Chapter 2: Prosodic Structure in Sylheti and Chokri

2.1. Introduction

Speech is characterized by the potential division of an utterance into units which can be further divided into smaller constituents. One of the major functions of intonation is fragmenting a continuous speech into groups that are separated by perceivable breaks. These breaks are similar to commas and full stops in written text (Hansson, 2003). Intonation divides connected speech into groups or phrases that are deemed complete by the listeners. Crystal (1969) termed such groups 'tone groups' and defined them as "the most readily perceivable, recurrent, maximal functional unit to which linguistic meanings can be attached." They are organized hierarchically where a unit of each level consists of one or more subordinate units (Selkirk, 1984, 1986; Nespor and Vogel, 1986), and such prosodic organization operates at both word and sentence levels (Gussenhoven, 2015). Most languages exploit supra-segmental features like syllable weight, pause, f0, and duration to group constituent speech units in different ways, giving rise to various prosodic domains. These units are syllable, foot, Phonological Word, Phonological Phrase, Intonational Phrase, and Utterance.

Prosodic units above the word level have received significant attention in both theories of syntax-prosody interface as well as intonational phonology. It has been long recognized that prosodic domains at various levels can correspond to syntactic constituent units (Nespor and Vogel, 1986; Selkirk, 1978; Truckenbrodt, 1995.) Edges of prosodic constituents, in particular, have often been seen to align with those of major syntactic units. Studies suggest that prosody can significantly aid in comprehending syntactic structure, as listeners have been shown to accurately pinpoint major syntactic boundaries based solely on prosodic cues (Collier and 't Hart, 1972.) Existing discourses in the syntax-prosody interface belong to two major camps: Indirect Reference Approaches and Direct Reference Approaches. **Indirect reference theories** like the **Prosodic structure theory** hold that prosodic constituent structure is not always identical to syntactic structures. Although representation in prosodic structure makes reference to or is derived from syntactic domains, it is subject to an independent set of prosodic well-formedness constraints (Selkirk, 1978 et seq.; Nespor and Vogel, 1986, Beckman and Pierrehumbert, 1986; Pierrehumbert and Beckman, 1988; Hayes, 1989;

Inkelas, 1990; Truckenbrodt, 1995, 1999; Ladd, 1996 2008; Elordieta, 2007a; Frota, 2000; Seidl, 2001; Dobashi, 2003; Gussenhoven, 2004). Empirical evidence for this approach comes from sandhi processes as well as phrase level intonation being influenced by factors like speech rate and discourse type (Bennett and Elfner, 2019). **Direct reference approaches** are based on the assumption that prosodic domains are directly derived from syntactic structures and systems of syntactic spell-out instead of prosodic structure. They are of the view that the domains that constitute prosodic structure are just syntactic domains, which implies that prosodic structure is redundant in the process of grammatical derivation. According to this view, phonological processes that are conditioned by domains *directly* refer to the syntactic structure (Cooper and Paccia-Cooper, 1980; Gee and Grosjean, 1983; Kaisse, 1985; Wagner, 2005; Chomsky, 2001).

Selkirk's (2011) Match Theory has been an influential addition to the works on syntax-prosody mapping. Leaning towards indirect reference, Match theory assumes that syntactic trees containing CPs, XPs, and X^0s can be mapped onto prosodic trees made up of phonological units like intonational phrases (ι), phonological phrases (φ), and prosodic words (ω).

In this mapping, syntactic inputs S (the syntactic domains) are fed to the phonological component that eventually produces prosodic outputs P (phonological domains).

CP	ι
XP	φ
X0	ω

These syntactic-prosodic constituent pairs constitute the Match constraint template Match (α,β). The mapping constraints can be bidirectional; in SP mapping, α is syntactic and β prosodic, and for PS mapping, α is prosodic and β syntactic. The following primary Match constraints govern them:

	Syntax → Prosody	*Prosody → Syntax*
MatchClause	Match(CP,ι)	Match(ι,CP)
MatchPhrase	Match(XP,φ)	Match(φ,XP)
MatchWord	Match(X0,ω)	Match(ω,X0)

