nepaleɹ] 'Nepal's'. This section examines the patterns of focus realization seen in complex-NP focus domains, and thus only the data from speakers whose productions of complex-NP focus domains were judged as grammatical by my primary consultant are considered here. By removing data in which the speaker either applied focus realization to a domain not elicited by the associated wh-question, or simply failed to apply focus realization at all, the number of sentences examined was reduced from 200 to 106.[88]

[88] A few speakers simply read the wh-question and corresponding answer as if they were two unrelated sentences, producing the wh-answer as a default sentence with no focus realization. More commonly, speakers applied wh-answer focus realization to a word outside the intended focus domain. Data including either of these types of mistakes were excluded from the analysis. The wh-answers involving the latter type of mistake (*i.e.* wrong focus) follow the same phonetic patterns described in the current model—just on the wrong word. With a different wh-question, they would be perfectly appropriate.

It was found that in complex-NP focus domains, a speaker may choose from one of two focus realization types: what I call *representative* and *repeated* realizations. The same string can be produced in either realization, and speakers can even vary between two repetitions of the same sentence, suggesting that the two realizations serve as phonological variants of the same semantic/pragmatic representation of focus.

Representative realization involves the application of focus realization (*i.e.* sharp rise L*+H) only to one NP in a complex-NP focus domain. The NP that bears the focus realization is called the *focus representative*, which is always the head of the focused NP—the rightmost NP in the domain, as Bengali is an overwhelmingly head-final language on the surface. All preceding material within the focus domain bears smooth rises (L*…Ha), which are normally associated with non-focused material. Representative realization is the more common of the two realization types.

An example of representative focus realization is provided in Figure 134 below, where the wh-question [ɹumu kad̪eɹ namgulo mone ɹakʰt̪e paɹe ni?] 'Whose names could Rumu not remember?' elicits wh-answer focus on the three-NPs [nepaleɹ ɹaniɹ mali d̪eɹ] 'of the gardeners of the queen of Nepal'. This focus domain bears only one sharp rise (L*+H), located on the focus representative [mali d̪eɹ] 'of the gardeners', while [nepaleɹ] 'of Nepal' and [ɹaniɹ] 'of the queen' bear default smooth rises (L*…Ha).

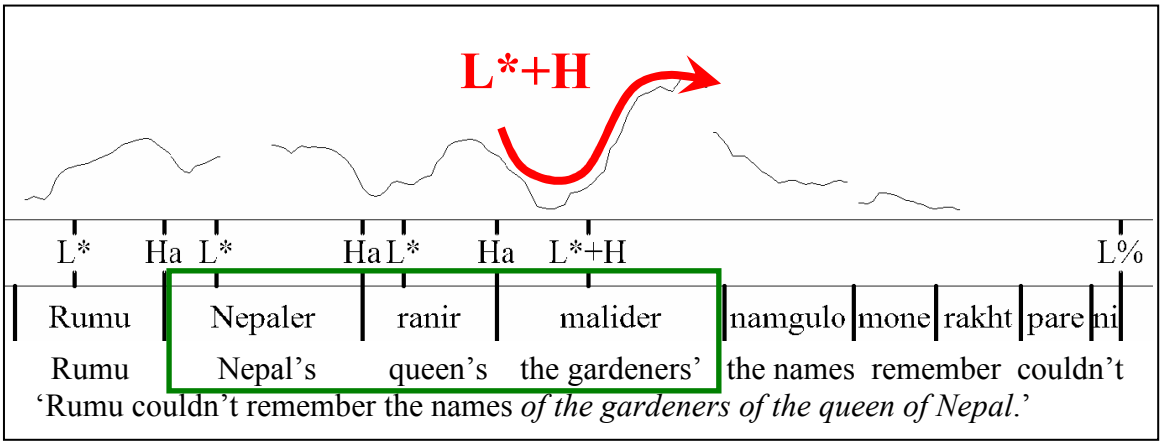


Figure 134. The complex-NP focus domain [nepaleɹ ɹaniɹ maliḍeɹ] 'of the gardeners of the queen of Nepal' bears only one sharp rise (L*+H),[89] located on the focus representative [maliḍeɹ] 'of the gardeners'. [Sf40]

Shortening the focus domain from Figure 134 by one word moves the rising pitch accent (L*+H) to the left in Figure 135 below, with [ɹaniɹ] 'of the queen' serving as the focus representative.

[89] The focus representative [maliḍeɹ] 'of the gardeners' is labeled as bearing a sharp rise (L*+H) instead of a focused smooth rise (L*…fHa), as the F_0 maximum occurs at the end of the second syllable (*i.e.* expected F_0 maximum location for L*+H) instead of the end of the third syllable (*i.e.* expected F_0 maximum location for fHa). The use of the sharp rise is consistent with other wh-answer focus domains.

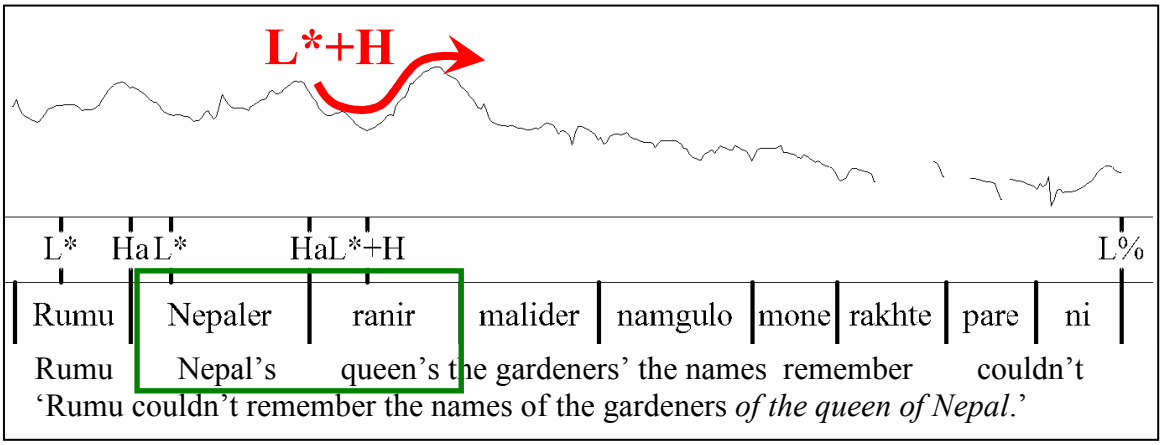


Figure 135. The complex-NP focus domain [nepaleɹ ɹaniɹ] 'of the queen of Nepal' bears only one sharp rise (L*+H),[90] located on the focus representative [ɹaniɹ] 'of the queen'. [To39]

The other realization type in complex-NP focus domains is repeated realization, which involves separately applying default focus realization (*i.e.* assignment of the sharp rise pattern L*+H) to multiple words in complex NPs within the focus domain. The quick succession of sharp rises (L*+H) causes each rising pitch accent's H component to merge into the following rising pitch accent's L component, making it virtually impossible to determine whether a high AP boundary tone (Ha) is projected between them. In Figure 136 below, the two-word focus domain [ɹaniɹ maliḍeɹ] 'of the gardeners of the queen' bears two sharp rises (L*+H), as an answer to the wh-question [ɹumu nepaleɹ kon

[90] The focus representative [ɹaniɹ] 'of the gardeners' is labeled as bearing a sharp rise (L*+H), but its realization is ambiguous between that and the focused smooth rise (L*…fHa), as the F_0 maximum is realized near the end of the syllable following the stressed syllable (*i.e.* expected location of F_0 maximum on L*+H), which is also near the right edge of the word (*i.e.* expected location of F_0 maximum on fHa). The primary reason for choosing the sharp rise (L*+H) label is consistency with other wh-answer examples showing less ambiguous realizations of the sharp rise (L*+H).

koɹmiḍeɹ namgulo mone ɹakʰt̪e paɹe ni?] ‘Rumu couldn’t remember the names of which workers of Nepal?’[91]

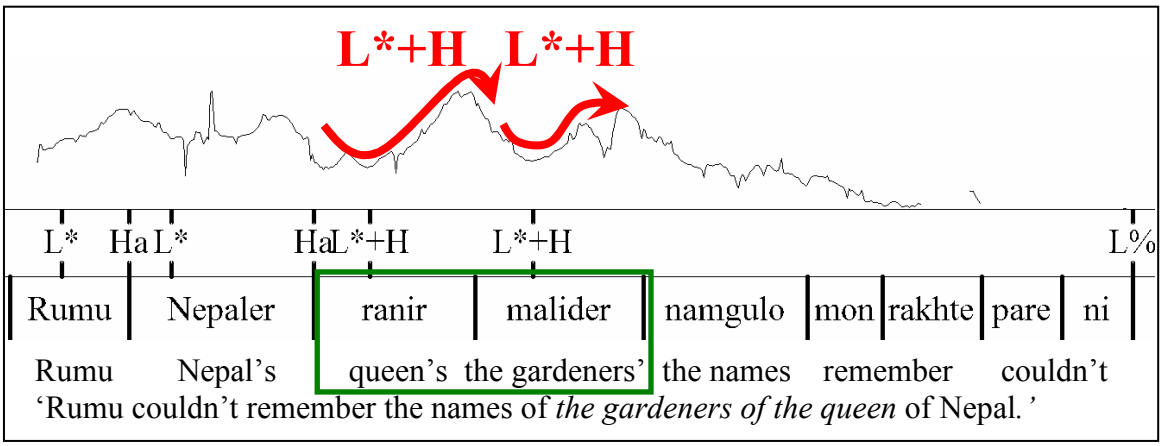


Figure 136. The two-word wh-answer focus domain [ɹaniɹ maliḍeɹ] ‘of the gardeners of the queen’ is marked with repeated focus realization (sharp rise exceeding the preceding Ha level) on the second and third words.[92] [To43]

The sentence in Figure 137 expands the focus domain seen in Figure 136 to include the word [namgulo] ‘the names’, allowing a third sharp rise (L*+H) to be assigned.

[91] As Bengali wh-questions do not require the same kinds of movement seen in English wh-questions, this particular utterance is not as awkward in Bengali as it sounds in English.

[92] As the F_0 maximum of the sharp rise (L*+H) tends to occur on the second or third syllable of the word, the two- and three-syllabled focused APs [ɹaniɹ] ‘of the queen’ and [maliḍeɹ] ‘of the gardeners’ are within the range of ambiguity between sharp rises (L*+H) and focused smooth rises (L*…fHa).

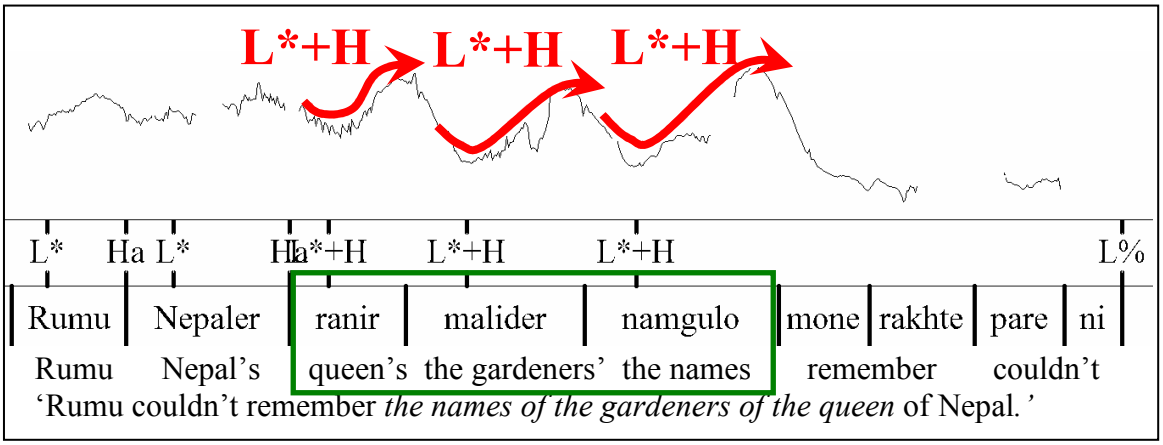


Figure 137. The three-word wh-answer focus domain [ɹaniɹ maliḏeɹ namgulo] 'the names of the gardeners of the queen' is marked with repeated focus realization (sharp rise exceeding the preceding Ha level) on each word. Each repetition of focus realization in this example is more exaggerated in pitch range than the previous; this is not consistent across all speakers, however.[93] [BM44]

Repeated realization does not have to occur on all of the APs in the focus domain; often, only the focus representative and one or two additional NPs will bear focus realization. In Figure 138, for example, the three-word wh-answer focus domain [nepaleɹ ɹaniɹ maliḏeɹ] 'of the gardeners of the queen of Nepal' only bears two sharp rises, one on the focus representative and one on the immediately preceding word.

[93] As the F_0 maximum of the sharp rise (L*+H) tends to occur on the second or third syllable of the word, the two- and three-syllabled focused APs [ɹaniɹ] 'of the queen', [maliḏeɹ] 'of the gardeners', and [namgulo] 'the names' are within the range of ambiguity between sharp rises (L*+H) and focused smooth rises (L*…fHa).

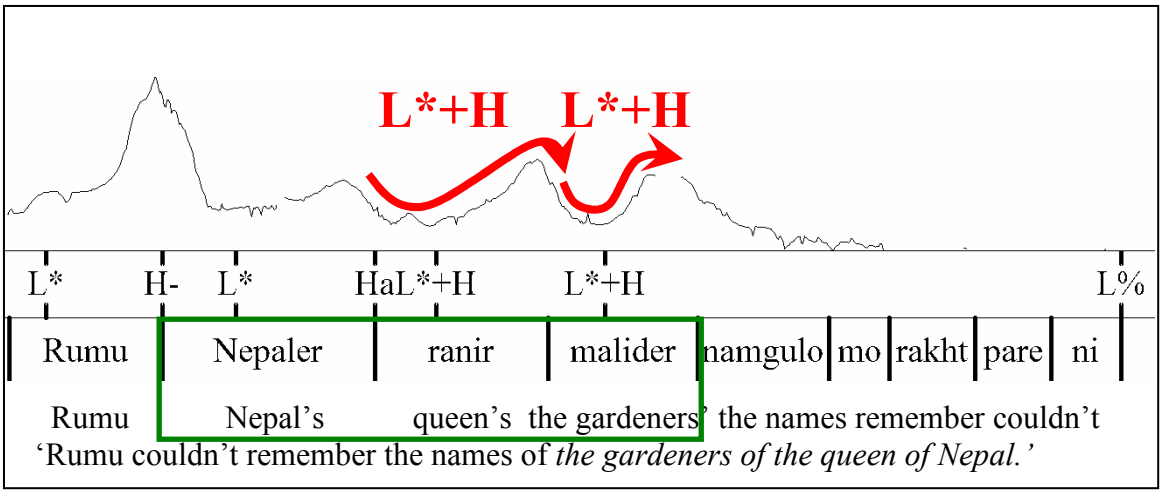


Figure 138. The three-word wh-answer focus domain [nepaleɹ ɹaniɹ maliḍeɹ] 'of the gardeners of the queen of Nepal' is marked with repeated focus realization (sharp rise exceeding the preceding Ha level) on the second and third words.[94] [To40]

Both repeated and representative realizations of complex-NP focus domains involve applying a sharp rise (L*+H) to the focus representative; the difference comes from whether or not other NPs within the domain also bear sharp rises (L*+H). Essentially, the two realization types can be considered phonological variants of each other, as they appear to be interchangeable. I now discuss properties of the post-focal string.

12.2 Post-focal tone compression and deletion

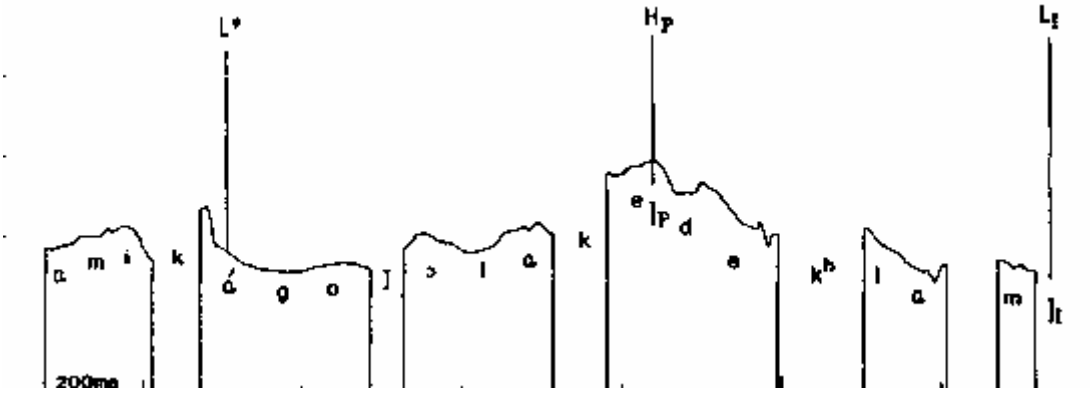
While many potential cues for the extent of focus domain can be found within the domain itself, one of the most potentially salient cues for focus is the deletion or compression of post-focal tones, where words following the surface realization of the

[94] As the F_0 maximum of the sharp rise (L*+H) tends to occur on the second or third syllable of the word, the two- and three-syllabled focused APs [ɹaniɹ] 'of the queen' and [maliḍeɹ] 'of the gardeners' are within the range of ambiguity between sharp rises (L*+H) and focused smooth rises (L*…fHa). Still, the F_0 maximum of [maliḍeɹ] 'of the gardeners' appears more like that of a sharp rise (L*+H), as it occurs just before the onset of the third syllable.

focus high tone (fH) do not bear any AP-level tones (*i.e.* pitch accents and AP boundary tones), or only bear phonetically weak AP-level tones. This phenomenon is often referred to as deaccenting or dephrasing in descriptions of other languages. Previous analyses of the underlying structure of post-focal material suggest there is a lack of tones between the focus domain and the right edge of the Intonation Phrase, and that pitch is determined simply by phonetic interpolation of F_0 between the F_0 maximum of the focused constituent and boundary tone borne on the Intonational Phrase boundary. Data from the current study regarding the shape of the F_0 contour post-focally are mixed; some examples appear to illustrate this simple phonetic interpolation, while others suggest an additional L target following the realization of the focus high tone (fH). Furthermore, data from the current study indicate that post-focal tone deletion begins immediately following the realization of the focus high tone (fH), and not necessarily following the focus domain's right boundary.

Hayes & Lahiri (1991) and Lahiri & Fitzpatrick-Cole (1999) explain that focused constituents must bear the most metrically-prominent syllable in the larger domain, and that the most metrically-prominent syllable in that domain (*i.e.* Intonation Phrase or I-phrase in their models) is consistently the rightmost pitch-accented syllable. These two factors work together to prevent the assignment of any pitch accents to post-focal material, which would in effect create a pitch accent-less—and thus also P-phrase boundary tone-less—string of words between the focused constituent and the I-phrase boundary tone (T%). Accompanying examples in both studies show long stretches of

falling and/or low pitch from the F_0 maximum of the focus domain's high P-phrase boundary tone (H_P) up to the I-phrase boundary tone (T_P), as shown in Figure 139 and Figure 140 (taken from Hayes & Lahiri 1991), and Figure 141 (taken from Lahiri & Fitzpatrick-Cole 1999) below.



ami	kágoǰɔlake	dekʰlam
[ami	kagodʑɔlake	d̪ekʰlam]
I	newspaperman-ACC	saw

'I saw the *newspaperman*.'