Prosodic domains are identified by features like the placement of pitch accents and boundary tones, the scope of segmental processes (often referred to as "sandhi"), and

various phonetic markers, including pauses, duration, and the scaling of tonal accents (Pierrehumbert, 1980; Nespor and Vogel, 1986; Selkirk, 1980, 1986). In prosodic hierarchy theories, the criteria for defining the prosodic units are based on their indirect reference to the syntactic units and their role as the domain of application of segmental phonological processes like sandhi. In intonational phonology, they are defined in terms of their phonetic forms, which are determined by their suprasegmental features like intonational events, pitch reset, final lengthening, and pauses. Pierrehumbert and Beckman's (1988) model of intonational phonology has been the most developed model adopted in analyses of many languages. Like prosodic hierarchy theories, this model also assumes that all constituents at one level of the phonological hierarchy are exhaustively parsed into constituents of the following level. However, Pierrehumbert and Beckman (1988) hold that the number of levels in the prosodic hierarchy is not uniform for all languages. It is rather language-specific. For instance, three prosodic levels above the word are identified for Japanese, the AP, the ip, and the IP, while English has only two, the ip and the IP.

Utterance is the largest prosodic unit, which is defined in prosodic phonology as the group of Intonational phrases that correspond to syntactic constituents dominated by the same single node X^n in the syntactic structure. The semantic content of the constituent words is also crucial in the identification of U.

In prosodic phonology, the intonational phrase is the domain throughout which intonational pitch contours are present, and they are also the locations of potential rule-governed junctures or pauses in languages like English (Nespor and Vogel, 1986). According to Jun (2005), cross-linguistically Intonation Phrase is the highest prosodic unit marked by intonation (IP). The Intonational Phonology of Pierrehumbert (1980) defines IP through the presence of boundary tones H% or L%. A phonological Phrase is a phonological unit that is very frequently considered to be essential for the formulation of phonological rules. It includes a phrasal head and whatever precedes (in head-final languages) or follows (in head-initial languages) it within the same phrase (Nespor and Vogel, 1986.) Selkirk and Tateishi (1988) distinguishes between phonological phrases in terms of levels in the hierarchical structure- minor and major phrases. The intonational units that correspond to the Phonological Phrase are the Intermediate Phrase (Major Phrase) (Beckman and Pierrehumbert, 1986), or the Accentual Phrase (Minor) (Selkirk and Tateishi, 1988). In English, The Intermediate Phrase is the domain of downstep,

which is bounded by an edge tone, i.e., phrase accent, H- or L- and/or final lengthening. In several languages like Japanese, French, and Korean, there is one more intonationally marked prosodic unit that contains more than one word. They have either or both pitch accents and accent boundary tones, constituting a phrase smaller in scope than an intermediate phrase. In most (non-tonal) south Asian languages, Accentual Phrases are distinctly marked with phrasal tones (Keane, 2014, for Tamil; Khan, 2014, for Bengali).

The phonological units identified in the intonational phonology have been integrated into the ToBI (Beckman et al., 2005) systems developed across languages for annotation of intonation in utterances. In addition to using the notations prescribed in the intonational phonology for pitch accents, phrase accents, and boundary tones, it also allows signifying boundaries of various levels of prosodic units with break indices, for example, break index 4 for IP boundary, 3 for ip Boundary, 2 for AP boundary, 1 for PW boundary and 0 for clitic boundary for a hypothetical language that has IP, ip as well as AP in its prosodic structure.

In intonational phonology, such phonological constituents are marked with intonational features like intonational tones (Pitch Accents and Boundary tones) and duration. Pierrehumbert and Beckaman's (1986) model of intonational phonology introduced an additional unit named intermediate phrase placed between IP and PP. While most intonation-only languages are seen to have various units marked with pitch accents, phrase accents, and/or boundary tones, only a little has been studied about prosodic units in tonal languages. Many African tone languages show two levels of intonationally marked prosodic constituents: the Phonological Phrase and the Intonational Phrase (Downing and Rialland, 2017). In many instances, the motivation for prosodic phrasing comes from the prosodic phonology approaches wherein the prosodic domains are conditioning factors for segmental or tonal processes. Utterances in these languages are sometimes parsed into smaller prosodic constituents based on phonological processes bound by the edges of some kinds of phrases. In Akan, phonological phrase boundary blocks the application of regressive vowel harmony that can exist across prosodic words (Kügler, 2015). High Tone Spreading (HTS) and Falling Tone Simplification (FTS) are tonal processes that demarcate prosodic domains in Bàsàá (Makasso, Hamloui, and Lee, 2017). Shingazidja is known to have domain-sensitive tone shift rules that cannot apply beyond a phonological phrase that aligns with maximal syntactic phrases (Patin, 2017). Bemba is another African tone language where H tone spreading is bound by the

Phonological Phrases (Kula and Bickmore, 2015). However, apart from phrase internal processes, prosodic domains in tonal languages can also be defined in terms of intonation features like pauses as seen in Kɔnni (Cahill, 2017) or the final lengthening observed in Shekgalagari (Hyman and Monaka, 2008), pitch reset or final lowering reported in the Akan language (Kügler, 2017).

2.2. Experimental Methodology

2.2.1. Materials:

Two datasets with 20 complete sentences each for both Chokri and Sylheti were prepared for this study through consultation with the native speakers. The scripted sentences followed the default word order of subject-object-verb and had sequences of different lexical tones. Three types of sentences were included to examine the possible intonational units- Simple declaratives of various lengths, compound sentences, and complex sentences. These sentences were representative of natural utterances of the language, and they were formed in a way that included different tonal specifications. The data was presented in Roman orthography as speakers of both languages informed competence in reading the same. Since the speakers are not accustomed to reading scripts with tonal notations, no tone marks were used in the corpus. Instead, English translations of the target utterances facilitated production of words with intended tones. Each speaker was asked to repeat the whole set 5 times, ensuring a considerable gap between each session. The total number of token sentences considered for the analysis are 2000 (20 sentences x 2 languages x 10 subjects x 5 repetitions =2000).

2.2.2. Speakers:

Five native speakers from each language, Sylheti and Chokri, participated in the production experiment conducted for this study. All five Chokri speakers were from the Thipuzu village of the Phek district in Nagaland, India, and were aged between 18 and 30 years. Two of them were female, and three were male. Other than Chokri, they were proficient in the use of English, Nagamese, and Tenyidie. The five Sylheti native speakers (four female, one male) were from northern Tripura, all between the ages of 20 and 35. The informants are also well versed in English and Standard Bengali.

2.2.3. Data Recording Procedures and Pre-processing

The Sylheti speech data used for the experiment was recorded in a soundproof booth in the Department of Linguistics and Language Technology, Tezpur University. The Chokri data was recorded during a field visit to Thipuzu village, which was conducted in an empty room, ensuring minimal background noise and echo effects. The speakers were shown each sentence on a laptop screen and were instructed to produce them in a manner that resembles natural speech. To eliminate possible background noise, a head-mounted microphone connected to a hand-held recording device was used. The recordings were digitized at a sampling frequency of 44.1 kHz and 32-bit resolution. The sentences were randomized and put across the corpus to elicit five iterations of all the sentences from each speaker. After the completion of the recording, the data was transferred through a USB cable to a PC for further processing.

The recordings were first manually segmented at sentence level on Praat (Boersma and Weenink, 2024), and individual sound files of each sentence were extracted using a Praat script. In the second stage, each sentence was segmented and labeled at the syllable level through visual observation and careful listening. After detailed examination, the sentences were further annotated in four TextGrid tiers. **The data annotation** format was customized for the two languages in order to capture each of their distinct intonational properties. The sound waves, spectrograms, and pitch tracks, along with the labeled TextGrids, were extracted using a modified version of *Create Pictures with Tiers v.6. Praat script* **(**Elvira García, 2022). For Sylheti, four different tiers are marked in the following structure:

Tier 1: marks phrasal tones following the AM framework (Pierrehumbert & Beckman, 1988) and the ToBI model (Silverman et al., 1992; Beckman et al., 2005).

Tier 2: IPA transcriptions of the words constituting the utterance, the H tone is marked with diacritic.

Tier 3: glossing of words.

Tier 4: relative strengths between speech units are marked with break indices, as per the ToBI conventions.

For Chokri, the annotation structure was:

Tier 1: marks the underlying lexical tones

Tier 2: IPA transcriptions of the syllables constituting the utterance

Tier 3: The glossing of words.

Tier 4: break indices, as per the ToBI conventions.

The representative diagrams of the spectrograms were drawn using a modified version of the script *Create Pictures with Tiers v.6. Praat script* **(**Elvira García, 2022), customized for this analysis.

2.2.4. Quantification and Statistical Analysis

In order to perform quantitative analysis, values of the relevant acoustic parameters, viz., f0, duration, and intensity, were extracted using the Prosody Pro Praat Script (Xu, 2013). Measurement of both average and time normalized f0 at 10 points of each syllable was done to have information about pitch level and pitch movement on different points of the f0 contour. The extracted data was stored in .csv files for preprocessing and subsequent statistical analysis. Statistical tests like repeated measure ANOVA were conducted using functions from the statsmodels library on Python. For visualization of quantitative analysis, graphs were produced using the smatplotlib library in Python.