Figure 139. After the focused constituent [kagodʑɔlake] 'newspaperman-ACC', there are no pitch accents (T*) or P-phrase boundary tones (T_P) up through the low I-phrase boundary tone (L_I). [H&L 18, §4.2, p. 59]

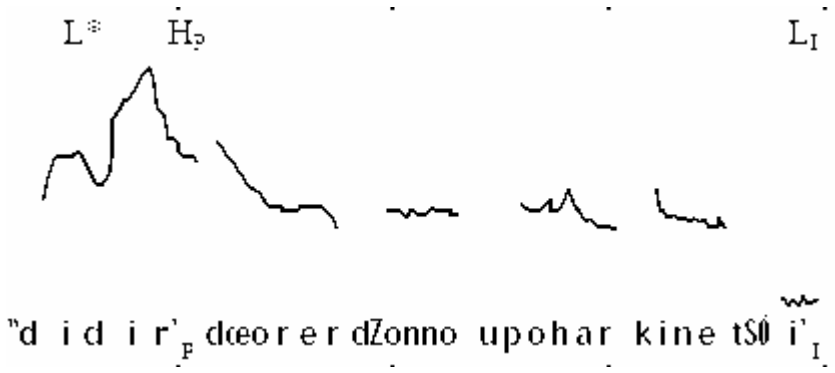
Note that the verb in Figure 139 above bears no pitch accent. When the focus domain is located further from the I-phrase boundary tone, the stretch of pitch accent-less and boundary tone-less pitch lengthens, as in Figure 140 below.

tumi	kón	mačʰer	matʰa	ranna	korle
[˻tumi	kon	matɕʰeɹ	maṯʰa	ɹanːa	koɹle]
you	which	fish's	head	cook	did

'*Which fish*-head did you cook?'

Figure 140. After the wh-question focused constituent [kon matɕʰeɹ] 'which fish's', there are no pitch accents (T*) or P-phrase boundary tones (T_P) up through the low I-phrase boundary tone (L_I). [H&L 21, §4.2, p. 61]

The same pattern is seen in Lahiri & Fitzpatrick-Cole's (1999) study, an example of which is shown in Figure 141 below.

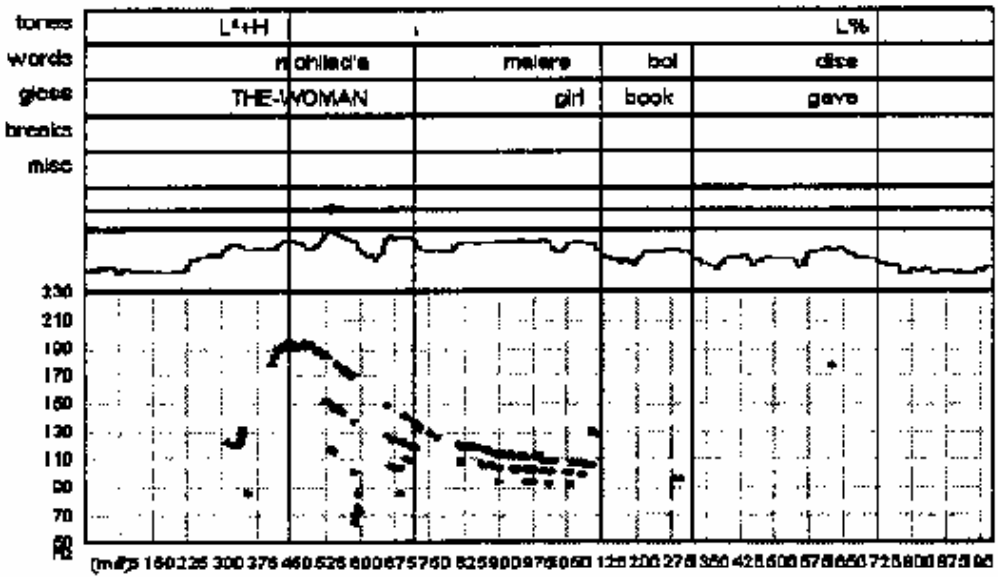


didir	dæorer	dZonno	upohar	kinetS^{h}i
[d̪id̪iɹ	d̪ɛɔ̯ɹeɹ	dʑon:o	upohaɹ	kinetɕ^{h}i]
sister's	brother-in-law	for	present	I bought

'I bought a present for *sister's* brother-in-law.'

Figure 141. After the wh-answer focused constituent [d̪id̪iɹ] 'elder sister's', there are no pitch accents (T*) or P-phrase boundary tones (T_P) up through the low I-phrase boundary tone (L_I). [L&F-C 27, §4]

Michaels & Nelson (2004) describe an "obligatory low plateau" following focused constituents, as shown in Figure 142 below. They interpret this phenomenon as a toneless stretch of text extending from the F_0 maximum (*i.e.* the H component) of the focused constituent's bitonal pitch accent (L*+H) to the F_0 minimum of the low IP boundary tone (L%), and explain that the effect of this tone deletion is the highlighting of the focused constituent as the most prosodically prominent item in the sentence.



mohilad'a maiaɹe boi dise
[mohilada majaɹe boj d̪ise]
the woman girl book gave
'*The woman* gave the girl the book.'

Figure 142. After the corrective focused constituent [mohilada] '*the woman*', there are no pitch accents (T*) or ip boundary tones (T-) up through the low IP boundary tone (L%). [M&N 22, §2.6.1, p. 22]

Data collected for the current study are largely consistent with these previous descriptions of the post-focal string. In the majority of cases, the realization of the focus high tone (fH)—either fHa, L*+H, or fH*—is the last AP-level tone (*i.e.* pitch accent or AP boundary tone) before the IP boundary tone (T%), regardless of the length of the string. In some cases, as in Figure 143 below, post-focal pitch is determined by simply interpolating between the realization of the focus high tone (fH) and the IP boundary tone (T%). In the current description, this tone deletion pattern is known as pure interpolation.

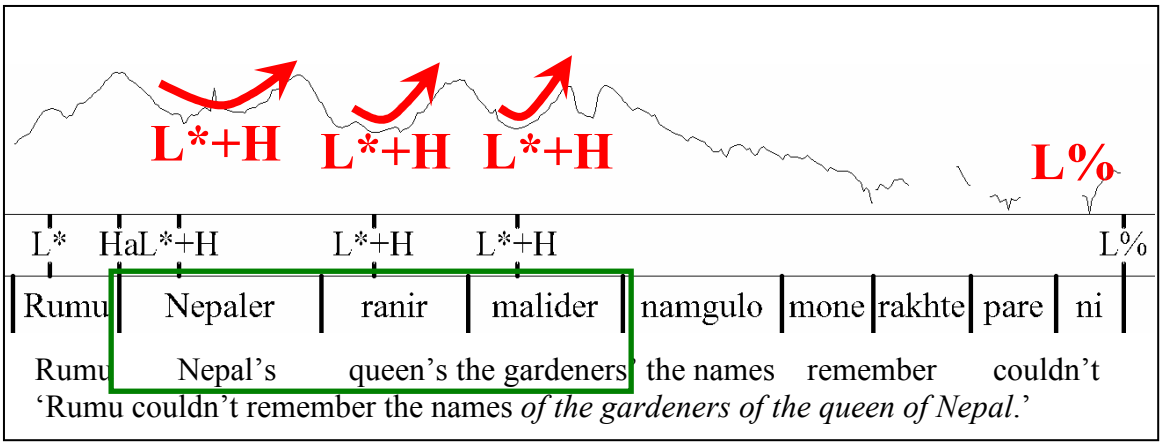


Figure 143. Pitch is interpolated directly from the pitch maximum of the focus representative [maliḏeɹ] 'of the gardeners' and the low IP boundary tone (L%). It is unclear whether the non-final APs in the focus domain [nepaleɹ ɹaniɹ maliḏeɹ] 'Nepal's queen's gardeners'' bear sharp rises (L*+H) or smooth rises (L*…Ha)**.**[95] [To41]

Other examples, however, show a sharper drop following the realization of the focus high tone (fH) before reaching a flatter pitch. In the current description, this pattern of tone deletion is called sharp initial drop. Note in Figure 144 below how the pitch drops somewhat sharply following the focused word [ɹaniɹ] 'queen-GEN', before flattening out towards the end of the sentence. This sharp initial drop is marked tentatively with a low ip boundary tone (L-) to mark what appears to be a low target following the focused constituent.

[95] Technically, this sentence is the incorrect answer to the question [ɹumu ki dʑiniʃ mone ɹakʰt̪e paɹe ni?] 'What (thing) did Rumu not remember?', as the focus domain of the answer should include the word [namgulo] 'the names'. Still, I feel it is a good example of pure interpolation.

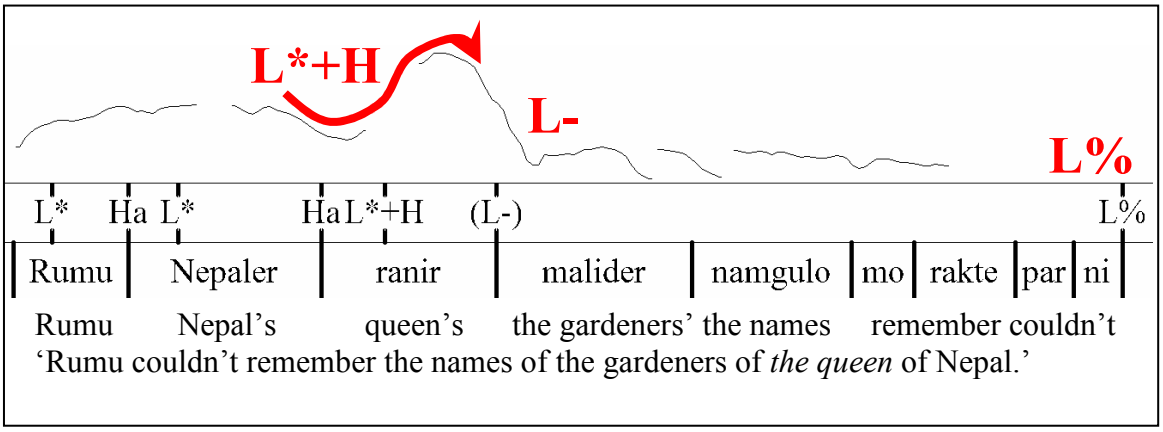


Figure 144. The pitch maximum of the rising pitch accent (L*+H) drops sharply into what may be a low ip boundary tone (L-)—identified by the pitch contour and moderate final lengthening—before flattening out into a low-pitched stretch of deleted tones. [Sf42]

Sometimes the drop towards low pitch occurs without the larger break associated with ip boundaries, as in below. In cases such as these, it is unclear what is causing the sharp initial drop.

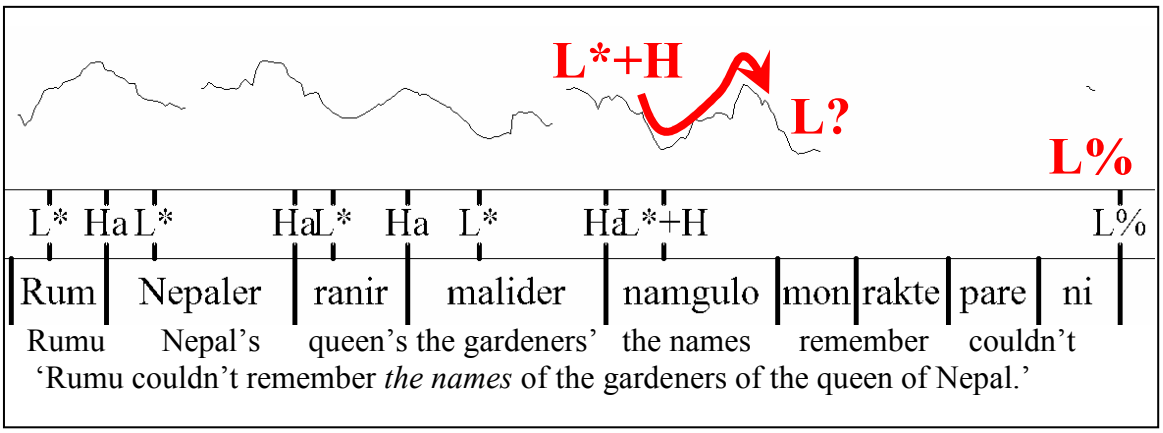


Figure 145. The pitch maximum of the rising pitch accent (L*+H) drops sharply before flattening out into a string of low pitch with no AP-level tones. It is unclear what the identity of this L target may be; a low pitch accent (L*) on [mone] 'mind-LOC' or a low

AP or ip boundary tone (La, L-) are possibilities. The pitch track is lost after the focused word [namgulo] 'the names' due to creaky phonation. [Fa47]

A similar weak L target is often seen between high tones, where a long stretch of high pitch is produced across the string of deleted tones from the focus high tone (fH) to a high or falling IP boundary tone (H%, HL%), as in Figure 146 below. In this example, an L target follows the focused constituent; the L target is possibly a phonetically reduced low pitch accent (L*), as I perceive (somewhat weak) metrical prominence on the stressed syllable of the word [ɹaniɹ] 'queen-GEN'.

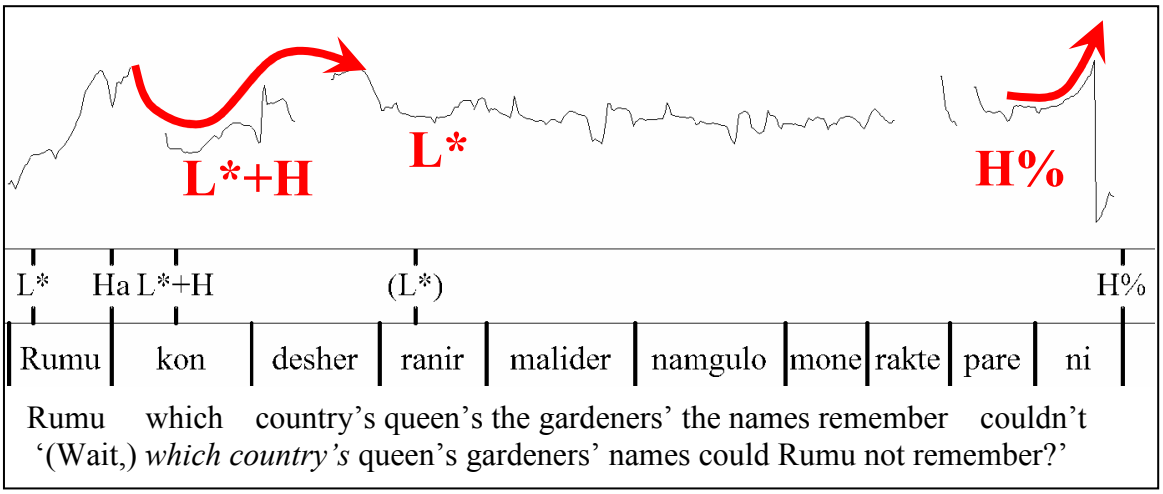


Figure 146. The high pitch of the sharp rise (L*+H) borne on [kon] 'which' falls somewhat into what may be a weakened low pitch accent (L*) on [ɹaniɹ] 'queen-GEN', labeled with the circle, before flattening out to a long stretch of high pitch reaching the high IP boundary tone (H%) associated with confirmation questions. Note that the pitch of the high IP boundary tone (H%) reaches such a great height that the software halves the F_0 value from about 400Hz to about 200Hz. [Fa38]

It is unclear if the sharp initial drop is caused by the same type of tonal target (*i.e.* L*, L-) or if it simply must be stipulated that focused constituents are often followed by a

sharp drop in pitch before the string of deleted tones. It is equally unclear if the two tone deletion patterns (*i.e.* pure interpolation, sharp initial drop) are in free variation, or if there is a tonal difference between them.

Regardless of whether or not there is an L target following the focused constituent, it is important to characterize the structure of the post-focal string. In Figure 147 below, the focused constituent [maliḏeɹ] 'of the gardeners' is followed by a string consisting of five toneless orthographic words [namgulo mone ɹakʰt̪e paɹe ni], which in other contexts could probably be parsed into three APs: [namgulo] 'the names', [mone ɹakʰt̪e] 'to remember' and [paɹe ni] 'wasn't able'.

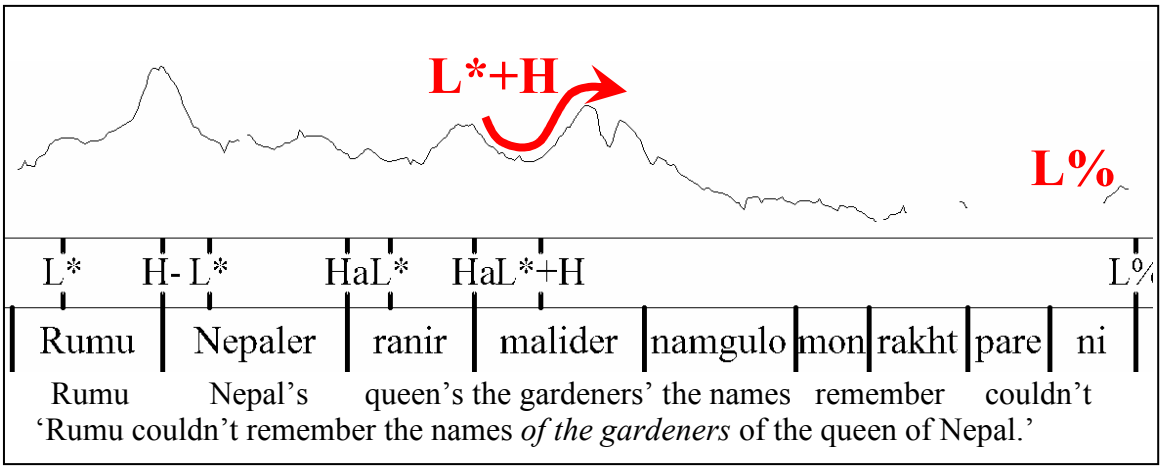


Figure 147. The words following the focused constituent [maliḏeɹ] 'of the gardeners' do not bear any pitch accents or boundary tones, and thus their pitch is entirely determined by phonetic interpolation of adjacent tones. It is unclear why the pitch maximum of [ɹaniɹ] 'queen's' breaks Ha downtrend. [To45]

The reader may be asking the same question I have: what is the structure of the string [namɡulo mone ɹakʰt̪e paɹe ni]? Unfortunately, because APs are primarily a tonal unit (see Chapter Three), and AP boundaries cannot be identified from their duration, *etc.*, it is difficult to understand what the phrasing of the post-focal string might be. I consider four possibilities:

(1) **Loss of metrical prominence:** The stressed syllables of post-focal words lose their otherwise-expected metrical prominence, and thus no pitch accents are attracted to those words. Thus, there are neither any pitch accents nor boundary tones following the realization of the focus high tone (fH), either underlyingly or on the surface. AP boundary tones cannot be projected post-focally, as there are no heads (*i.e.* pitch accents) to project them. The focused constituent is the last AP in the sentence, violating the Exhaustivity component of the Strict Layer Hypothesis (Selkirk 1984, 1986; Nespor & Vogel 1986). This analysis is schematized below in Figure 148.