2.3. Results and Analysis: Prosodic Structure in Sylheti

2.3.1. Structure of Prosodic Domains in Sylheti

In this production experiment, we examined different levels of prosodic domains present in Sylheti. Our analysis proposes four levels of prosodic structure in the language: PrWd, AP, ip, and IP. Of these, the three higher domains are marked by phrasal tones, relative juncture, and boundary tones. The motivation for PrWd as a tonally-relevant prosodic unit comes not from the presence of an intonational tone per se but from the process of tonal polarity observed in the language, the application of which is limited to grammatical words. Despite being a tonal language, Sylheti does exhibit intonational pitch accents that participate in the AP unit. In addition to that, Sylheti's tonal inventory also contains boundary tones that occur at the edges of all three prosodic units.

2.3.2. Prosodic Words

The motivation for our identification of the PrWd as a tonally relevant unit of Sylheti prosodic structure has been drawn from the previous findings by Gope (2016) and Mahanta and Gope (2018), who argued that both derivational and inflectional affixations in Sylheti obey rules of tonal polarity. In the case of nouns, the suffixes that attach

themselves to the roots are underlyingly toneless. When affixed, they are realized with a tone opposite to the one carried by the root, i.e., if the root carries an underlying L tone, the suffix receives an H tone and vice versa. Consider the following examples in **Figure 2.1 and 2.2** (reproduced from Gope, 2016; Mahanta and Gope, 2018).

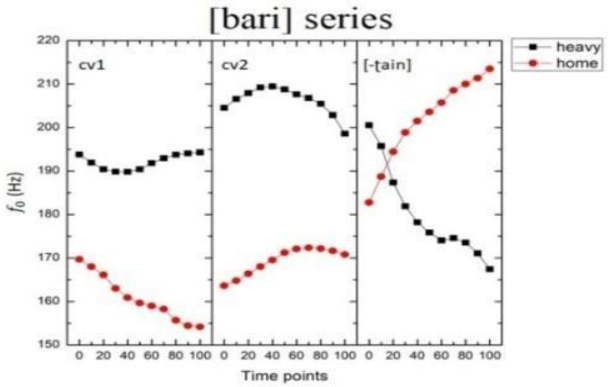

Figure 2.1: Averaged non-normalized pitch track of the word [bàri] 'home' and [bári] 'heavy' produced by 9 speakers with the inanimate plural suffix [-ʈain]. (The details can be seen in Gope, 2016; Mahanta and Gope, 2018).

The verbal inflectional affixes, on the other hand, underlyingly carry an H tone that requires that the root's tonal specification be L regardless of its underlying status.

The boundaries of grammatical words bound tonal polarity in this language. This strict adherence to tonal polarity serves as motivation for considering stem+suffixes as one prosodic unit (PrWd).

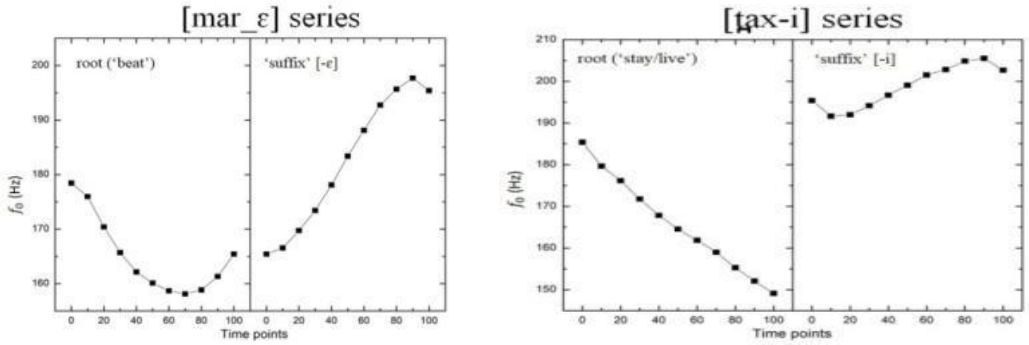

Figure 2.2: averaged normalized pitch track of all the tokens of the words [mar] with the 3rd person indefinite suffix [ɛ] (left panel) and [ʈax] with the 1st person indefinite suffix [i] (right panel) by 9 speakers (the details can be seen in Gope, 2016, and Mahanta and Gope, 2018).