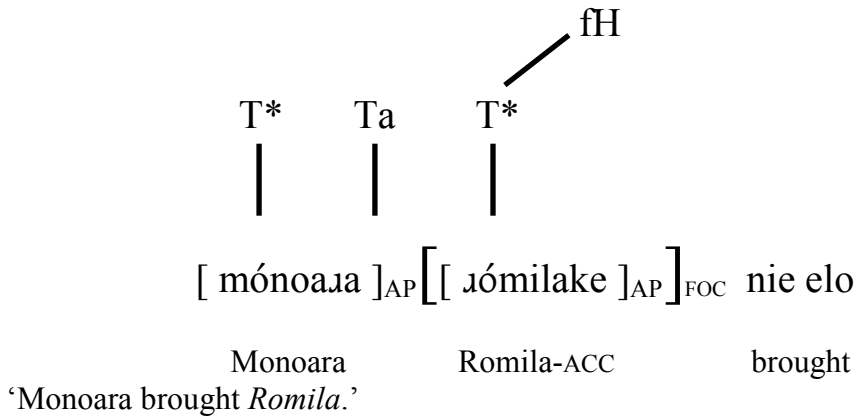


Figure 148. Schematic illustration of post-focal tone deletion as the result of a lack of sufficient metrical prominence (note the lack of the acute accent mark representing stress on [nie]) following the focused word, which bears the underlying focus high tone (fH) adjoined to the low pitch accent (L*), creating the bitonal pitch accent (L*+H).[96]

(2) **Prevention of pitch accent assignment:** The stressed syllables of post-focal words in fact maintain their metrical prominence, but pitch accents are nevertheless not allowed to be associated with those syllables, due to an additional constraint against the assignment of tones to post-focal words. AP boundary tones cannot be projected post-focally, as there are no heads (*i.e.* pitch accents) to project them. The focused constituent is thus the last AP in the sentence, violating Exhaustivity. This analysis is schematized below in Figure 149.

[96] An additional constraint would need to be stipulated in the "lack of metrical prominence" analysis in order to explain the lack of the high AP boundary tone (Ha) on focused constituents bearing the rising pitch accent (L*+H).

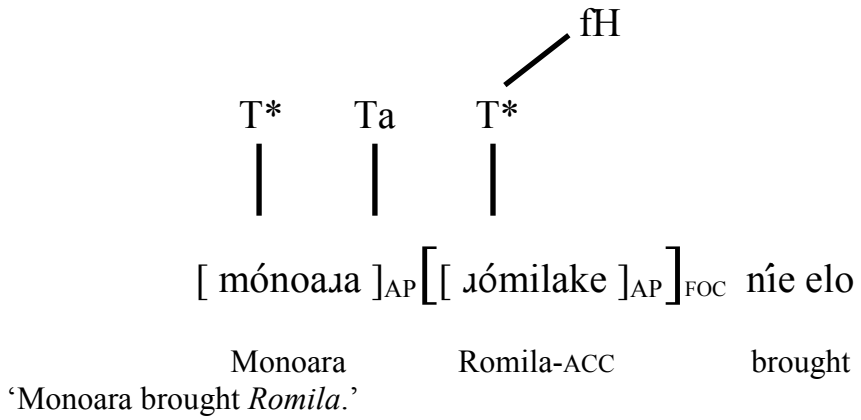
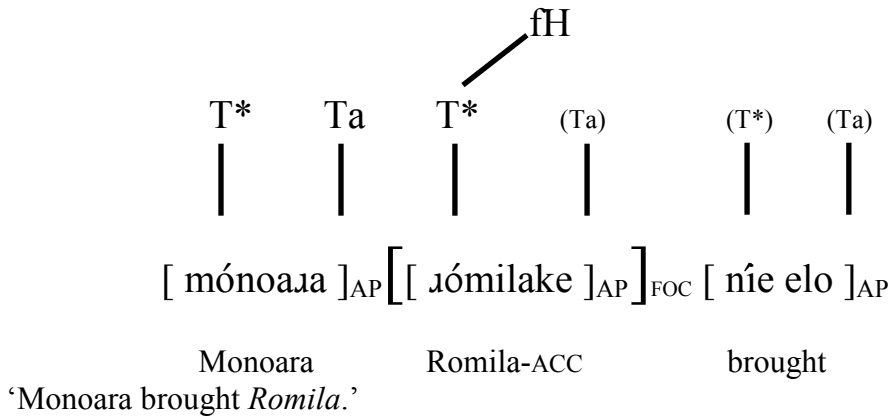


Figure 149. Schematic illustration of post-focal tone deletion as the prevention of pitch accent assignment following the realization of the underlying focus high tone (fH), which is seen here adjoined to the low pitch accent (L*) to create the bitonal pitch accent (L*+H).[97]

(3) **Pitch range compression:** The stressed syllables of post-focal words in fact maintain their metrical prominence, but the tones they attract and project are subject to a compression in post-focal pitch range. Thus, pitch accents and the AP boundary tones they project in fact exist post-focally, but their realization is phonetically reduced, often to the point that many pitch accents and AP boundary tones become largely indistinguishable amongst microprosodic changes in the pitch track. This analysis is schematized below in Figure 150.[98]

[97] As with the "loss of metrical prominence" analysis, an additional constraint must be stipulated in the "prevention of pitch accent assignment" analysis to explain the lack of the high AP boundary tone (Ha) on focused words bearing the rising pitch accent (L*+H).

[98] Compression of post-focal tones is also attested in Swedish (Bruce 1977).



'Monoara brought *Romila*.'

Figure 150. Schematic illustration of post-focal tone deletion as extreme compression of AP-level tones following the realization of the underlying focus high tone (fH), which is seen here adjoined to the low pitch accent (L*) to create the bitonal pitch accent (L*+H). Smaller font and parentheses represent compression.

(4) **Dephrasing**: Post-focal words are grouped together with the focused constituent into one very large AP, so they all share one pitch accent and project one AP boundary tone, which is overridden by the following IP boundary tone. This analysis is schematized below in Figure 151.

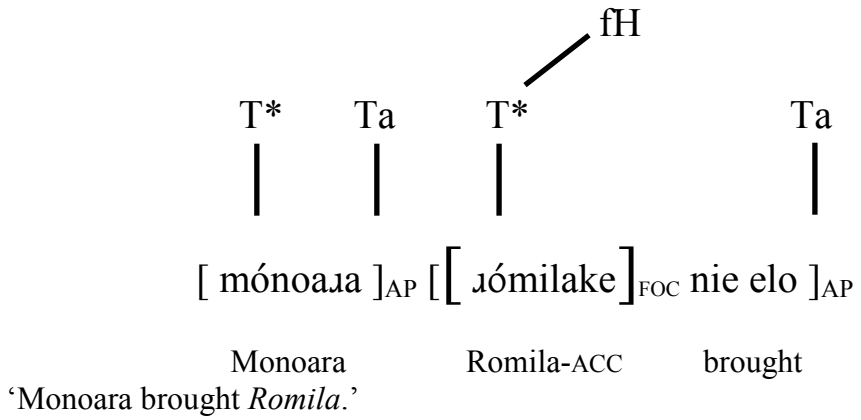


Figure 151. Schematic illustration of post-focal tone deletion as dephrasing of words following the realization of the underlying focus high tone (fH), which is seen here adjoined to the low pitch accent (L*) to create the bitonal pitch accent (L*+H).

It is clear in instances where the focused constituent bears a focused high AP boundary tone (fHa) that option (4)—dephrasing—is unlikely. The post-focal string in such cases follows the focused constituent's AP boundary tone, and thus cannot be parsed into the focused constituent's AP. Dephrasing is a possibility in situations where the focused word bears a rising pitch accent (L*+H) and no AP boundary tone. It is harder to differentiate the other three options. Post-focal tone deletion is not obligatory; there are numerous examples of sentences in which the focused constituent bears the focus high tone (fH) but is also followed by pitch accents and AP boundary tones, although often with phonetic reduction, supporting the pitch range compression theory (3) provided above.[99] Even the same speaker can vary between tonal deletion and compression across productions of the same sentence or similar sentences. For example, compare Figure 152

[99] Many examples of sentences containing focused constituents not followed by tonal deletion, however, were considered ungrammatical by my primary consultant, given the focus context elicited. See Nagahara (1994) for similar variation in Japanese.

and Figure 153 below; note that in Figure 152, the focus-encliticized constituent [monoaɹaj] ‘(only) *Monoara*’ triggers total tone deletion on the following words [ɹomilake] ‘Romila-ACC’ and [nie elo] ‘brought’, which normally form separate APs. The very similar sentence in Figure 153 contains the focus-encliticized constituent [monoaɹaə̯] ‘even *Monoara*’, but the following words are parsed into two APs complete with pitch accents and boundary tones, although the pitch range for the tones may be somewhat compressed.

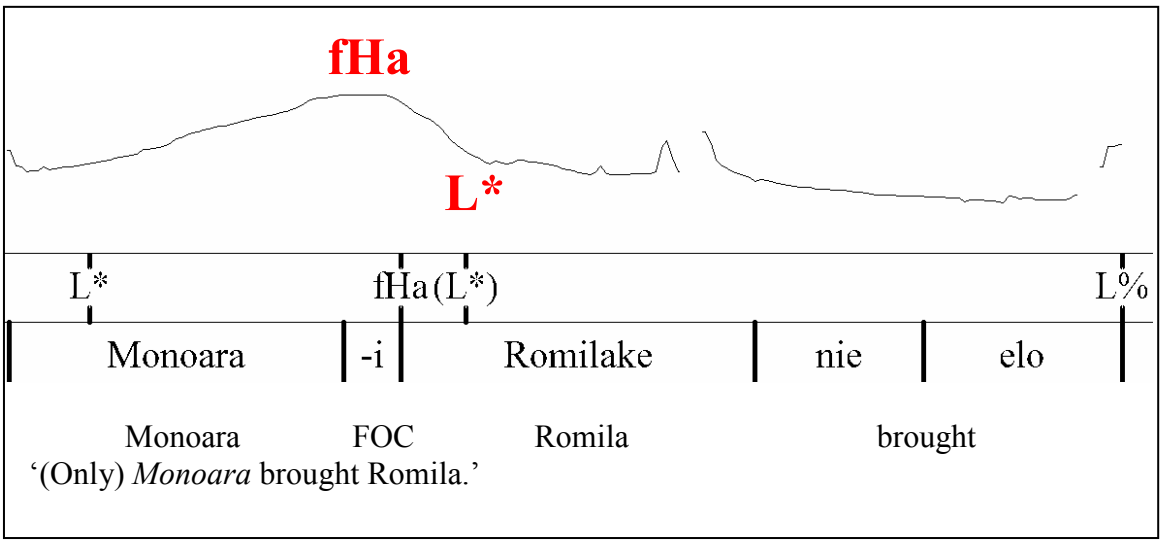


Figure 152. In this sentence, the speaker deletes at least two AP-level tones (*i.e.* one pitch accent, one AP boundary tone) following the surface realization of the focus high tone (fH). The low pitch accent (L*) immediately following the focus domain is evident in the pitch contour, although the perceived prominence during that syllable is weak. [Na11]

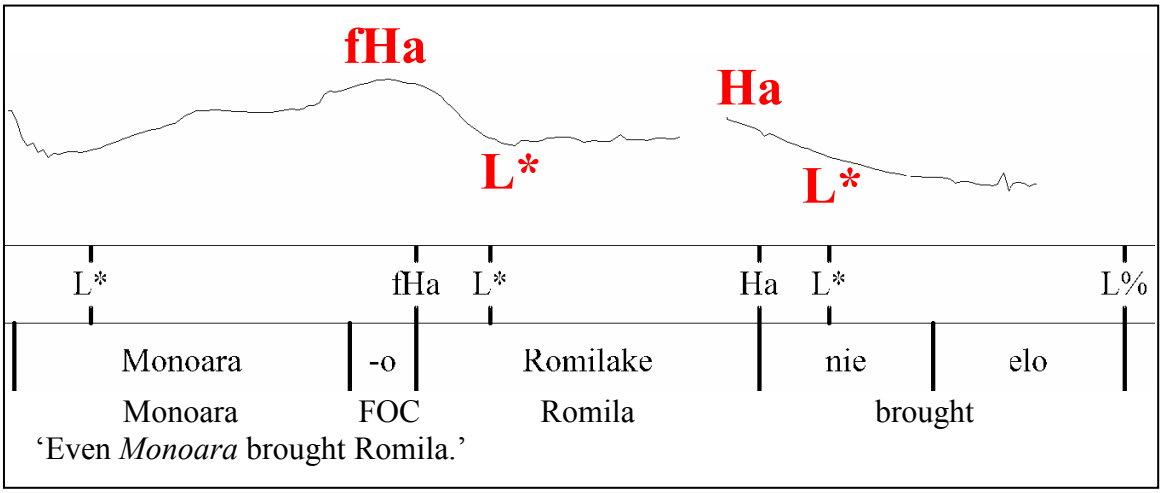


Figure 153. In this sentence, the same speaker maintains all AP-level tones following the realization of the focus high tone (fH), although the pitch range of the post-focal tones appears somewhat compressed. [Na13]

In some cases, only some post-focal words are stripped of their AP-level tones, while others are assigned AP-level tones, albeit with some phonetic reduction (*i.e.* compression), as in Figure 154 below.

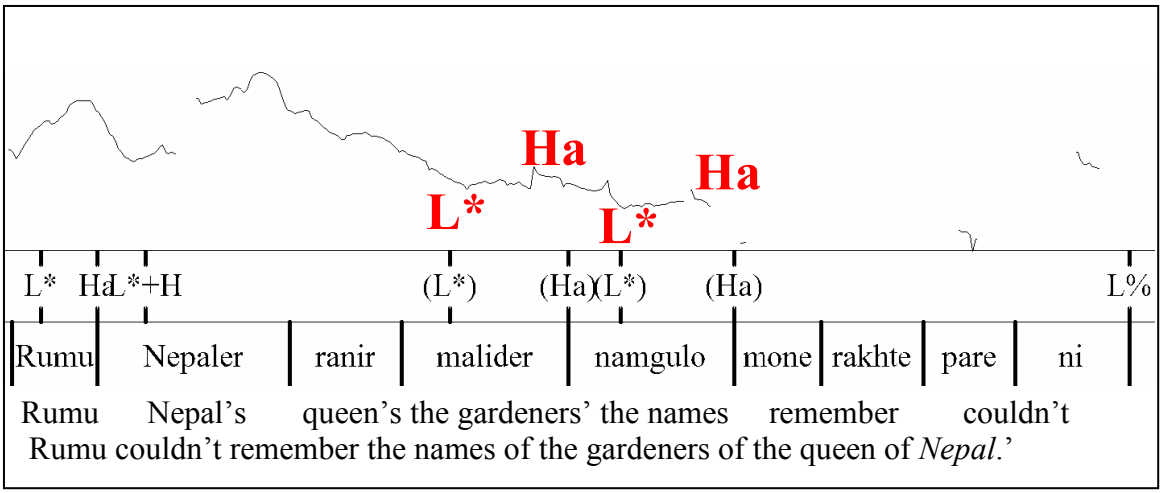


Figure 154. In this sentence, many of the post-focal words bear no AP-level tones, but two of them—[malid̪eɹ] 'gardeners-GEN' and [namgulo] 'the names'—bear (somewhat reduced) smooth rises (L*…Ha). Note the irregular pitch track from the edge of [namgulo] due to creaky phonation. [Fa38]

Given the variation and phonetic reduction seen in the post-focal string, it appears that it cannot be said with any certainty at this time how to characterize its prosodic makeup. What can be said is that the domain of post-focal tone deletion—when it occurs—is no smaller than the IP. Post-focal words can be stripped of AP-level tones even following a high ip boundary tone (H-) projected to the right of the focused word, as in Figure 155, indicating that ip boundary tones (T-) neither block nor undergo post-focal tone deletion.[100]

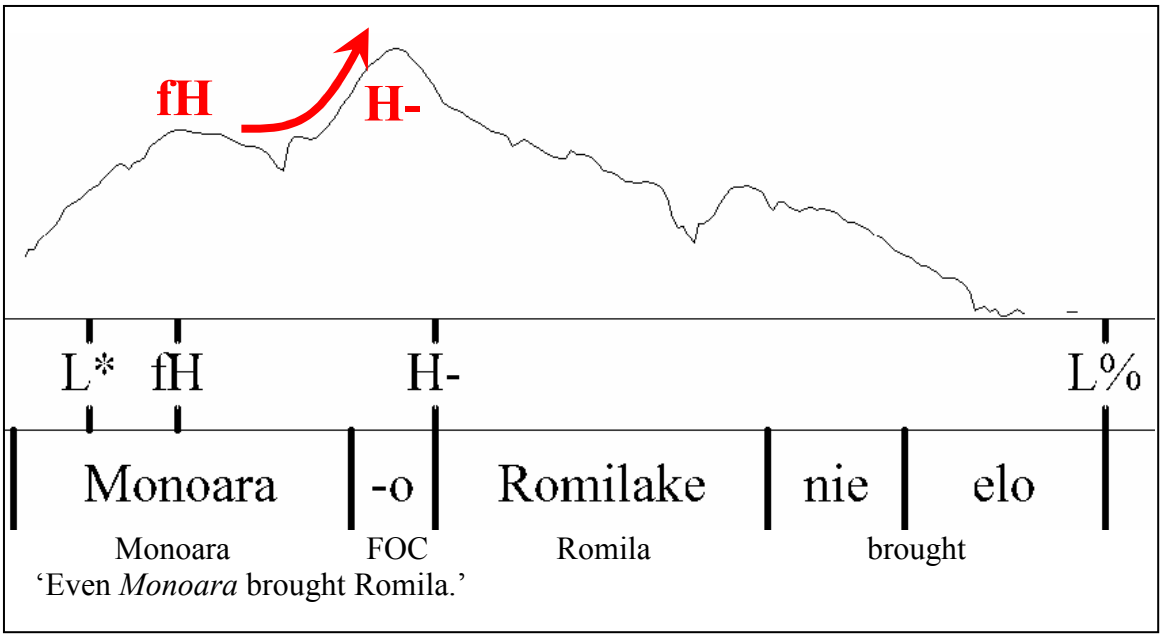


Figure 155. Although [ɹomilake] 'Romila-ACC' and [nie elo] 'brought' normally bear pitch accents and, in the case of [ɹomilake], an AP boundary tone, these tones are suppressed as they follow the focus high tone (fH) on [monoaɹaə̯] 'even *Monoara*'.[100] The intervening high ip boundary tone (H-) does not block tone deletion**.** [Fe13]

I now move on to discuss the interaction of the focus high tone (fH) with other tones.

12.3 Tonal interaction

This section is concerned with the interaction of boundary tones of higher prosodic units (*i.e.* T%, T-) and two of the three focused tonal patterns: focused smooth rise (L*…fHa) and sharp rise (L*+H). The focused smooth rise (L*…fHa) reacts differently to different cooccurring boundary tones, and is described in §12.3.1. In §12.3.2, I show how the sharp rise (L*+H) can cooccur with almost all higher level boundary tones, unlike the focused tonal patterns seen in models of other varieties of Bengali.

[100] See §12.3.1 for a discussion of this particular realization of the focus high tone (fH).

12.3.1 Focused smooth rise (L*…fHa)

When the focus high tone (fH) fuses with the high AP boundary tone (Ha) to create the focused high AP boundary tone (fHa), the main feature identifying its AP as being focused is the relative height of the boundary tone. Therefore, it is presumably of high importance to maintain this boundary tone, even when faced with the danger of concurrent boundary tone overriding. As post-focal tone deletion only compresses or suppresses tones of the AP-level (*i.e.* pitch accents and AP boundary tones), it cannot affect the tones of higher prosodic units (*i.e.* ip and IP boundary tones). Thus, when the focused high AP boundary tone appears ip-finally, it must find another way to avoid concurrent boundary tone overriding. Depending on the type of tone with which it is cooccurring, the expanded high AP boundary tone (fHa) of the focused smooth rise (L*…fHa) can either adjoin to the higher level boundary tone or shift away from it. I first describe the adjunction of the expanded high AP boundary tone (fHa) to L boundary tones (*i.e.* L-, L%), and then move on to high tone shift.

Adjunction to L boundary tones (fHaL-, fHaL%)

When a focus-encliticized constituent occurs ip-finally before a low ip or IP boundary tone (L-, L%), it avoids being overridden by the higher boundary tone by adjoining to it, forming a stacked tone (*i.e.* fHaL-, fHaL%) similar to the higher boundary tones of English (*e.g.* L-H%). Observe the pitch contour during the IP-final word in Figure 156 and Figure 157 below. Note how the IP-final focus enclitics -[i]~[j] 'only',

‘indeed’ and -[o]~[ɔ̰] ‘also’, ‘even’ force the raised high AP boundary tone (fHa) to be realized despite the concurrent low IP boundary tone (L%). The combined tone (fHaL%) is realized with a pitch maximum (for the fHa) immediately preceding the pitch minimum (for the L%), concentrated at the end of the IP-final syllables [lo-i] and [lo-o].[101] Just as the raised high AP boundary tone (fHa) is distinguishable from other high AP boundary tones (Ha) by its refusal to obey Ha downtrend, it seems that this violation of concurrent boundary tone overriding helps to amplify the realization of the encliticized constituent’s focused status.

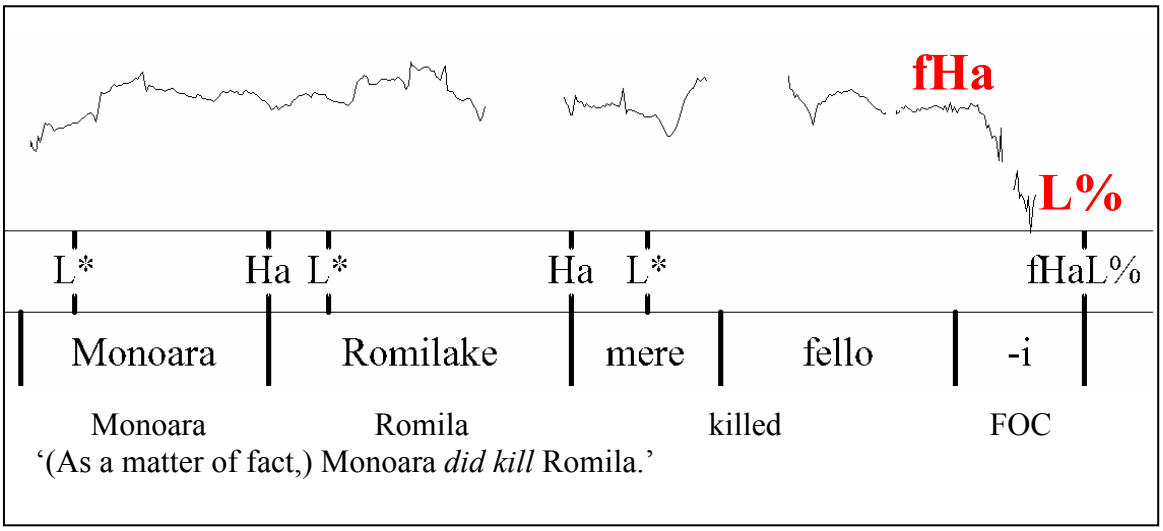


Figure 156. The focus-encliticized verb [meɹe felːo-i] ‘*did kill*’ bears its focused high AP boundary tone (fHa) despite the concurrent low IP boundary tone. The AP tone (fHa) is realized earlier than the IP tone (L%), although on the same **sy**llable. [FoFSA5]

[101] Although the string /…lo-i/ would normally syllabify as […loj], the stacking of tones presumably causes the clitic -/i/ to maintain its syllabicity. The relatively strong formants during the clitic -/i/ suggest that the string should be transcribed […lo.i]. A similar transcription could be made for /…lo-o/, giving […lo.o].

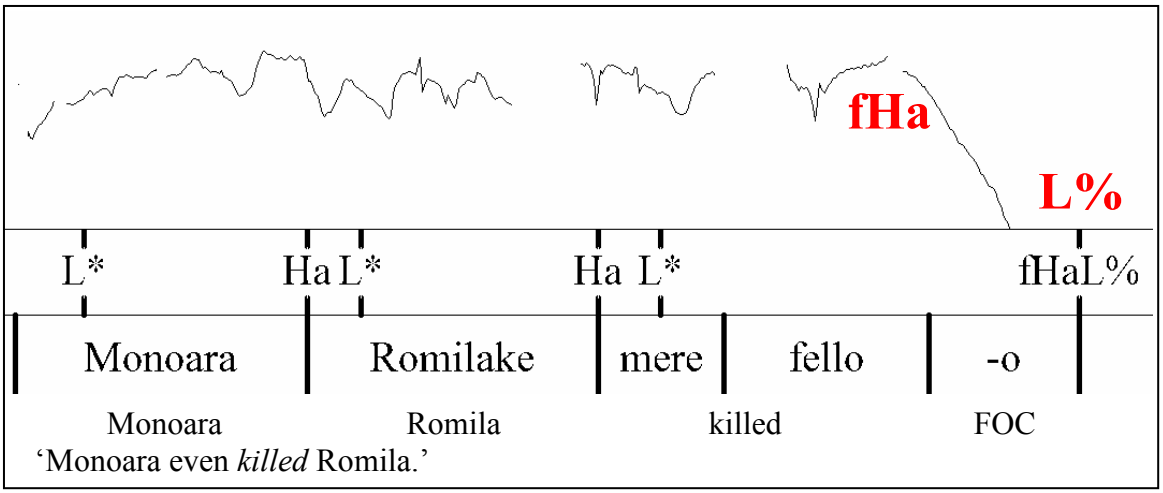


Figure 157. The focus-encliticized verb [meɹe felːo-o] 'even *killed*' bears its focused high AP boundary tone (fHa) despite the concurrent low IP boundary tone. The AP tone (fHa) is realized earlier than the IP tone (L%), although on the same syllable**.** [FoFSA3]

This pattern of adjunction to an L tonal target is not appropriate for when the focused high AP boundary tone (fHa) comes into contact with a higher boundary tone of the H category such as the high ip boundary tone (H-). See §12.3.1 for a discussion of high tone shift.

High tone shift

Although constituents marked by focus enclitics normally bear the highest pitch at their right edge (as the focus high tone fH fuses with the AP boundary tone), as introduced in §11.1, this is not the case when the right edge of a focus-encliticized constituent is concurrent with a high ip boundary tone (H-). As discussed previously, attempting to assign the focused smooth rise tonal pattern (L*…fHa) to ip-final position would cause the focused high AP boundary tone (fHa) to be overridden by the concurrent

ip boundary tone (T-), leaving the encliticized constituent with no tonal vestiges of its focused status. Adjoining the focused high AP boundary tone (fHa) to the high ip boundary tone (H-)—as is done with low ip and IP boundary tones (L-, L%)—would also be fruitless, as it would presumably be difficult to distinguish the pitch maxima of the two adjoined tones. To avoid both of these troubling situations, it appears that speakers shift the pitch maximum from the AP boundary tone (Ha) leftward towards the pitch accent (L*), typically appearing in the syllable immediately following the main stress. This shift allows some separation of the two H targets, as the H target of the ip boundary tone is only realized on the final syllable (due to the ip boundary tone locality constraint), leaving a slight sag in pitch between the two H targets, as shown in Figure 158 below.

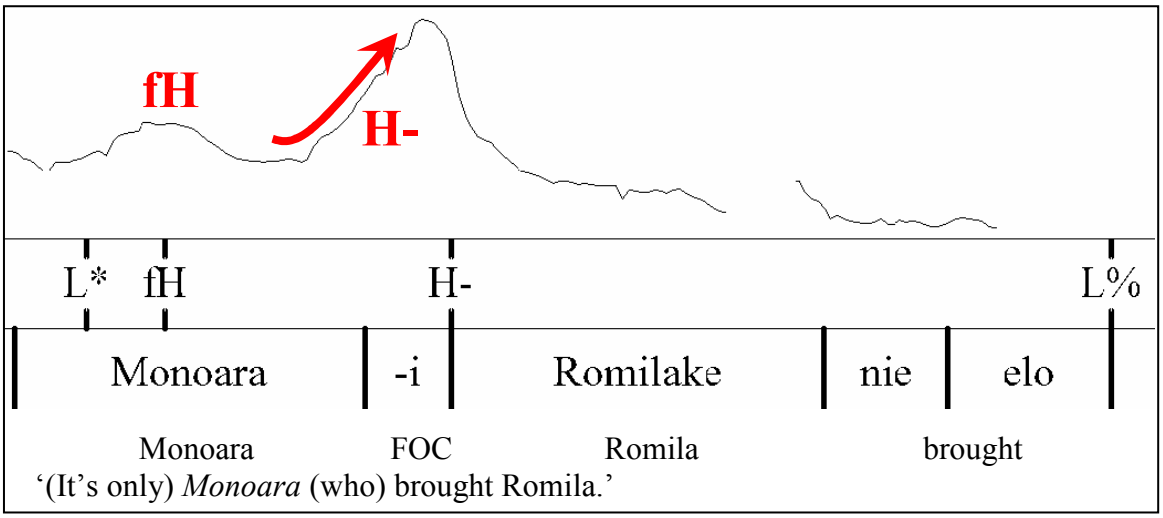


Figure 158. Although the encliticized AP [monoaɹaj] '(only) *Monoara*' is expected to bear the focus high tone (fH) at its right edge, the tone appears during the second syllable [no] instead, avoiding a clash with the high ip boundary tone (H-)—which can be identified by its ip-final rise in pitch from mid to high, and moderate final syllable lengthening**.** The representation of the focus high tone (fH) is left ambiguous. [Ba11]

The same pattern is seen in sentence-medial position, as in Figure 159 below.

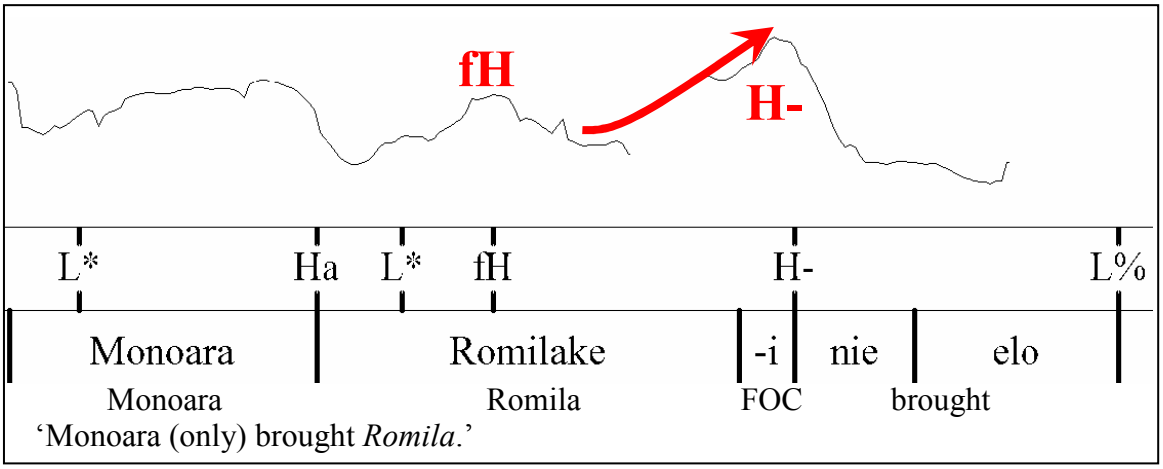


Figure 159. Although the encliticized AP [ɹomilakej] '(only) *Romila*-ACC' is expected to bear the focus high tone (fH) at its right edge, the tone appears during the second syllable [no] instead, avoiding a clash with the high ip boundary tone (H-)—which can be identified by its ip-final rise in pitch from mid to high, and moderate final syllable lengthening. The representation of the focus high tone (fH) is left ambiguous. [Ba12]

Although it is not entirely surprising that the pitch maximum would shift leftward to avoid overriding of the focus high tone (fH), it is unclear what the phonological representation of this shift would be. Three possible analyses come to mind:

(1) **Early realization (>fHa):** The focus high tone (fH) attaches to the high AP boundary tone (Ha) of the smooth rise (L*…Ha), but is realized immediately following the pitch accented syllable—much earlier than would be normally expected of an AP boundary tone.

(2) **Reassignment (L*+H):** The focus high tone (fH) attaches to the low pitch accent (L*) of the smooth rise (L*…Ha) instead of to the high AP boundary tone (Ha), which is overridden. The adjunction of the focus high tone (fH) to the low pitch accent (L*) creates a rising pitch accent (L*+H), already seen in constituents bearing corrective/wh-answer focus.

(3) **Detachment (fH):** The focus high tone (fH) attaches to neither the low pitch accent (L*) nor the high AP boundary tone (Ha) of the smooth rise (L*…Ha), and is realized instead as a floating tone between the pitch accent and overridden AP boundary tone.

These three analyses are schematized below in Figure 160.

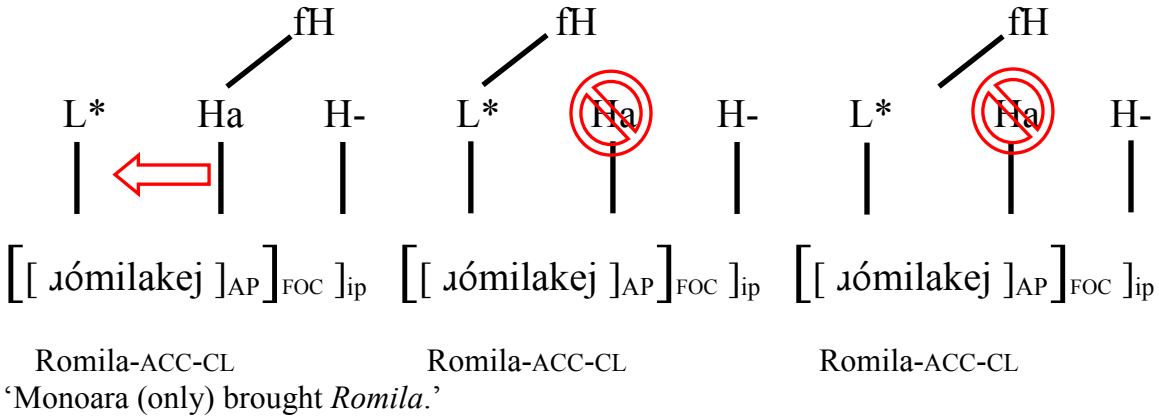


Figure 160. Schematic illustration of the three analyses of the leftward shifting of the focus high tone (fH) adjacent to a high ip boundary tone (H-): early realization, reassignment, and detachment. Crossed circles represent overridden tones.

Each of these three analyses (*i.e.* early realization, reassignment, detachment) has its own particular drawbacks. Early realization of the focused high AP boundary tone

(fHa)—which could be transcribed using the early realization diacritic “>” used in other ToBI systems, giving >fHa—is preferable in that it makes use of the basic pattern already seen in other encliticized-focused constituents, while the reassignment analysis (giving L*+H) requires that the sharp rise be posited not only for wh-answer/corrective focus but also for encliticized focus when preceding a high ip boundary tone (H-). The detachment analysis requires that an entirely new pattern be posited.

Both the early realization (>fHa) and reassignment (L*+H) analyses would be preferred over the detached or floating tone analysis in that they allow the focus high tone (fH) to be aligned to a preexisting tone (*i.e.* either the pitch accent or the AP boundary tone). In all situations seen thus far, the focus high tone (fH) is required to align to one of these preexisting tones in order to be realized.

Reassignment of the focus high tone (fH) is preferable over the other two analyses in that it can derive the location of the pitch maximum without additional stipulations; the rising pitch accent (L*+H) is already described as typically projecting its pitch maximum onto the syllable following the main stress, and a very similar pattern is seen in these cases of putative reassignment. In both the early realization and detachment analyses, it would have to be stipulated that the pitch maximum is realized specifically on the syllable following the main stress.

Lastly, a drawback of both the early realization and reassignment analyses is that the pitch maximum seen in these cases does not exceed the preceding pitch maximum—thus not violating Ha downtrend in the way that both the rising pitch accent (L*+H) and focused high AP boundary tone (fHa) are known to do. The pros and cons of adopting each of the three analyses are summarized below in Table 36.

	Early realization (L*…>fHa…H-)	**Reassignment (L*+H…H-)**	**Detachment (L*…fH…H-)**
Pros	1. Makes use of preexisting tones 2. Does not require positing two structures for encliticized focus	1. Makes use of preexisting tones 2. Derives location of F_0 max without stipulation	
Cons	1. Does not explain lack of downtrend violation 2. Needs stipulation to derive F_0 max location	1. Requires positing two patterns for encliticized focus (L* fHa, L*+H) 2. Does not explain lack of downtrend violation	1. Requires positing two patterns for encliticized focus (L* fHa, L* fH) 2. Does not make use of preexisting tones 3. Needs stipulation to derive F_0 max location

Table 36. Pros and cons of adopting the three analyses of the representation of the focus high tone (fH) on encliticized focus constituents preceding the high ip boundary tone (H-): early realization, reassignment, and detachment.

To truly differentiate the predictions of these three analyses, additional data using encliticized focused constituents of different lengths (*i.e.* to see if the fH tone moves leftward or rightward) and found in a variety of sentence positions (*i.e.* preceded by longer and shorter words to check for downtrend patterns) would be required. In the absence of these data, I leave this matter as an open question at this time. I now discuss the interactions between the rising pitch accent (L*+H) and IP boundary tones (T%).

12.3.2 Sharp rise (L*+H)

The sharp rise focus tonal pattern (L*+H) can cooccur with four of the five IP boundary tones (*i.e.* L%, H%, LH%, HL%). The tritonal dipping IP boundary tone (HLH%) is the only higher level boundary tone that cannot cooccur with a sharp rise (L*+H). The appearance of these four IP boundary tone types following corrective and wh-answer focused constituents suggests that, unlike in the proposed models of previous studies, focused constituents are not restricted to cooccurrence with any particular boundary tone at the IP level.

While the Hayes & Lahiri (1991), Lahiri & Fitzpatrick-Cole (1999), and Selkirk (2006) models of Kolkata Standard Bengali posit an OCP constraint prohibiting a sequence of adjacent H tones, data from the current study indicate that this constraint is not active in Bangladeshi Standard Bengali. Sequences of the rising pitch accent (L*+H) and a high or falling IP boundary tone (H%, HL%) are among those of Bangladeshi Standard Bengali that are not seen in Kolkata Standard Bengali. I provide examples of all sequences below.

As shown in Figure 161, declarative sentences containing focused constituents bear the low IP boundary tone (L%), as do declarative sentences without focused constituents.

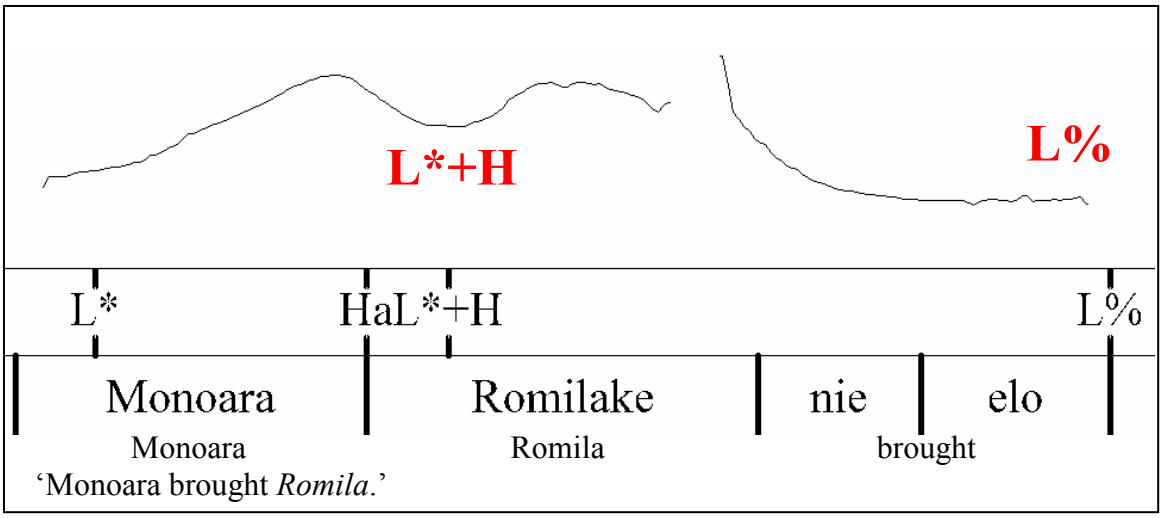


Figure 161. The rising pitch accent (L*+H) most often cooccurs with the low IP boundary tone (L%) associated with default declaratives as well as other sentence types**.** [Na23]

The rising IP boundary tone (LH%), normally associated to default wh-questions, frequently cooccurs with focused elements, as in Figure 162 below.