2.3.3. Accentual Phrase

Apart from PrWd, we propose three additional levels of prosodic structure: the Accentual Phrase (AP), intermediate phrase (ip), and intonational phrase (IP). These prosodic categories in Sylheti are marked by distinctive phrasal tones (both pitch accents and

boundary tones), different degrees of pauses between phrases, and pitch resetting in some cases.

The AP is the lowest tonally-marked intonational unit in Sylheti, a feature similar to standard varieties of Bengali (Hayes and Lahiri, 1991; Jun, 2005; Khan, 2014). It serves as the domain for two types of phrasal tones: (i) a pitch accent (T*), which consistently appears on the initial syllable, and (ii) the AP boundary tone (Ta), which consistently appears on the final syllable. In pre-nuclear positions, two types of pitch accents are possible: L*, i.e., a local pitch valley, and L*H, a low tone immediately followed by a rise to a high target. All pre-nuclear APs have a high right boundary tone (Ha). The interpolation of the phrase-initial pitch accent to the Ha boundary tone results in a rising contour. The entire utterance is divided into multiple consecutive APs, producing a repeatedly rising f0 contour (the *RRC* reported in Khan, 2019). In many instances, APs consist of singular PrWds, as shown in **Figure 2.3,** where the subject and object each constitute individual APs. While the PA always attaches to the initial syllable, the H target of the Ha boundary tone may be reached as early as the AP's penultimate syllable. APs may also contain more than one PrWd, a lexical head, along with neighbouring functional elements, as in **Figure 2.4**.

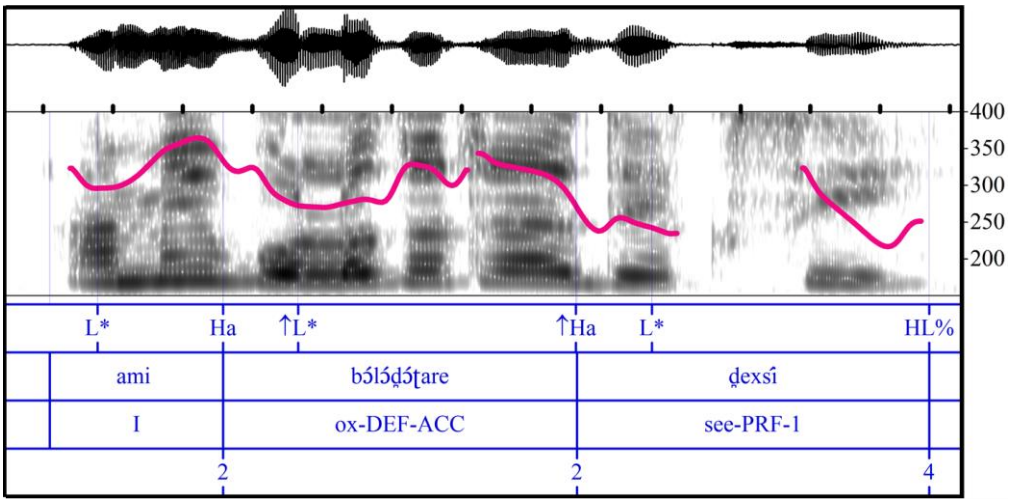

Figure 2.3: pitch track of the sentence [ami bɔlɔɖɔ ʈare ɖexsí] 'I have seen the ox' produced by a female speaker.

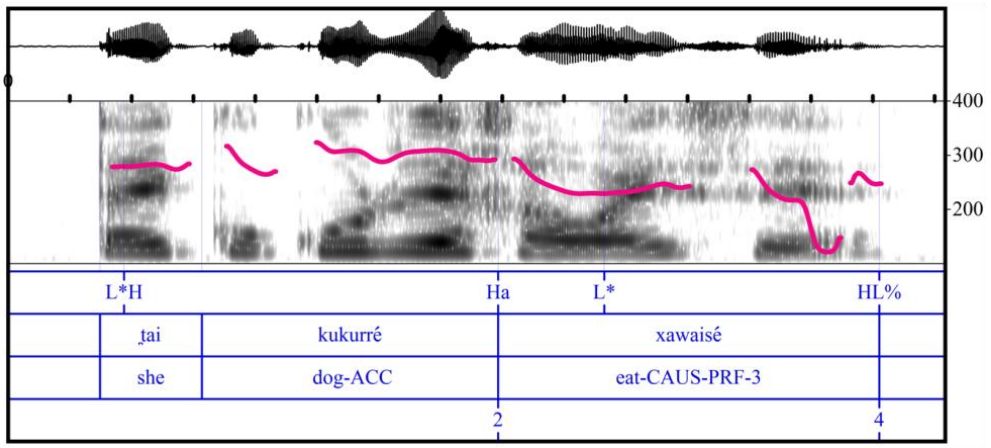

Figure 2.4: pitch track of the sentence [t̪ai kukurré xawaisé] 'she has fed the dog' produced by a female speaker.

The most frequently occurring type of AP in Sylheti, however, is the *recursive AP*; a feature reported in other South Asian languages as well (Khan, *to appear*); realized in the (L*/L*H)-Ha-Ha form. After the Ha tone of the initial AP, the f0 contour continues to be high till the end of the following AP, skipping the projection of the second or third PA. Such recursive APs mostly comprise lexical words followed by functional elements, as shown in **Figure 2.5**. It is to be noted that the domain of recursive APs does not always correspond to syntactic phrase boundaries.

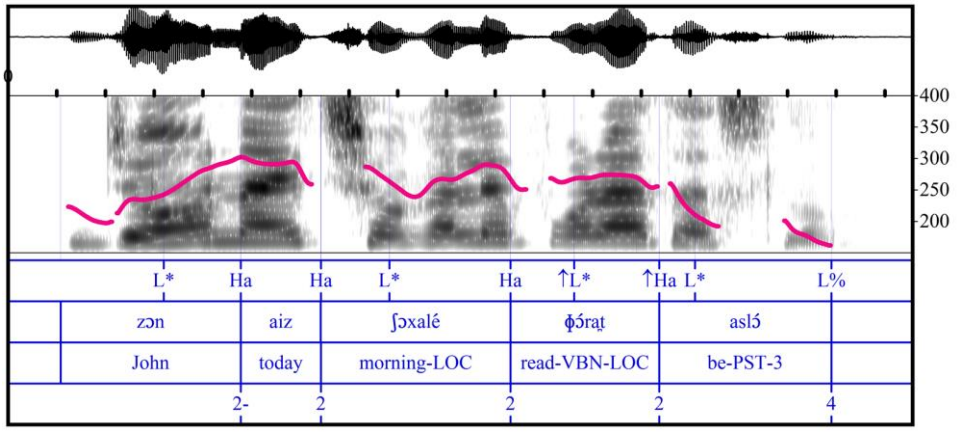

Figure 2.5: f0 contour of the sentences [zɔn aiz ʃɔxalé ɸɔ́rat̪ aslɔ́], 'John was studying this morning' produced by a female speaker

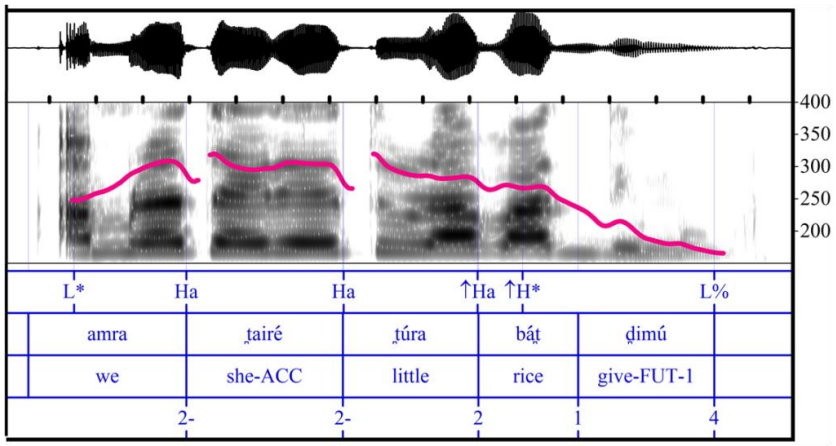

Figure 2.6: pitch track of the sentence [amra t̪airé t̪úra bát̪ d̪imú] 'We will give her some rice' produced by a female speaker

APs in nuclear positions contain one of three probable Pitch Accents: L*, H*, L*H. As nuclear APs are the final APs in an utterance, their boundary tones are overridden by the boundary tones of higher intonational units, i.e., ip or IP.