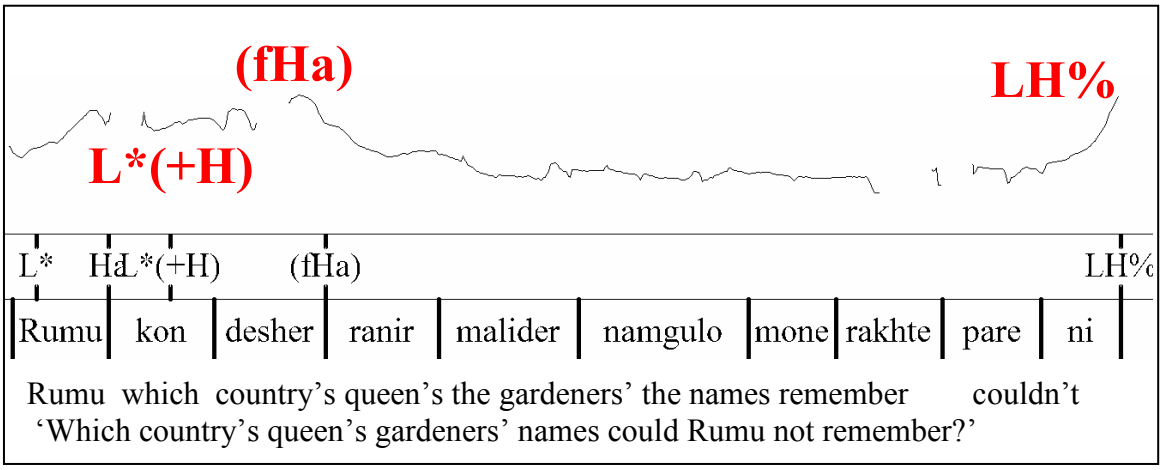


Figure 162. The F_0 maximum of [kon d̪eʃeɹ] 'which country's'—whether associated to the bitonal pitch accent (L*+H) or focused high AP boundary tone (fHa)—cooccurs with the rising IP boundary tone (LH%) associated with default wh-questions. [Tu38]

Of particular interest is the falling IP boundary tone (HL%), which marks default yes/no questions regardless of whether they contain focused constituents. Note in Figure 163 below how the pitch maximum of the rising pitch accent (L*+H)—or alternatively, the focused high AP boundary tone (fHa)—interpolates directly to that of the falling IP boundary tone (HL%).

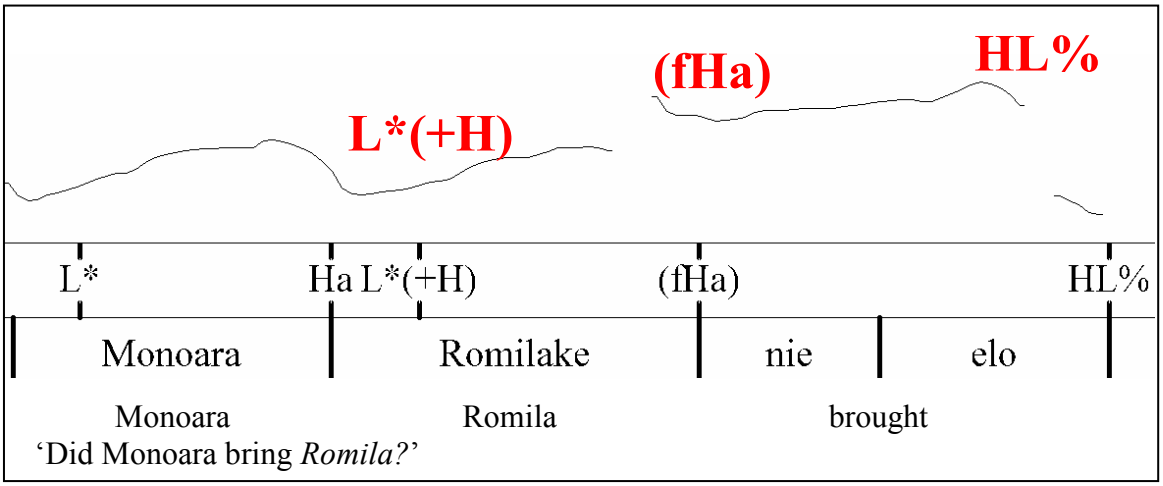


Figure 163. This yes/no question includes two adjacent H targets: the trailing H portion of the rising pitch accent (L*+H) assigned to the focused object [ɹomilake] 'Romila-ACC', and the initial H portion of the falling IP boundary tone (HL%) assigned to yes/no questions. The tonal makeup of the focused word is ambiguous, and in fact could be composed of a focused smooth rise (L*…fHa) or a hybrid of the two sequences (L*+H…fHa). In any case, however, the H targets would be adjacent. [Az02]

A similar pattern is seen when another speaker applied focus to the subject of the same sentence in Figure 164 below.

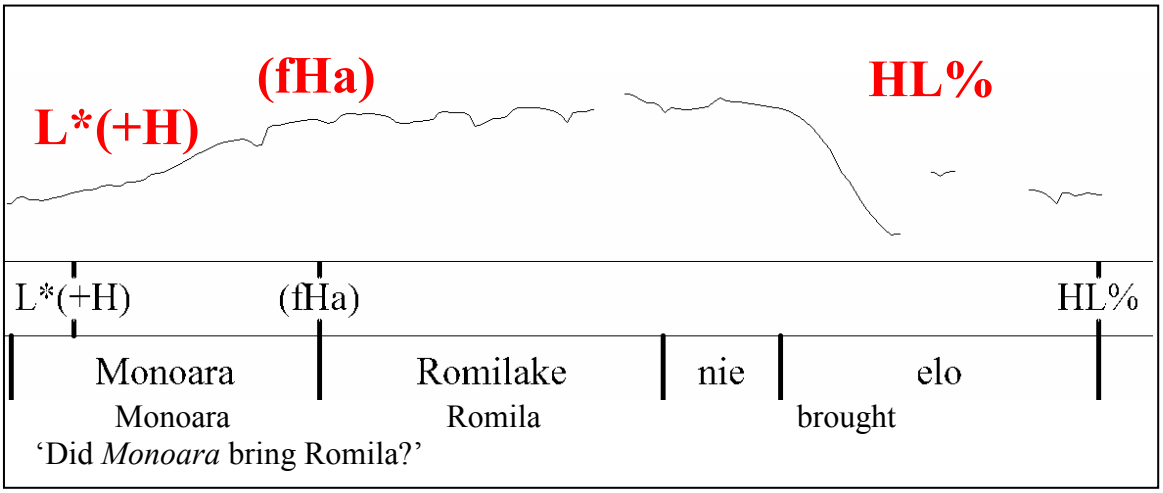


Figure 164. This yes/no question includes two adjacent H targets: the trailing H portion of the rising pitch accent (L*+H) assigned to the focused subject [monoaɹa] 'Monoara', and the initial H portion of the falling IP boundary tone (HL%) assigned to yes/no questions. The tonal makeup of the focused word is ambiguous, and in fact could be composed of a focused smooth rise (L*…fHa), or a hybrid of the two sequences (L*+H…fHa). In any case, however, the H targets would be adjacent.[102] [Fa02]

Furthermore, interrogative sentences that would bear a high IP boundary tone (H%) under neutral focus also bear the same boundary tone when under focus. Note in Figure 165 how the pitch maxima of the rising pitch accent (L*+H) and high IP boundary tone (H%) interpolate.

[102] Note that in this speaker's production, the lengthening of the final syllable [lo] is extreme; although it is not clear in this representation, the falling pitch occurs primarily through the onset [l]. The pitch track during the final [o] is irregular due to creaky voice.

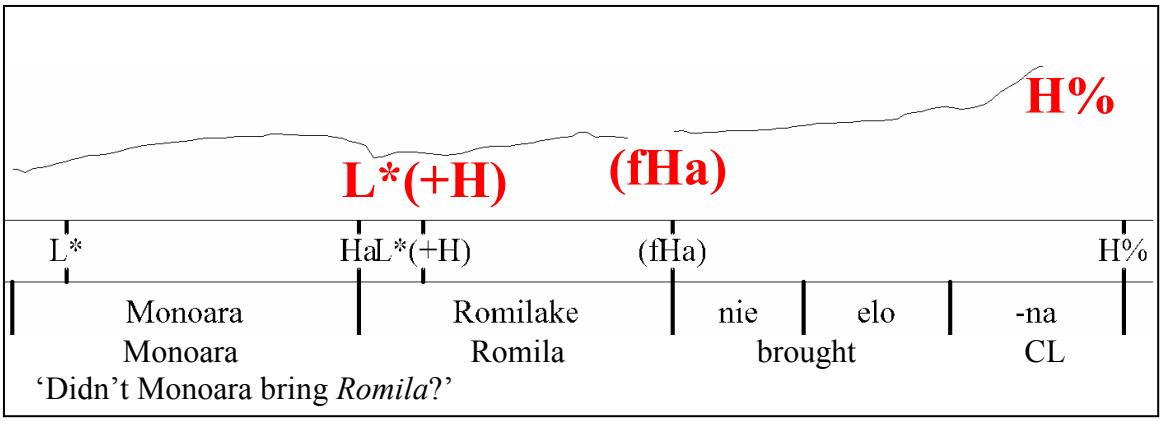


Figure 165. This confirmation yes/no question includes two adjacent H targets: the trailing H portion of the rising pitch accent (L*+H) assigned to the focused object [ɹomilake] 'Romila-ACC', and the high IP boundary tone (H%) assigned to confirmation questions. The tonal makeup of the focused word is ambiguous, and in fact could be composed of a focused smooth rise (L*…fHa), or a hybrid of the two sequences (L*+H…fHa). In any case, however, the H targets would be adjacent. [Re06]

Given the ability of the rising pitch accent (L*+H) to cooccur with four of the five IP boundary tones, including those that begin with H targets, it appears that the variant of Bengali examined in the current study is not affected by the OCP constraint against adjacent H tones in Kolkata Standard Bengali (Hayes & Lahiri 1991; Lahiri & Fitzpatrick-Cole 1999; Selkirk 2006).

13 Conclusion

Focused constituents in Bangladeshi Standard Bengali can bear one of three focus realization patterns, each of which incorporates a surface realization of the focus high tone (fH): focused smooth rise (L*…fHa), sharp rise (L*+H), and focused high pitch accent (fH*). Both the low pitch accent (L*) and the high AP boundary tone (Ha) of the smooth rise AP tonal pattern (L*…Ha) can serve as docking points for the focus high

tone (fH). Words attached to focus enclitics use the boundary tone as the docking point and bear the focused smooth rise tonal pattern (L*…fHa), while corrective and wh-answer focus is realized through the use of the sharp rise AP tonal pattern (L*+H), in which the low pitch accent (L*) serves as a point of adjunction with the focus high tone (fH). Words denoting particularly surprising information bear the focused high pitch accent (fH*), which is the result of fusion between the H target in the smooth fall (H*…La) and the focus high tone (fH). All three realizations of the focus high tone (fH) can be identified in their violation of Ha or H* downtrend, and their tendency to trigger tone deletion or compression to their right. Furthermore, the sharp rise (L*+H) can be distinguished from its non-focused variant (L*…Ha) through observations of the differences in rising contour shapes and F_0 maximum location. When in contact with higher level boundary tones (*i.e.* ip and IP boundary tones), the docking location of the focus high tone (fH) can be changed or modified in such a way that it avoids concurrent boundary tone overriding, either through adjunction with L tones or leftward shift away from H tones. In either situation, it is clear that the OCP constraint against adjacent H tonal targets in Indian (Kolkata) Standard Bengali is not active in Bangladeshi Standard Bengali.

Multi-word single NP focus domains have a tendency to bear a single pitch accent even when the same phrase would be parsed into multiple APs without focus. Complex NPs divide into two realization types: in representative realization, the head of the complex NP (*i.e.* the focus representative)—which is also the rightmost NP—bears the

focus high tone (fH) alone, while in repeated realization, multiple focus high tones (fH) are borne on the focus domain, with as many as one per constituent AP. Following the realization of the focus high tone (fH), AP-level tones (*i.e.* pitch accents and AP boundary tones) are weakened and often deleted (*i.e.* post-focal tone deletion/compression), helping to further draw attention to the focused constituent, as it is the last tonally-marked part of the sentence before the higher level boundary tones (*i.e.* T-, T%).

With the help of the realization of the focus high tone (fH), the violations of Ha and H* downtrend and of concurrent boundary tone overriding, the use of special phrasing, and the optional compression or deletion of post-focal AP-level tones, focused constituents are presumably accentuated in such a way to be easily identified by the listener as the most salient part of the sentence.

CHAPTER FIVE
SUMMARY, DISCUSSION, AND CONCLUSIONS

This chapter provides a brief but comprehensive summary of the tones, prosodic units, and phonological phenomena discussed in Chapters Three and Four, as described in the framework of Intonational Phonology (§14), introduced in Chapter One. A discussion of the similarities and differences between the model proposed in the current study and those of previous studies is presented (§15), along with cross-linguistic comparisons and conclusions of the current study.

14 SUMMARY OF THE INTONATIONAL PHONOLOGY OF BENGALI

By adopting the autosegmental-metrical theory of intonational phonology (Pierrehumbert 1980, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988, among others) as a research framework, the current model of the prosody of Bangladeshi Standard Bengali interprets pitch contours as a series of low (L) and high (H) tonal targets aligned to metrically-prominent positions (*i.e.* pitch accents) and to prosodic boundaries (*i.e.* boundary tones). The model adopts a three-level hierarchy of tonally-marked prosodic units, with the largest phrases spanning clauses and sentences and the smallest phrases spanning individual words or small groups of words. A sentence can be prosodically parsed into these phrases in multiple ways, depending on factors including speech rate, syntax, word length, and focus, among others. The smallest tonally-marked prosodic unit in the phrasing of a sentence is the accentual phrase (AP)—the domain of a

single pitch accent. Each non-focused AP hosts one of the three basic pitch accents—low (L*), high (H*), or rising (L*+H).[103] The two monotone pitch accents (*i.e.* L*, H*) project the AP boundary tone of the opposite tonal target, so that the low pitch accent (L*) projects a high AP boundary tone (Ha) and the high pitch accent (H*) projects a low AP boundary tone (La). This creates either smooth rise (L*…Ha) or smooth fall (H*…La) AP tonal patterns, keeping the number of tones within each basic AP at two. In both of these patterns, the H element is subject to a downtrend pattern—H* downtrend and Ha downtrend—where each AP-level H tone is lower than the preceding one. Exceptions to Ha and H* downtrend can occur due to word length, word type, or focus prominence, as longer words, content words, and focused words bear higher H tones than shorter words, function words, and non-focused words.

As the bitonal pitch accent (L*+H) is by definition composed of two tonal targets, it does not project an AP boundary tone. In the vast majority of cases, the bitonal pitch accent (L*+H) is a marker of focus; the lack of an AP boundary tone corresponding to this pitch accent may in fact be related to the deletion of AP-level tones (*i.e.* pitch accents and AP boundary tones) post-focally (see below).

In addition to these AP tonal patterns are two larger prosodic units: the intonation phrase (IP) and the intermediate phrase (ip). The IP is the largest tonally-marked unit in the system, roughly corresponding to a clause or sentence. Between the AP and the IP is

[103] Although rising pitch accents (L*+H) are typically associated with focused constituents, they can at times be borne on non-focused constituents at the end of a small ip. See Chapter Three, §2.1.3

the intermediate phrase (ip), which marks the edges of smaller phrases, especially in longer sentences. Each of these higher prosodic domains (*i.e.* IP, ip) projects its own boundary tone: the ip boundary tones—high (H-), low (L-), rising (LH-), and falling (HL-)—are realized on the final syllable of the phrase, while the IP boundary tones—high (H%), low (L%), rising (LH%), falling (HL%), and dipping (HLH%)—are realized across all the syllables occurring between the last pitch accent and the edge of the phrase. Because AP tones (Ta), ip tones (T-), and IP tones (T%) all can cooccur at the same point in a sentence, one may expect that phrases can bear multiple tones at the right edge. In truth, however, boundary tones are subject to concurrent boundary tone overriding, which gives priority to the boundary tones of higher prosodic units over other boundary tones projected at the same location. Violations of concurrent boundary tone overriding are always indicative of encliticized focus realization, where the underlying focus high tone (fH) surfaces as a focused high AP boundary tone (fHa) that adjoins to adjacent boundary tones (*e.g.* fHaL-, fHaL%).

Words under narrow focus host the underlying focus H tone (fH), which can surface phonetically in a variety of ways depending on the type of focus (*i.e.* corrective, wh-answer, encliticized, or surprise) and the existence of particular adjacent tones of higher prosodic domains. The focus high tone (fH) attaches to words that would in non-focused contexts bear smooth rises (L*…Ha) either by fusing with the high AP boundary tone (fHa) or by adjoining to the low pitch accent (L*). It also attaches to words that would in non-focused contexts bear smooth falls (H*…La) by fusing with the high pitch accent

(H*). Regardless of the particular realization pattern assumed, the focus high AP boundary tone (fHa) can be identified as an AP-level pitch maximum that exceeds the pitch maximum of the preceding AP, and that often triggers post-focal pitch accent compression or deletion to its right.

The first pattern of focus tone realization (*i.e.* fusion with Ha) is used on words whose focused status is marked segmentally by the focus enclitics -[i]~[j] 'only', 'indeed' or -[o]~[o̯] 'also', 'even', creating a focused smooth rise (L*…fHa). Because the high AP boundary tone (Ha) here is responsible for hosting the focus high tone (fH), it cannot be deleted by the focused constituent's pitch accent or overridden by concurrent boundary tones.

The second pattern of focus tone realization (*i.e.* adjunction to L*) is used on words under corrective focus or serving as the answer to a wh-question. By adjoining the low pitch accent (L*) to the focus high tone (fH), the bitonal pitch accent (L*+H) is generated; words bearing this tone show their pitch maximum (*i.e.* the phonetic realization of fH) within the first two syllables following the pitch accented syllable (which bears a low tone), regardless of the length of the word, indicating that the focus high tone (fH) realized in the rising pitch accent (L*+H) is not aligned to the edge of the word. Due to the earlier realization of the focus high tone (fH) in the rising pitch accent (L*+H), part of the focused constituent itself (especially in longer words) along with the post-focal words can exhibit pitch accent compression or total deletion.

The third pattern of focus tone realization (*i.e.* fusion with H*) is used on words bearing extra emphasis representing particularly surprising or unexpected information. Its realization is the simplest in that it only causes amplification of the high pitch accent (H*) (breaking H* downtrend) and tends to trigger post-focal compression or deletion of AP-level tones, including the loss of the low AP boundary tone (La) normally projected by the high pitch accent (H*).

In addition to the three realization patterns of the abstract focus high tone (fH), focus phrases can be identified by their distinctive phrasing. Multiple-word focused constituents composing a single NP are far more likely than their non-focused counterparts to be parsed as one AP. Complex NPs can exhibit either representative or repeated realization; while many complex NPs only bear one focus high tone (fH) on the rightmost NP (*i.e.* the focus representative), the same complex NP can be produced with multiple focus high tones (fH), with as many focus high tones (fH) as there are embedded NPs within the complex structure. Material to the right of the realization of the focus high tone (fH) is typically (but not always) stripped of its AP-level tones up until the IP-right boundary; often this tone deletion is incomplete, and only results in a compression of F_0 range for the pitch accents and corresponding AP boundary tones.