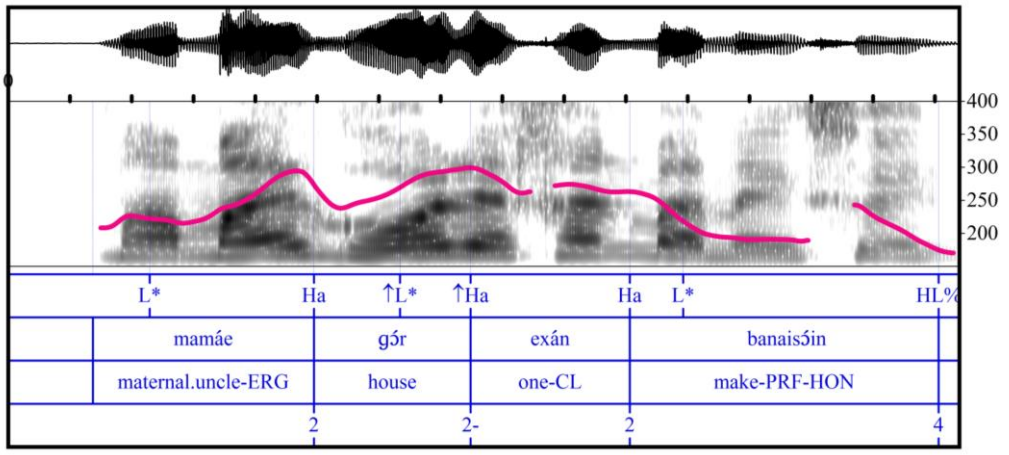

Figure 2.7: pitch track of the sentence [mamáe gɔ́r exán banaisɔ́in], '(my maternal) uncle has built a house' produced by a female speaker

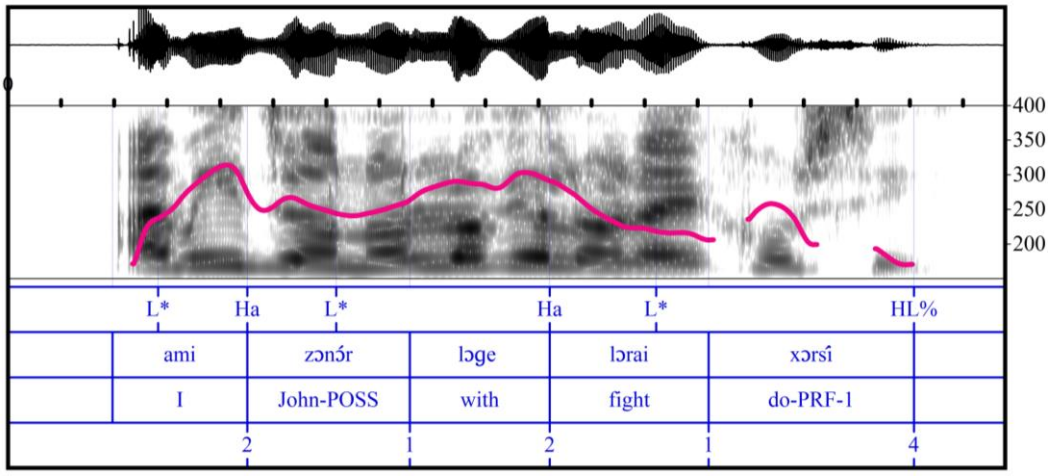

Figure 2.8: pitch track of the sentence [ami zɔnɔ́r lɔgé lɔrai xɔrsí] 'I've fought with John' produced by a female speaker

The derived pitch tracks for utterances used in this study exhibit the presence of a downtrend that affects the AP H tones. This entails that, as a phonetic effect, the f0 of the H tone of APs occurring later in an AP succession will have a lower f0 than that of the preceding AP(s). In **Figure 2.9**, the f0 levels reached by each AP's H target are 355 Hz, 319 Hz, 302 Hz, and 277 Hz[1]. The evidence for AP downstep comes from the analysis of time normalized average f0 of all iterations of the sentence produced by the five speakers as seen in **Figure 2.10.** Each successive Ha is realized with a comparatively lower f0 than that of the previous one.