The recordings used to propose the current intonational phonological model of Bangladeshi Standard Bengali were thoroughly annotated using the transcription system

proposed in the current study: the Bengali Tones and Break Indices (B-ToBI) system. The B-ToBI transcription model is based on the more general ToBI framework of prosodic transcription (Silverman *et al.* 1992, Beckman & Hirschberg 1994), also used for languages such as English (Beckman & Ayers Elam 1997), German (Grice*, et al.* 2005), Japanese (Venditti 1997), Korean (Jun 1999), Serbo-Croatian (Godjevac 2001), and others. In each utterance, a visual representation of the pitch is labeled across four adjoining tiers. The "tones" tier marks tonal events using the labels introduced in Chapters Three and Four, including all pitch accents and boundary tones. The "words" tier includes a Romanized version of each orthographic word in the recording. The "break indices" tier includes numbers representing perceived disjuncture size between the orthographic words in the recording, with larger numbers denoting larger perceived breaks. In addition to the existence and type of boundary tones at phrase edges, disjuncture size can be perceived by various non-tonal cues, including the degree of lengthening of phrase-final syllables, the existence and duration of interphrasal pauses, voice quality, and the presence of segmental alternations (*i.e.* lenition, diphthongization). The fourth tier is reserved for labeling any additional information such as "disfluency" or "reading pause".

Having briefly summarized the material presented in Chapters One through Four, I now proceed to the discussion and conclusions of the current study.

15 DISCUSSION AND CONCLUSIONS

By collecting a corpus of data recorded from a larger number of subjects speaking in a wider range of contexts, the current intonational phonological model of Bangladeshi Standard Bengali supplements the established work of earlier models of Bengali prosody in discovering tonal contours, prosodic units, and phonological patterns not previously described. I begin by discussing phenomena first described in the current model (§15.1), and then move on to comparisons between the current model and models of other languages (§15.2).

15.1 New findings

Findings unique to the current model of Bengali include additional boundary tones (§15.1.1), a third layer of tonally-marked prosodic structure (§15.1.2), and an underlying focus tone (§15.1.3).

15.1.1 Boundary tones

While the number of pitch accents in the current model simply combines the inventories described in previous work (*i.e.* L* and H* from Hayes & Lahiri 1991, Lahiri & Fitzpatrick-Cole 1999, Selkirk 2004; L* and L*+H from Michaels & Nelson 2004), the current model discovers several boundary tones in both Bangladeshi Standard and nonstandard varieties of Northern, Eastern, and Central Bengali dialects. At the accentual phrase (AP) level, the current model finds a low counterpart (La) to the high AP

boundary tone (Ha) identified in previous work as a high phonological phrase or intermediate phrase boundary tone (H-). At the intonation phrase (IP) level, a tritonal dipping boundary tone (HLH%) adds to the inventory first proposed by Hayes & Lahiri (1991). Furthermore, as the intermediate phrase (ip) level has not been distinguished before from the AP and IP levels, all four ip boundary tones (*i.e.* L-, H-, LH-, HL-) are new discoveries.

15.1.2 An additional layer of tonally-marked prosodic structure

Unlike intonational phonological models of Bengali put forth in the past, the current model of the Bangladeshi Standard variety of Bengali adds a third level of tonally-marked phrasing to the prosodic structure, giving an Intonation Phrase (IP), Intermediate Phrase (ip), and Accentual Phrase (AP), with very clear distinctions in their distribution, durational cues, and tonal realization. Despite the fact that the same tonal targets can occur as boundary tones for each of the three phrases, they can be distinguished by their relative height and their particular contour shape (*e.g.* the T- pitch elbow caused by the ip boundary tone locality constraint). In addition, the duration of phrase-final lengthening increases as the level of the phrase rises in the hierarchy. These three levels of tonally-marked prosodic phrasing are also seen in other languages, including Basque (Hualde 1988, Jun 2005), Farsi (Jun 2005, Scarborough 2007, Esposito & Barjam 2007, Arbisi-Kelm 2007), K'iche' (Nielsen 2005), and more recent analyses of Korean (Jun 2007).

15.1.3 Underlying focus tone

One particularly interesting finding of the current study is the underlying focus high tone (fH), which surfaces in three different manners depending on the type of focus applied and the existence and type of adjacent tones. As previously mentioned, the three surface reflexes of the underlying focus high tone (fH) are in complementary distribution: the focus high tone (fH) fuses with the high AP boundary tone (Ha) in encliticized focus constituents, fuses with the high pitch accent (H*) in surprise focus constituents, and adjoins with the low pitch accent (L*) in corrective and wh-answer focus constituents. The relationship between these three "allo-realizations" is clear in that they share particular phonetic properties, and because there may be evidence for some interchangeability between two of the forms given a specific tonal environment. The phonetic properties shared by the three surface reflexes of the underlying focus high tone (fH) are that they exceed the pitch of the previous AP's F_0 maximum and that they tend to trigger some degree of post-focal tone compression or deletion immediately following their realization. Possible evidence for the interchange between two of the surface realizations of the tone (fH) is seen when a focus-encliticized constituent occurs before a high ip boundary tone (H-). In order to maintain a distinction between the H tonal targets of the focus high tone (fH) and the high ip boundary tone (H-), the focus high tone (fH) is shifted leftwards away from the ip boundary. Following the "reassignment analysis" presented in Chapter Four, the focus high tone (fH) attaches to the low pitch accent (L*) instead of to the otherwise-expected high AP boundary tone (Ha). Thus, the reassignment analysis of this leftward shift regards encliticized words in this position as effectively substituting the focus smooth rise (L*…fHa) pattern with the sharp rise pattern (L*+H)

normally seen on wh-answer and corrective focus constituents, further suggesting that the sharp rise (L*+H) and focused smooth rise (L*…fHa) are conditioned variants of each other.

Previous accounts of different varieties of Bengali prosody only looked at one or two types of focus each, and thus either did not collect data on or could not see the relationship between the three focus patterns observed in the current study's corpus. Lahiri & Fitzpatrick-Cole (1999) were the first to differentiate encliticized and non-encliticized focused constituents, and their encliticized pattern of Indian (Kolkata) Standard Bengali (L*…H*) resembles the corresponding patterns seen in the current study of Bangladeshi Standard Bengali. Their non-encliticized pattern (L*…H_P), however, may be specific to Indian (Kolkata) Standard Bengali. Michaels & Nelson's (2004) study of Eastern Bengali identified the rising pitch accent (L*+H) as a marker for wh-answer and corrective focus, and their data and analysis closely resemble those of the current model. However, as Michaels & Nelson (2004) did not study focus enclitics, and as no previous model examined data resembling the current model's focused high pitch accent (fH*), the current study is the first to posit an underlying focus high tone (fH) with three realizations.

15.2 Comparison with other languages

As interest in Indo-Iranian intonation has grown in the past ten years, the current study contributes additional data to the growing body of typological research begun by

Mahanta and others, on languages described as having phonetically weak stress patterns and a rising contour across a phrase spanning about the size of a word. As Bengali is known to possess many phonological similarities with other languages of the Eastern Group of Indic languages (part of the larger Indo-Iranian branch of the Indo-European language family), it would be of no surprise if the prosodic system of Bangladeshi Standard Bengali was found to share the most features with other Eastern Group languages (*e.g.* Assamese, Oriya, Bishnupriya Manipuri, Tripuri, Sylheti, Chittagonian, *etc.*). Parallels can also be drawn with other Indic languages such as Hindi, as well as with more distantly-related languages such as Farsi, French, and Spanish.

Evidence on the intonation systems of two Eastern Group languages suggest that the default smooth rise AP tonal pattern (L*…Ha) of Bengali is also a commonly-recurring prosodic unit in Assamese (Mahanta & Vijayakrishnan 2002, Mahanta 2002a/b) and Oriya (Mahanta 2002a). These studies describe a rising pitch contour whose domain is about the size of the word or somewhat larger, suggesting the existence of a similar AP tonal pattern in all of these languages. Furthermore, Mahanta & Vijayakrishnan (2002) and Mahanta (2002a) suggest a shared feature of left-edge prominence characterized by a low tone (L*) in all three major Eastern Group languages—Bengali, Assamese, and Oriya.[104] Further evidence from Hindi (Harnsberger 1999, 1996; Harnsberger & Judge 1996), an Indic language outside the Eastern Group, suggests that this pattern is more

[104] A few studies, however, such as Das's (2001) study of Tripuri Bengali and Shaw's (1984) study of Indian (Kolkata) Standard Bengali describe quantity-sensitive stress as opposed to fixed initial stress.

widespread in the language branch.[105] The more distantly-related Farsi (Mahjani 2003, Esposito & Barjam 2007, Scarborough 2007, Sadat-Tehrani 2007) also makes use of a rising pitch contour for default phrases, although with final stress on most words (…L+H* Ha), unlike Bengali (L*…Ha). Other AP languages, both within the Indo-European family (*e.g.* French: Jun & Fougeron 1995, 2000, 2002) and outside it, may also pattern like these Indo-Iranian languages—for example, unrelated but geographically-neighboring languages such as Tamil also show strikingly similar patterns with Bengali prosody (Keane 2007), including sequences of rising contours on all non-final content words.

It would be of particular interest to examine additional data from other parts of the Bengali-speaking region and from related and unrelated languages to see how much of the current model of Bangladeshi Standard Bengali can be applied to analyses of other prosodic systems. Considering the diversity in dialect backgrounds even within the current study's corpus of data, it would be appropriate to study what tonal sequences are legal in individual dialects. In addition, testing the perceptual salience of the structures and processes proposed could shed more light on the psychological reality of the current model. The current study's analysis of focus prosody also prompts questions of the interface between the phonetic/phonological realization of focus and semantic theories of the focus feature. I hope that with the data collected for the current study and the

[105] The rhythm of Hindi and Bengali are, however, very different; much of the differences in rhythm stem from the fact that Hindi has a length distinction in both consonants and vowels while Bengali only contrasts length in consonants. Furthermore, Hindi allows a larger number of complex codas, thus allowing stress (or at least what sounds like stress) to appear in different parts of the word.

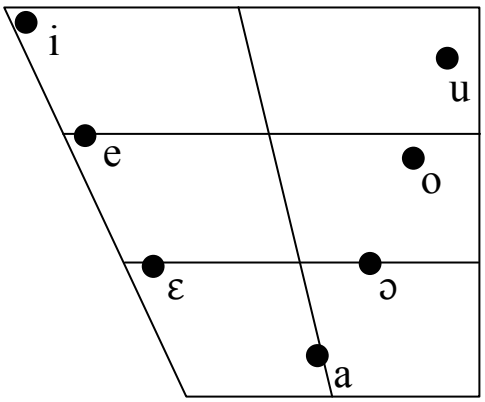
corresponding intonational phonological model as a starting point, other researchers will be able join me in studying the prosody of Bengali from all subfields of linguistic research.

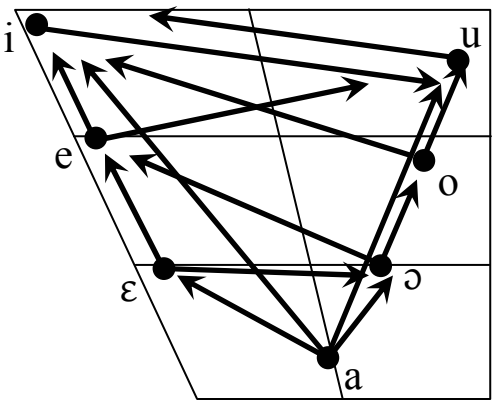
APPENDIX

16 LANGUAGE BACKGROUND

I include a sketch of the phoneme inventory of Eastern Standard Bengali, the style of speech used by all experimental subjects for the current study. The language description is drawn from Khan (2008).

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Velar	Glottal
Plosive	p b bʱ		t̪ d̪ t̪ʰ d̪ʱ		t d tʰ dʱ	k g kʰ gʱ	
Affricate					tɕ dʑ tɕʰ dʑʱ		
Nasal	m		n			ŋ	
Fricative	f		s		ʃ		h
Approx.					ɹ		
Lateral Appr.			l				





17 COMPARISON OF ROMANIZATION SCHEMES

The data collected for the current study are transcribed in IPA, following the transcriptions used in Khan (2008). However, due to font limitations in Praat (Boersma & Weenink 2005), a simplified romanization scheme can be seen in images copied from Praat pitch tracks. This simplified transcription ignores some of the phonemic contrasts in Bengali (*e.g.* /t/ *vs.* /t̪/, /ng/ *vs.* /ŋ/, *etc.*) A table comparing the different romanizations used in the current study and previous studies with the corresponding IPA transcription (drawn from Khan 2008) is provided below.

Bengali	IPA	Praat	S 2006	M&N 2004	L&F-C 1999	H&L 1991	R *et al.* 1966	Ch 1921
অ	ɔ	ô		ô		ɔ	O	ɔ
আ	a	a	a	a	a	ɑ	a	ɑ
ই	i	i	i	i	i	i	i	i
উ	u	u	u	u	u	u	u	u
এ	e	e	e	e	e	e	e	e
ও	o	o	o	o	o	o	o	o
অ্যা	ɛ~æ	ê		ê	æ	æ	E	æ

Bengali	IPA	Praat images	S 2006	M&N 2004	L&F-C 1999	H&L 1991	R *et al.* 1966	Ch 1921
ক	k	k	k	k	k	k	k	k
খ	k^h	kh	k^h	kh	k^h	k^h	kh	kh
গ	ɡ	g	ɡ	g	g	ɡ	g	ɡ
ঘ	$ɡ^ɦ$	gh	$ɡ^h$	gh	g^h	$ɡ^h$	gh	gh
ঙ	ŋ	ng		ng				
চ	tɕ~ts	ch, ts	č	c	tS	č	c	ĉʃ
ছ	$tɕ^h$~s	chh, s	$č^h$	ch, s	tS^h	$č^h$	ch	ĉʃh
জ	dʑ~z	j, dz	ǰ	j, z	dZ	ǰ	j	ʤ
ঝ	$dʑ^ɦ$~z	jh, z	$ǰ^h$	jh, z	dZ^h	$ǰ^h$	jh	ʤh
ট	t	t	ʈ	t’		ṭ	T	ṭ
ঠ	t^h	th	$ʈ^h$	t’h		$ṭ^h$	Th	ṭh
ড	d	d	ɖ	d’		ḍ	D	ḍ
ঢ	$d^ɦ$	dh	$ɖ^h$	d’h		$ḍ^h$	Dh	ḍh
ত	t̪	t	t	t	t	t	t	t
থ	$t̪^h$	th	t^h	th	t^h	t^h	th	th
দ	d̪	d	d	d	d	d	d	d
ধ	$d̪^ɦ$	dh	d^h	dh	d^h	d^h	dh	dh
ন	n	n	n	n	n	n	n	n
প	p	p	p	p	p	p	p	p
ফ	f~p^h	f	p^h	f	p^h	p^h	ph	ph
ব	b	b	b	b	b	b	b	b
ভ	$b^ɦ$	bh	b^h	bh	b^h	b^h	bh	bh
ম	m	m	m	m	m	m	m	m
র	ɹ~ɾ	r	r	r	r	r	r	r
ল	l	l	l	l	l	l	l	l
শ	ʃ	sh	š	ś	S	š	S	ʃ
স	s	s	s	s	s	s	s	s
হ	h	h	h	h	h	h	h	h
ড়	ɹ~ɾ~ɽ	r		r		r	R	ṛ

18 FULL LIST OF STIMULI

18.1 Experiment I – Scripted Production Experiment

57 recordings, 20 speakers

1. মোনোয়ারা রমিলাকে নিয়ে এল।
[monoaɹa ɹomilake nie elo]
Monoara Romila-ACC taken came
‘Monoara brought Romila.’

2. মোনোয়ারা রমিলাকে নিয়ে এল?
[monoaɹa ɹomilake nie elo?]
Monoara Romila-ACC taken came
‘Did Monoara bring Romila?’

3. মোনোয়ারা কি রমিলাকে নিয়ে এল?
[monoaɹa ki ɹomilake nie elo?]
Monoara CL Romila-ACC taken came
‘Did Monoara bring Romila?’

4. মোনোয়ারা রমিলাকে নিয়ে এল কি?
[monoaɹa ɹomilake nie elo ki?]
Monoara Romila-ACC taken came CL
‘Did Monoara bring Romila?’

5. মোনোয়ারা না রমিলাকে নিয়ে এল?
[monoaɹa na ɹomilake nie elo?]
Monoara CL Romila-ACC taken came
‘Didn’t Monoara bring Romila?’

6. মোনোয়ারা রমিলাকে নিয়ে এল না?
[monoaɹa ɹomilake nie elo na?]
Monoara Romila-ACC taken came CL
‘Didn’t Monoara bring Romila?’

7. মোনোয়ারা নাকি রমিলাকে নিয়ে এল?
[monoaɹa naki ɹomilake nie elo?]
Monoara CL Romila-ACC taken came
‘So, apparently, Monoara brought Romila?’

8. মোনোয়ারা রমিলাকে নিয়ে এল নাকি?

[monoaɹa ɹomilake nie elo naki?]
Monoara Romila-ACC taken came CL
‘So did Monoara bring Romila?’

9. মোনোয়ারা তো রমিলাকে নিয়ে এল!
[monoaɹa t̪o ɹomilake nie elo!]
Monoara CL Romila-ACC taken came
‘But Monoara brought Romila!’

10. মোনোয়ারা রমিলাকে নিয়ে এল তো!
[monoaɹa ɹomilake nie elo t̪o!]
Monoara Romila-ACC taken came CL
‘But Monoara brought Romila!’

11. মোনোয়ারাই রমিলাকে নিয়ে এল।
[monoaɹaj ɹomilake nie elo]
Monoara-FOC Romila-ACC taken came
‘*Monoara* brought Romila.’

12. মোনোয়ারা রমিলাকেই নিয়ে এল।
[monoaɹa ɹomilakej nie elo]
Monoara Romila-ACC-FOC taken came
‘Monoara brought *Romila*.’

13. মোনোয়ারাও রমিলাকে নিয়ে এল।
[monoaɹaɔ̃ ɹomilake nie elo]
Monoara-FOC Romila-ACC-FOC taken came
‘Even *Monoara* brought Romila.’

14. মোনোয়ারা রমিলাকেও নিয়ে এল।
[monoaɹa ɹomilakeo nie elo]
Monoara Romila-ACC-FOC taken came
‘Monoara even brought *Romila*.’

15. মোনোয়ারা যে রমিলাকে নিয়ে এল, সেটা তো সবাই জানে।
[monoaɹa dʑe ɹomilake nie elo, ʃeta t̪o ʃɔbaj dʑane]
Monoara CL Romila-ACC taken came, that CL everyone knows
‘But everyone knows that Monoara brought Romila.’

16. মোনোয়ারা রমিলাকে নিয়ে এল যে, সেটা তো সবাই জানে।
[monoaɹa ɹomilake nie elo dʑe, ʃeta t̪o ʃɔbaj dʑane]
Monoara Romila-ACC taken came CL, that CL everyone knows
‘But everyone knows that Monoara brought Romila.’

17. মোনোয়ারা নুন নিয়ে এল।
[monoaɹa nun nie elo]
Monoara salt taken came
‘Monoara brought salt.’

18. মোনোয়ারা মরিচ নিয়ে এল, তাই না? না, না। মোনোয়ারা নুন নিয়ে এল।
[monoaɹa moɹitɕ nie elo, t̪aj na? na, na. monoaɹa nun nie elo]
Monoara peppers taken came, right? No, no. Monoara salt taken came
‘Monoara brought chili peppers, right? No, no. Monoara brought salt.’

19. মোনোয়ারা মাকে নিয়ে এল।
[monoaɹa make nie elo]
Monoara mom-ACC taken came
‘Monoara brought mom.’

20. মোনোয়ারা বাবাকে নিয়ে এল, তাই না? না, না। মোনোয়ারা মাকে নিয়ে এল।
[monoaɹa babake nie elo, t̪aj na? na, na. monoaɹa make nie elo]
Monoara dad-ACC taken came, right? No, no. Monoara mom-ACC taken came
‘Monoara brought dad, right? No, no. Monoara brought mom.’

21. মোনোয়ারা নিনাকে নিয়ে এল।
[monoaɹa ninake nie elo]
Monoara Nina-ACC taken came
‘Monoara brought Nina.’

22. মোনোয়ারা রুমুকে নিয়ে এল, তাই না? না, না। মোনোয়ারা নিনাকে নিয়ে এল।
[monoaɹa ɹumuke nie elo, t̪aj na? na, na. monoaɹa ninake nie elo]
Monoara Rumu-ACC taken came, right? No, no. Monoara Nina-ACC taken came
‘Monoara brought Rumu, right? No, no. Monoara brought Nina.’

23. মোনোয়ারা রুমুকে নিয়ে এল, তাই না? না, না। মোনোয়ারা রমিলাকে নিয়ে এল।
[monoaɹa ɹumuke nie elo, t̪aj na? na, na. monoaɹa ɹomilake nie elo]
Monoara Rumu-ACC taken came, right? No, no. Monoara Romila-ACC taken came
‘Monoara brought Rumu, right? No, no. Monoara brought Romila.’

24. মোনোয়ারা লীনা মামীকে নিয়ে এল।
[monoaɹa lina mamike nie elo]
Monoara Lina Aunt-ACC taken came
‘Monoara brought Aunt Lina.’

25. মোনোয়ারা রুমুকে নিয়ে এল, তাই না? না, না। মোনোয়ারা লীনা মামীকে নিয়ে এল।

[monoaɹa ɹumuke nie elo, t̪aj na? na, na. monoaɹa lina mamike nie elo]
Monoara Rumu-ACC taken came, right? No, no. Monoara Lina Aunt-ACC taken came
‘Monoara brought Rumu, right? No, no. Monoara brought Aunt Lina.’

26. মোনোয়ারা রমিলা মামীকে নিয়ে এল।
[monoaɹa ɹomila mamike nie elo]
Monoara Romila Aunt-ACC taken came
‘Monoara brought Aunt Romila.’

27. মোনোয়ারা রুমুকে নিয়ে এল, তাই না? না, না। মোনোয়ারা রমিলা মামীকে নিয়ে এল।
[monoaɹa ɹumuke nie elo, t̪aj na? na, na. monoaɹa ɹomila mamike nie elo]
Monoara Rumu-ACC taken came, right? No, no. Monoara Romila Aunt-ACC taken came
‘Monoara brought Rumu, right? No, no. Monoara brought Aunt Romila.’

28. মোনোয়ারা রমিলা মামানীকে নিয়ে এল।
[monoaɹa ɹomila mamanike nie elo]
Monoara Romila Aunt-DIM-ACC taken came
‘Monoara brought dear Aunt Romila.’

29. মোনোয়ারা রুমুকে নিয়ে এল, তাই না? না, না। মোনোয়ারা রমিলা মামানীকে নিয়ে এল।
[monoaɹa ɹumuke nie elo, t̪aj na? na, na. monoaɹa ɹomila mamanike nie elo]
Monoara Rumu-ACC taken came, right? No, no. Monoara Romila Aunt-DIM-ACC taken came
‘Monoara brought Rumu, right? No, no. Monoara brought dear Aunt Romila.’

30. বল তো, তুমি কী কী ভাষা জান? আমি বাংলা, হিন্দী, ওড়িয়া, মলয়ালম, তামিল, আর উর্দু জানি।
[bɔlo t̪o, t̪umi ki ki bʰaʃa dʑano? ami baŋla hind̪i oɾia mɔlaɕalom t̪amil aɹ uɹd̪u dʑani]
say CL, you what what language know? I Bengali Hindi Oriya Malayalam Tamil and Urdu know
‘Tell me, what languages do you know? I know Bengali, Hindi, Oriya, Malayalam, Tamil, and Urdu.’

31. বল তো, আজকের দাওয়াতে কে কে আসবে? শুনলাম শুধু রুমু, রেনু, মুনিমা, আমিনা, আর মোনোয়ারা আসবে।
[bɔlo t̪o, adʑkeɹ d̪aɔ̯at̪e ke ke aʃbe? ʃunlam ʃud̪ʰu ɹumu, ɹenu, munima, amina, aɹ monoaɹa aʃbe]
say CL, today-GEN party-LOC who who will.come? I.heard only Rumu, Renu, Munima, Amina, and Monoara will.come
‘Tell me, who all is coming to today’s party? I heard only Rumu, Renu, Munima, Amina, and Monoara are coming.’

32. আমার মনে নেই রুমুর নামের মানে কী।
[amaɹ mone nej ɹumuɹ nameɹ mane ki]
I-GEN mind-LOC there.isn’t Rumu-GEN name-GEN meaning what

‘I don’t remember what Rumu’s name means.’

33. রুমুর নামের মানে কী আমার মনে নেই।
[umuɹ nameɹ mane ki amaɹ mone nej]
Rumu-GEN name-GEN meaning what I-GEN mind-LOC there.isn’t
‘I don’t remember what Rumu’s name means.’

34. মীরার নানা মারা গেলেন বলে আমার নারায়ণগঞ্জে যাওয়া হল না।
[miɹaɹ nana maɹa gelen bole, amaɹ naɹaɛ̯ɔngɔndʑe dʑaɔ̯a holo na]
Mira-GEN grandfather dead went-HON said, I-GEN Narayanganj-LOC going happened not
‘Because Mira’s grandfather passed away, I didn’t end up going to Narayanganj.’

35. যেহেতু মীরার নানা মারা গেলেন, আমার নারায়ণগঞ্জে যাওয়া হল না।
[dʑehet̪u miɹaɹ nana maɹa gelen, amaɹ naɹaɛ̯ɔngɔndʑe dʑaɔ̯a holo na]
because Mira-GEN grandfather dead went-HON, I-GEN Narayanganj-LOC going happened not
‘Because Mira’s grandfather passed away, I didn’t end up going to Narayanganj.’

36. আমার নারায়ণগঞ্জে যাওয়া হল না মীরার নানা মারা গেলেন বলে।
[amaɹ naɹaɛ̯ɔngɔndʑe dʑaɔ̯a holo na, miɹaɹ nana maɹa gelen bole]
I-GEN Narayanganj-LOC going happened not, Mira-GEN grandfather dead went-HON said
‘I didn’t end up going to Narayanganj, because Mira’s grandfather passed away.’

37. আমার নারায়ণগঞ্জে যাওয়া হল না কারণ মীরার নানা মারা গেলেন।
[amaɹ naɹaɛ̯ɔngɔndʑe dʑaɔ̯a holo na, kaɹon miɹaɹ nana maɹa gelen]
I-GEN Narayanganj-LOC going happened not, reason Mira-GEN grandfather dead went-HON
‘I didn’t end up going to Narayanganj, because Mira’s grandfather passed away.’

Each stimulus question in 38-47 is followed by the following response:
রুমু নেপালের রানীর মালীদের নামগুলো মনে রাখতে পারে নি।
[ɹumu nepaleɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni]
Rumu Nepal-GEN queen-GEN gardener-DEF-PL-GEN name-DEF-PL mind-LOC keep-to can hasn’t
‘Rumu couldn’t remember the names of the gardeners of the queen of Nepal.’

38. রুমু কোন দেশের রানীর মালীদের নামগুলো মনে রাখতে পারে নি?
[ɹumu kon d̪eʃeɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni?]
Rumu which country-GEN queen-GEN gardener-DEF-PL-GEN name-DEF-PL mind-LOC keep-to can hasn’t
‘Which country’s queen’s gardeners’ names could Rumu not remember?’

39. রুমু কার মালীদের নামগুলো মনে রাখতে পারে নি?

[ɹumu kaɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni?]
Rumu who-GEN gardener-DEF-PL-GEN name-DEF-PL mind-LOC keep-to can hasn't
'Whose gardeners' names could Rumu not remember?'

40. রুমু কাদের নামগুলো মনে রাখতে পারে নি?
[ɹumu kaḏeɹ namgulo mone ɹakʰt̪e paɹe ni?]
Rumu who-GEN-PL names-DEF mind-LOC keep-to can hasn't
'Whose names could Rumu not remember?'

41. রুমু কী জিনিস মনে রাখতে পারে নি?
[ɹumu ki dʑiniʃ mone ɹakʰt̪e paɹe ni?]
Rumu what thing mind-LOC keep-to can hasn't
'What (thing) could Rumu not remember?'

42. রুমু নেপালের কোন নেত্রীর মালীদের নামগুলো মনে রাখতে পারে নি?
[ɹumu nepaleɹ kon net̪ɹiɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni?]
Rumu Nepal-GEN which ruler-FEM-GEN gardener-DEF-PL-GEN name-DEF-PL mind-LOC keep-to can hasn't
'Which ruler of Nepal's gardeners' names could Rumu not remember?'

43. রুমু নেপালের কোন কর্মীদের নামগুলো মনে রাখতে পারে নি?
[ɹumu nepaleɹ kon koɹmiḏeɹ namgulo mone ɹakʰt̪e paɹe ni?]
Rumu Nepal-GEN which worker-PL-GEN gardener-DEF-PL-GEN name-DEF-PL mind-LOC keep-to can hasn't
'Which workers of Nepal's names could Rumu not remember?'

44. রুমু নেপালের কী জিনিস মনে রাখতে পারে নি?
[ɹumu nepaleɹ ki dʑiniʃ mone ɹakʰt̪e paɹe ni?]
Rumu Nepal-GEN what thing mind-LOC keep-to can hasn't
'What (thing) of Nepal could Rumu not remember?'

45. রুমু নেপালের রানীর কোন কর্মীদের নামগুলো মনে রাখতে পারে নি?
[ɹumu nepaleɹ ɹaniɹ kon koɹmiḏeɹ namgulo mone ɹakʰt̪e paɹe ni?]
Rumu Nepal-GEN queen-GEN which worker-PL-GEN gardener-DEF-PL-GEN name-DEF-PL mind-LOC keep-to can hasn't
'Which workers of the queen of Nepal's names could Rumu not remember?'

46. রুমু নেপালের রানীর কী জিনিস মনে রাখতে পারে নি?
[ɹumu nepaleɹ ɹaniɹ ki dʑiniʃ mone ɹakʰt̪e paɹe ni?]
Rumu Nepal-GEN queen-GEN what thing mind-LOC keep-to can hasn't
'What (thing) of the queen of Nepal could Rumu not remember?'

47. রুমু নেপালের রানীর মালীদের কী জিনিস মনে রাখতে পারে নি?

[ɹumu nepaleɹ ɹaniɹ maliḏeɹ ki dʑiniʃ mone ɹakʰt̪e paɹe ni?]
Rumu Nepal-GEN queen-GEN gardener-DEF-PL-GEN what thing mind-LOC keep-to can hasn't
'What (thing) of the gardeners of the queen of Nepal could Rumu not remember?'

48. আরি? আমরা রমনায় নামলাম কেন?
[aɹi? amɹa ɹɔmnaɛ̰ namlam kɛno?]
INTERJ? we Ramna-LOC descended why
'Wait a second. Why did we get off at Ramna?'

49. আজ দুপুর বেলায় জুম্মার নামাজে শুনলাম মুম্বাইয়ের রেলগাড়ী বোমাহামলায় রমিলার নানা-নানীরা মারা গেলেন।
[adʑ ḏupuɹ bɛlaɛ̰ dʑum:aɹ namadʑe ʃunlam mumbajeɹ ɹelgaɾi bomahamlaɛ̰ ɹomilar nana-naniɹa maɹa gelen]
today early.afternoon time-LOC Friday.congregation-GEN prayer-LOC heard Mumbai-GEN rail-car bomb-attack-LOC Romila-GEN grandfather-grandmother-DEF-PL dead went-HON
'This afternoon at Friday prayers I heard that Romila's grandparents passed away in Mumbai's train bombing.'

50. রুমু নেপালের রানীর মালীদের নামগুলো মনে রাখতে পারে নি।
[ɹumu nepaleɹ ɹaniɹ maliḏeɹ namgulo mone ɹakʰt̪e paɹe ni]
Rumu Nepal-GEN queen-GEN gardener-DEF-PL-GEN name-DEF-PL mind-LOC keep-to can hasn't
'Rumu couldn't remember the names of the gardeners of the queen of Nepal.'

51. আমার মামাও রুমু যেই নামগুলো মনে রাখতে পারে নি, সেই নামগুলো ভুলে গেলেন।
[amaɹ mamaɔ̰ ɹumu dʑej namgulo mone ɹakʰt̪e paɹe ni, ʃej namgulo bʱule gelen]
my uncle-also Rumu which name-DEF-PL mind-LOC keep-to can hasn't, those name-DEF-PL forgotten went-HON
'Even my uncle forgot the names that Rumu couldn't remember.'

52. রুমু যেই নামগুলো মনে রাখতে পারে নি, আমার মামাও সেই নামগুলো ভুলে গেলেন।
[ɹumu dʑej namgulo mone ɹakʰt̪e paɹe ni, amaɹ mamaɔ̰ ʃej namgulo bʱule gelen]
Rumu which name-DEF-PL mind-LOC keep-to can hasn't, my uncle-also those name-DEF-PL forgotten went-HON
'Even my uncle forgot the names that Rumu couldn't remember.'

53. ওই ছেলেটি লম্বা? হ্যাঁ, ওই ছেলেটি লম্বা। কী বলেছ? ওই ছেলেটি মোটা? না না, ওই ছেলেটি লম্বা।
[oj tɕʰeleti lɔmba? hɛ̃, oj tɕʰeleti lɔmba. ki boletɕʰo? oj tɕʰeleti mota? na na, oj tɕʰeleti lɔmba]
that boy-DEF tall? yes, that boy-DEF tall. what have.you.said? that boy-DEF fat? no no, that boy-DEF tall
'Is that boy tall? Yes, that boy is tall. What did you say? That boy is fat? No, no, that boy is tall.'

54. এই আয়নাগুলো মুনিমা মামী কিন্তু পছন্দ করেন না।
[ej aɛ̯nagulo munima mami kin̪tu pɔtɕʰon̪do kɔɹen na]
these mirror-DEF-PL Munima aunt CL like does-HON not
'These mirrors, mind you, Aunt Munima doesn't like them.'

55. এই আয়নাগুলো মুনিমার মামী কিন্তু পছন্দ করেন না।
[ej aɛ̯nagulo munimaɹ mami kin̪tu pɔtɕʰon̪do kɔɹen na]
these mirror-DEF-PL Munima-GEN aunt CL like does-HON not
'These mirrors, mind you, Munima's aunt doesn't like them.'

56. আয়নাগুলো যে মুনিমার, মামী কিন্তু পছন্দ করেন না।
[aɛ̯nagulo dʑe munimaɹ, mami kin̪tu pɔtɕʰon̪do kɔɹen na]
mirror-DEF-PL CL Munima-GEN aunt CL like does-HON not
'The fact that these mirrors are Munima's, mind you, auntie doesn't like that.'

57. এই আয়নাগুলো মুনিমার। মামী কিন্তু পছন্দ করেন না।
[ej aɛ̯nagulo munimaɹ. mami kin̪tu pɔtɕʰon̪do kɔɹen na]
these mirror-DEF-PL. Munima-GEN aunt CL like does-HON not
'These mirrors are Munima's. Auntie, mind you, doesn't like that.'

18.2 Experiment III – Scripted Production Experiment

24 recordings, 1 speaker

Dialect S (Standard Bengali), Set A (SOV order)

1. মোনোয়ারা রমিলাকে মেরে ফেলল।
[monoaɹa ɹomilake meɹe felːo]
Monoara Romila-ACC beaten dropped
'Monoara killed Romila.'

2. মোনোয়ারা রমিলাকে মেরেও ফেলল।
[monoaɹa ɹomilake meɹeo felːo]
Monoara Romila-ACC beaten-CL dropped
'Monoara even *killed* Romila.'

3. মোনোয়ারা রমিলাকে মেরে ফেললও।
[monoaɹa ɹomilake meɹe felːoo]
Monoara Romila-ACC beaten dropped-CL
'Monoara even *killed* Romila.'

4. মোনোয়ারা রমিলাকে মেরেই ফেলল।

[monoaɹa ɹomilake meɹej fel:o]
Monoara Romila-ACC beaten-CL dropped
‘Monoara just *killed* Romila.’

5. মোনোয়ারা রমিলাকে মেরে ফেললই।
[monoaɹa ɹomilake meɹe fel:oj]
Monoara Romila-ACC beaten dropped-CL
‘Monoara just *killed* Romila.’

6. মোনোয়ারা রমিলাকে *মেরে ফেলল*।
[monoaɹa ɹomilake meɹe fel:o]
Monoara Romila-ACC beaten dropped
‘Monoara *killed* Romila.’

Dialect S (Standard Bengali), Set B (OVS order)

1. রমিলাকে মেরে ফেলল মোনোয়ারা।
[ɹomilake meɹe fel:o monoaɹa]
Romila-ACC beaten dropped Monoara
‘Monoara killed Romila.’

2. রমিলাকে মেরেও ফেলল মোনোয়ারা।
[ɹomilake meɹeo fel:o monoaɹa]
Romila-ACC beaten-CL dropped Monoara
‘Monoara even *killed* Romila.’

3. রমিলাকে মেরে ফেললও মোনোয়ারা।
[ɹomilake meɹe fel:oo monoaɹa]
Romila-ACC beaten dropped-CL Monoara
‘Monoara even *killed* Romila.’

4. রমিলাকে মেরেই ফেলল মোনোয়ারা।
[ɹomilake meɹej fel:o monoaɹa]
Romila-ACC beaten-CL dropped Monoara
‘Monoara just *killed* Romila.’

5. রমিলাকে মেরে ফেললই মোনোয়ারা।
[ɹomilake meɹe fel:oj monoaɹa]
Romila-ACC beaten dropped-CL Monoara
‘Monoara just *killed* Romila.’

6. রমিলাকে *মেরে ফেলল* মোনোয়ারা।
[ɹomilake meɹe fel:o monoaɹa]

Romila-ACC beaten dropped Monoara
‘Monoara *killed* Romila.’

Dialect G (Eastern Bengali), Set A (SOV order)

1. মোনোয়ারা রমিলারে মাইরা ফালাইল।
[monoaɹa ɹomilaɹe majɹa falajlo]
Monoara Romila-ACC beaten dropped
‘Monoara killed Romila.’

2. মোনোয়ারা রমিলারে মাইরাও ফালাইল।
[monoaɹa ɹomilaɹe majɹao falajlo]
Monoara Romila-ACC beaten-CL dropped
‘Monoara even *killed* Romila.’

3. মোনোয়ারা রমিলারে মাইরা ফালাইলও।
[monoaɹa ɹomilaɹe majɹa falajloo]
Monoara Romila-ACC beaten dropped-CL
‘Monoara even *killed* Romila.’

4. মোনোয়ারা রমিলারে মাইরাই ফালাইল।
[monoaɹa ɹomilaɹe majɹaj falajlo]
Monoara Romila-ACC beaten-CL dropped
‘Monoara just *killed* Romila.’

5. মোনোয়ারা রমিলারে মাইরা ফালাইলই।
[monoaɹa ɹomilaɹe majɹa falajloj]
Monoara Romila-ACC beaten dropped-CL
‘Monoara just *killed* Romila.’

6. মোনোয়ারা রমিলারে *মাইরা ফালাইল*।
[monoaɹa ɹomilaɹe majɹa falajlo]
Monoara Romila-ACC beaten dropped
‘Monoara *killed* Romila.’

Dialect G (Eastern Bengali), Set B (OVS order)

1. রমিলারে মাইরা ফালাইল মোনোয়ারা।
[ɹomilaɹe majɹa falajlo monoaɹa]
Romila-ACC beaten dropped Monoara
‘Monoara killed Romila.’

2. রমিলারে মাইরাও ফালাইল মোনোয়ারা।

[ɹomilaɹe majɹao falajlo monoaɹa]
Romila-ACC beaten-CL dropped Monoara
‘Monoara even *killed* Romila.’

3. রমিলারে মাইরা ফালাইলও মোনোয়ারা।
[ɹomilaɹe majɹa falajloo monoaɹa]
Romila-ACC beaten dropped-CL Monoara
‘Monoara even *killed* Romila.’

4. রমিলারে মাইরাই ফালাইল মোনোয়ারা।
[ɹomilaɹe majɹaj falajlo monoaɹa]
Romila-ACC beaten-CL dropped Monoara
‘Monoara just *killed* Romila.’

5. রমিলারে মাইরা ফালাইলই মোনোয়ারা।
[ɹomilaɹe majɹa falajloj monoaɹa]
Romila-ACC beaten dropped-CL Monoara
‘Monoara just *killed* Romila.’

6. রমিলারে *মাইরা ফালাইল* মোনোয়ারা।
[ɹomilaɹe majɹa falajlo monoaɹa]
Romila-ACC beaten dropped Monoara
‘Monoara *killed* Romila.’

BIBLIOGRAPHY

Anderson, J. D. (1962). *A Manual of the Bengali Language*. Cambridge University Press.

Arvaniti, Amalia & Mary Baltazani (2000). 'Greek ToBI: A system for the annotation of Greek speech corpora'. *Proceedings of Second International Conference on Language Resources and Evaluation (LREC2000)* **2**: 555–562.

Arvaniti, Amalia & Mary Baltazani (2005). 'Intonational Analysis and Prosodic Annotation of Greek Spoken Corpora'. In Jun, Sun-Ah (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*: 84-117. Oxford University Press.

Barjam, Patrick (2004). 'Intonational Phonology of Porteño Spanish'. UCLA, MA thesis.

Beaver, Davis & Brady Clark (2008). *Sense and Sensitivity: How Focus Determines Meaning*. Malden, MA: Blackwell Publishers.

Beckman, Mary (1986). *Stress and Non-Stress Accent*. Dordrecht, Netherlands: Foris.

Beckman, Mary (1996). 'The parsing of prosody'. *Language and Cognitive Processes* **11**: 17-67.

Beckman, Mary & Gayle Ayers Elam (1994, 1997) 'Guidelines for ToBI Labeling'. Unpublished manuscript. Ohio State University. Version 3 (1997). Available online at http://ling.ohio-state.edu/Phonetics/etobi_homepage.html.

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.

Beckman, Mary, Manuel Díaz-Campos, Julia Tevis McGory, & T.A. Morgan (2002). 'Intonation across Spanish, in the Tones and Break Indices framework'. *Probus* **14**: 9-36.

Birkeland, Harris (1954). 'Stress patterns in Arabic'. In Dybwad, Jacob. *Avhandlinger utgitt av Det Norske Videnskaps-Akademi i Oslo: II. Hist.-filos.* **3**.

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

Brandon, Frank Robert (1975). 'A constraint on deletion in Swahili'. In *The Proceedings of the Sixth Conference on African Linguistics. Ohio State University Working Papers in Linguistics* **20**: 241-259.

Bruce, Gösta (1977). *Swedish Word Accents in Sentence Perspective*. Gleerup, Lund.

Büring, Daniel (1997). *The Meaning of Topic and Focus: The 59th Street Bridge Accent*. Routledge Studies in German Linguistics.

Büring, Daniel & Sun-Ah Jun (2006). 'Focus and Intonation'. UCLA proseminar handout.

Bykova, E. M. (1981). *The Bengali Language*. (English translation: M. E. Feldman) Moscow, USSR: Nauka Publishing House, Central Department of Oriental Literature.

Calhoun, Sasha (2007). 'Predicting Focus through Prominence Structure'. In *Proceedings of Interspeech 2007*. Antwerp.

Carlson, Rolf, Julia Hirschberg, & Marc Swerts (2005). 'Cues to upcoming Swedish prosodic boundaries: subjective judgment studies and acoustic correlates'. *Speech communication*, **46(3/4)**: 326-333.

Chatterji, Suniti Kumar (1921). 'Bengali Phonetics'. *Bulletin of the School of Oriental Studies.* University of London.

Cho, Taehong & Patricia A. Keating (2001). 'Articulatory strengthening at the onset of prosodic domains in Korean', *Journal of Phonetics* **28**: 155-190.

Cowan, J.M. (1993). *Arabic-English Dictionary: The Hans Wehr Dictionary of Modern Written* Arabic. Spoken Language Services.

Das, Shyamal (2001). 'Some Aspects of the Prosodic Phonology of Tripura Bangla and Tripura Bangla English'. Unpublished Ph.D. dissertation, Central Institute of English and Foreign Languages: Hyderabad, India.

Dasgupta, Probal (2003). *Bangla.* In Cardona, G. & Jain, D. (eds.). *The Indo-Aryan Languages.* Routledge.

É.Kiss, Katalin (1998). 'Identification focus and information focus'. *Language* **74**: 245-273.

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**. Los Angeles: University of California.

Ferguson, Charles A. & Munier Chowdhury (1960). 'The Phonemes of Bengali'. *Language.* Linguistic Society of America.

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

Fitzpatrick-Cole, Jennifer (1994). The Prosodic Domain Hierarchy in Reduplication. Ph.D. dissertation, Stanford University.

Fitzpatrick-Cole, Jennifer (1996). "Reduplication Meets the Phonological Phrase in Bengali". *The Linguistic Review* **13**: 305–356.

Fitzpatrick-Cole, Jennifer (1999). 'The Alpine Intonation of Bern Swiss German'. *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS),* San Francisco, CA.

Fougeron, Cécile (1999). 'Prosodically conditioned articulatory variation: A Review'. *UCLA Working Papers in Phonetics* **97**: 1-73.

Fougeron, Cécile & Patricia A. Keating (1997). 'Articulatory strengthening at edges of prosodic domains', *Journal of the Acoustical Society of America* **101**: 3728-3740.

Frota, Sónia (2000). *Prosody and focus in European Portuguese: Phonological Phrasing and Intonation.* New York: Garland.

Godjevac, Svetlana (2001). 'Serbo-Croatian ToBI (SC_ToBI)'. Available online at: http://www.ling.ohiostate.edu/~tobi/

Goldsmith, John A. (1976/79). 'Autosegmental Phonology'. Ph. D. dissertation, Massachusetts Institute of Technology, Garland Publishing Company.

Gordon, Raymond G., Jr. (ed.) (2005). Ethnologue: Languages of the World, Fifteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.

Goswami, Krishnapada (1944). 'Linguistic notes on Chittagong Bengali'. *Indian Linguistics: Journal of the Linguistic Society of India* **8**. Poona, British India: Deccan College Centre of Advanced Study in Linguistics jointly with the Linguistic Society of India.

Grice, Martine, Stefan Baumann, & Ralf Benzmüller (2005). 'German Intonation in Autosegmental-Metrical Phonology'. In Jun, Sun-Ah (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing.* Oxford University Press.

Grierson, George A. (1928). *Linguistic Survey of India*. Calcutta, British India. Available online: http://joao-roiz.jp/LSI/

Gutiérrez-Palma, Nicolás & Alfonso Palma-Reyes (2007). 'On the use of lexical stress in reading Spanish'. In *Reading and Writing*. Netherlands: Springer.

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

Harnsberger, James D. & Jasmeet Judge (1996). 'Pitch range and focus in Hindi'. Presented at the 131st Meeting of the Acoustical Society of America.

Harnsberger, James D. (1996). 'Towards an intonational phonology of Hindi'. Presented at Laboratory Phonology V.

Harnsberger, James D. (1999). 'The role of metrical structure in Hindi intonation'. Presented at the South Asian Analysis Roundtable.

Hayes, Bruce (1989). 'The prosodic hierarchy in meter'. In *Rhythm and meter,* Paul Kiparsky and Gilbert Youmans (eds.), 201–260. Orlando, Florida: Academic Press.

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

Hirst, Daniel and Albert Di Cristo (1998). 'A Survey of Intonation Systems'. In Hirst, Daniel & Albert Di Cristo (eds.) *Intonation Systems: Survey of Twenty Languages*. Cambridge, UK: Cambridge University Press.

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

Huang, Tsan (2001). 'The interplay of perception and phonology in Tone 3 sandhi in Chinese Putonghua'. In Hume, Elizabeth & Keith Johnson (eds.) *Studies on the interplay of speech perception and phonology (OSU working papers in linguistics* **55***)*: 23-42.

Inkelas, Sharon (1989). 'Prosodic constituency in the lexicon'. Doctoral dissertation, Stanford University, Palo Alto.

Itô, Junko & Armin Mester (1992). 'Weak Layering and Word Binarity'. Linguistic Research Center, LRC-92-09, University of California, Santa Cruz.

Jun, Sun-Ah (1993). 'The phonetics and phonology of Korean prosody'. Doctoral dissertation, Ohio State University, Columbus.

Jun, Sun-Ah (1995). 'Asymmetrical prosodic effects on the laryngeal gesture in Korean'. In Connell, B. & A. Arvaniti (eds.) *Phonology and phonetic evidence: Papers in Laboratory Phonology* **IV**: 235-253, Cambridge University Press, England.

Jun, Sun-Ah (1996). 'Influence of microprosody on macroprosody: a case of phrase initial strengthening'. *UCLA Working Papers in Phonetics* **92**: 97-116.

Jun, Sun-Ah (1998). 'The Accentual Phrase in the Korean prosodic hierarchy'. *Phonology* **15**, 189-226. United Kingdom: Cambridge University Press.

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.

Jun, Sun-Ah (2007). 'The Intermediate Phrase in Korean Intonation: Evidence from Sentence Processing'. In Gussenhoven, Carlos & T. Riad (eds.) *Tones and Tunes: Studies in Word and Sentence Prosody*. Berlin: Mouton de Gruyter.

Jun, Sun-Ah & Cécile Fougeron (1995). 'The Accentual Phrase and the Prosodic Structure of French'. *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Stockholm, Sweden, **2**: 722-725.

Jun, Sun-Ah & Cécile Fougeron (2000). 'A Phonological Model of French Intonation'. A. Botinis (ed.) *Intonation: Analysis, Modeling and Technology*. Dordrecht, Netherlands: Kluwer Academic. 209-242.

Jun, Sun-Ah & Cécile Fougeron (2002). 'Realizations of the Accentual Phrase in French Intonation'. J. Hualde (ed.) *Probus* **14**:147-172. A special issue on *Intonation in the Romance Languages.*

Jun, Sun-Ah, Chad Vicenik, & Ingvar Löfstedt (2007). 'Intonational Phonology of Georgian'. *UCLA Working Papers in Phonetics* **106**: 41-57. Los Angeles: University of California.

Kálmán, László, & Noor van Leusen (1993). 'The interpretation of free focus'. In *ILLC Prepublication Series* **CL-93-01**, University of Amsterdam.

Kawasaki, Haruko & Stefanie Shattuck-Hufnagel (1988). 'Acoustic correlates of stress in four demarcative-stress languages'. *Journal of the Acoustical Society of America* **84** (S1): S98.

Keane, Elinor (2007). 'Distribution and alignment of F_0 contours in Tamil'. Presented at the International Congress of Phonetic Sciences, Saarbrücken.

Keating, Patricia, Taehong Cho, Cécile Fougeron, and Chai-Shune Hsu (2003). 'Domain-initial articulatory strengthening in four languages'. In J. Local, R. Ogden, R. Temple (eds.) *Phonetic Interpretation (Papers in Laboratory Phonology* **6***)*, Cambridge University Press: 143-161.

Khan, Sameer ud Dowla (2006). 'Bengali p-word phenomena'. Handout presented at the Proseminar on the Prosodic Word, 5 December. UCLA Ms.

Khan, Sameer ud Dowla (2008). 'Bengali, Bangladeshi Standard'. Ms.

Kreiman, Jody (1982). 'Perception of sentence and paragraph boundaries in natural conversation'. Journal of Phonetics 10, 163-175.

Kristofferson, Gjert (2003). 'The tone bearing unit in Swedish and Norwegian tonology'. In Jacobsen, Henrik Galberg, Dorthe Bleses, Thomas O. Madsen, & Pia Thomsen (eds.) *Take Danish - for instance*: 189-98. Odense: University Press of Southern Denmark.

Ladd, D. Robert (1980). *The structure of intonational meaning*, Indiana University Press, Bloomington.

Ladd, D. Robert (1983). 'In Defense of a Metrical Theory of Intonational Downstep.' In van der Hulst, H. & K. Snider (eds.) *The Phonology of Tone: The Representation of Tonal Register*. Berlin: Mouton de Gruyter. 109-32.

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, Massachusetts Institute of Technology.

Lieberman, Philip (1965). 'On the acoustic basis of the perception of intonation by linguists'. *Word* **21**: 40-54.

Mahanta, Shakuntala (2002a). 'Prominence as a left edge parameter in Oriya, Bangla and Oxomiya – A typological study'. Presented at the 2nd meeting of the Generative Linguists of Hyderabad, India.

Mahanta, Shakuntala (2002b). 'The Prosodic Phonology of Negation in Assamese'. Poster presented at the International conference on the Architecture of Grammar, CIEFL, Hyderabad, India.

Mahanta, Shakuntala & K.G. Vijayakrishnan (2002). 'Low rise on Prominence'. Paper presented at the International conference on the Architecture of Grammar, CIEFL, Hyderabad, India.

Mahjani, Behzad (2003). *An instrumental study of prosodic features and intonation in Modern Farsi (Persian).* MS thesis, Dept. of Theoretical and Applied Linguistics, University of Edinburgh. Available online at: http://www.ling.ed.ac.uk/teaching/postgrad/mscslp/archive/dissertations/2002-3/behzad_mahjani.pdf

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.

Nagahara, Hiroyuki (1994). 'Phonological Phrasing in Japanese'. UCLA PhD dissertation.

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

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

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.

de Pijper, Jan Roelof & Angelien Sanderman (1994). 'On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues'. *Journal of the Acoustical Society of America* **96(4)**: 2037-2047.

Pittayaporn, Pittayawat (2007). 'Prosody of Final Particles in Thai'. Poster presented at the International Congress of Phonetic Sciences Satellite Workshop on Intonational Phonology: Understudied or Fieldwork Languages, Saarbrücken.

Potisuk, Siripong, Jack Gandour, & Mary P. Harper (1996). 'Acoustic correlates of stress in Thai'. In *Phonetica* **53(4)**: 200-220.

Prieto, Pilar, Lourdes Aguilar, Ignasi Mascaró, Francesc Josep Torres, & Maria del Mar Vanrell (2007). 'Catalan Tones and Break Indices'. http://seneca.uab.es/atlesentonacio/cat-tobi/Cat-ToBI.pdf

Ray, Punya Sloka, Muhammad Abdul Hai, & Lila Ray (1966). *Bengali Language Handbook*. Center for Applied Linguistics. Washington, D.C.

Sadat-Tehrani, Nima (2007). 'The Intonational Grammar of Persian'. Poster presented at the International Congress of Phonetic Sciences Satellite Workshop on Intonational Phonology: Understudied or Fieldwork Languages, Saarbrücken.

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

Selkirk, Elisabeth O. (1980). 'Prosodic Domains in Phonology: Sanskrit Revisited'. In Mark Aronoff and Mary-Louise Kean (eds.) *Juncture*. Anma Libri, Saratoga, CA.

Selkirk, Elisabeth O. (1983). 'Phonology and Syntax: the Relation between Sound and Structure'. Massachusetts Institute of Technology Press, Cambridge.

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. (1995). 'Sentence prosody: intonation, stress, phrasing'. In Goldsmith. John A. (ed.) *The handbook of phonological theory*. Blackwell: Cambridge.

Selkirk, Elisabeth O. (1996). 'The prosodic structure of function words'. In Morgan, J. & K. Demuth (eds) *Signal to syntax: bootstrapping from speech to grammar in early acquisition.*

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.

Selkirk, Elisabeth O. & Koichi Tateishi (1989). 'Constraints on Minor Phrase Formation in Japanese'. *Proceedings of the CLS* **24**(1): 316-336.

Shahidullah, Muhammad (2000). *বাংলাদেশের আঞ্চলিক ভাষার অভিধান (A lexicon of Bangladeshi dialects)*. Dhaka, Bangladesh: Bangla Academy.

Shattuck-Hufnagel, Stefanie & Alice E. Turk (1996). 'A prosody tutorial for investigators of auditory sentence processing'. *Journal of Psycholinguistic Research*, **25(2)**: 193-247.

Shaw, Rameshwar (1984). 'Stress-patterns in Bengali and Hindi: A comparative study'. In B. B. Rajapurohit (eds.) *Papers in phonetics and phonology: proceedings of an institute.* Mysore: Central Institute of Indian Languages.

Silverman, Kim, Mary Beckman, John Pitrelli, Mori Ostendorf, Colin W. Wightman, Patti Price, Janet Pierrehumbert, & Julia Hirschberg (1992): 'ToBI: A standard scheme for labeling prosody'. *Proceedings of the Second International Conference on Spoken Language Processing*, 867-870.

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

Suomi, Kari & Riikka Ylitalo (2002). 'Word stress and duration in Finnish'. In *Proceedings of Fonetik, TMH-QPSR* **44** (1). KTH.

Thorsen, N. (1983). 'Two issues in the prosody of standard Dutch'. In Cutler, A. & R. Ladd (eds.) *Prosody: models and measurements*. Heidelberg: Springer: 27-38.

Trager, George Leonard & Henry Lee Smith, Jr. (1951). 'An outline of English structure'. *Studies in Linguistics: occasional papers* **3**. Oklahoma: Norman.

Truckenbrodt, Hubert (1995). 'Phonological Phrases: Their Relation to Syntax, Focus, and Prominence'. Doctoral dissertation, Massachusetts Institute of Technology.

Truckenbrodt, Hubert (2003). 'Variation in p-phrasing in Bengali.' *Linguistic Variation Yearbook* **2**: 259-303.

Venditti, Jennifer J. (1997) *Japanese ToBI labelling guidelines*. http://www.ling.ohio-state.edu/phonetics/J_ToBI/

Viana, Céu & Sónia Frota (2007), Towards a P_ToBI. http://www.fl.ul.pt/dlgr/SonseMelodias/P-ToBI/P-ToBI.htm (collaborators: Isabel Falé, Flaviana Fernandes, Isabel Marcarenhas, Ana Isabel Mata, Helena Moniz & Marina Vigário)

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

Wong, Wai Yi P., Marjorie K. M. Chan, & Mary E. Beckman (2005). 'An Autosegmental-Metrical Analysis and Prosodic Annotation Conventions for Cantonese'. In Sun-Ah Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press.

Woodbury, Anthony C. (1987). 'Rhetorical structure in a central Alaskan Yupik Eskimo traditional narrative'. In Sherzer, J. & A. C. Woodbury (eds.) *Native American Discourse: Poetics and Rhetoric*: 176-239, Cambridge University Press, Cambridge UK.

Wuolle, Aino (2002). *Finnish-English English-Finnish Dictionary*. New York: Hippocrene Books.