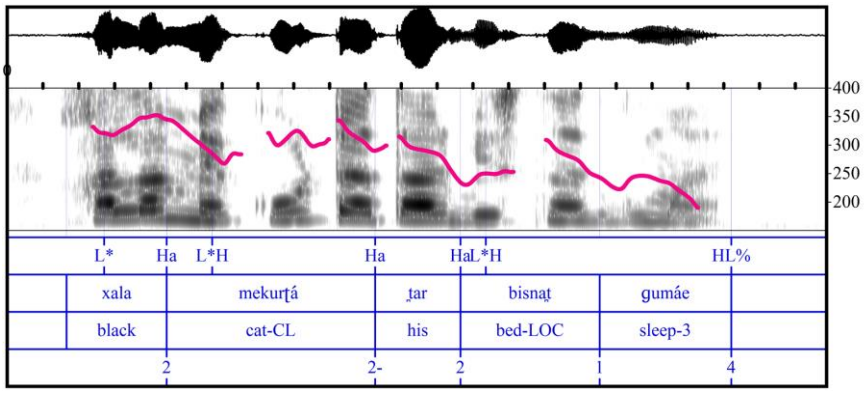

Figure 2.9: pitch track of the sentence [xala mekurʈá ʈar bisnaʈ gumáe] 'the black cat sleeps on his bed' produced by a female speaker

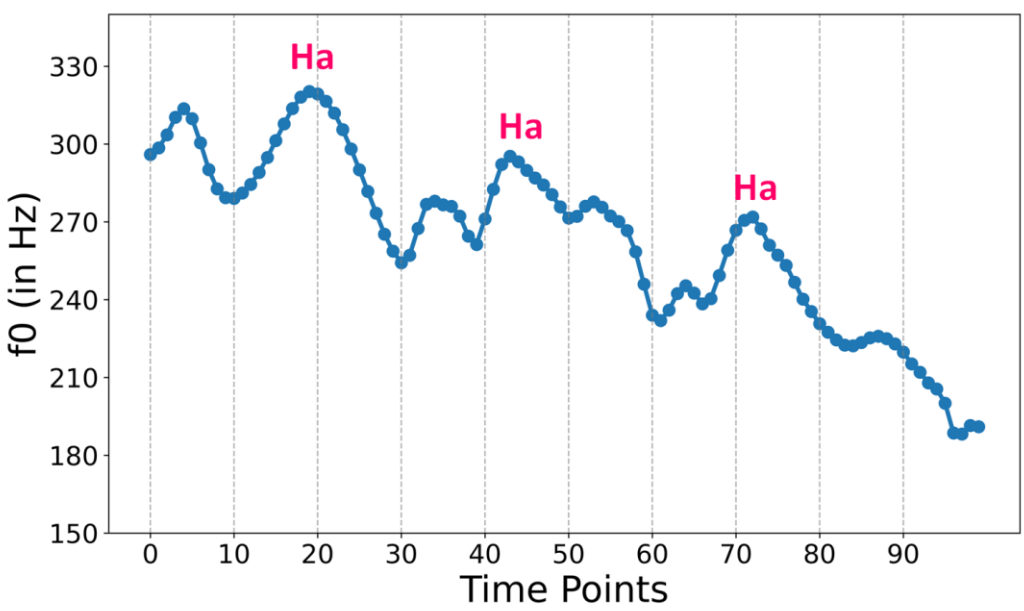

Figure 2.10: time normalized average f0 of the sentence [xala mekurʈá ʈar bisnaʈ gumáe] 'the black cat sleeps on his bed' of all tokens by five speakers. The Ha AP boundary tones are labeled on the diagram. Each 10 points in the x-axis corresponds to one syllable.

[1] The average f0 values of each word was obtained using the 'get pitch' feature on PRAAT

2.3.4. Interaction of AP phrasal tones and Lexical tones

At first glance, pitch contours in Sylheti seem to be made up entirely of intonational tones, resembling related languages such as Bangladeshi Standard Bengali (Khan, 2014), Kolkata Standard Bengali (Hayes and Lahiri, 1991), and other South Asian languages (Khan, 2019). The repeated rising contour of consecutive APs strengthens this assumption. However, a closer look into the f0 realization of words with contrastive lexical tones presents a different picture, setting Sylheti apart from its more widely-described neighbors. In order to examine this tone-intonation interaction, we considered sentences with different lexical tone sequences, finding that the manifestation of lexical tone varies depending on the intonational position. For example, sentence-initial AP rising contours completely override the underlying lexical tone specification for the words contained within it. This results in rising f0 tracks with similar scaling for words with L, H, HL, and LH lexical tones.[2]

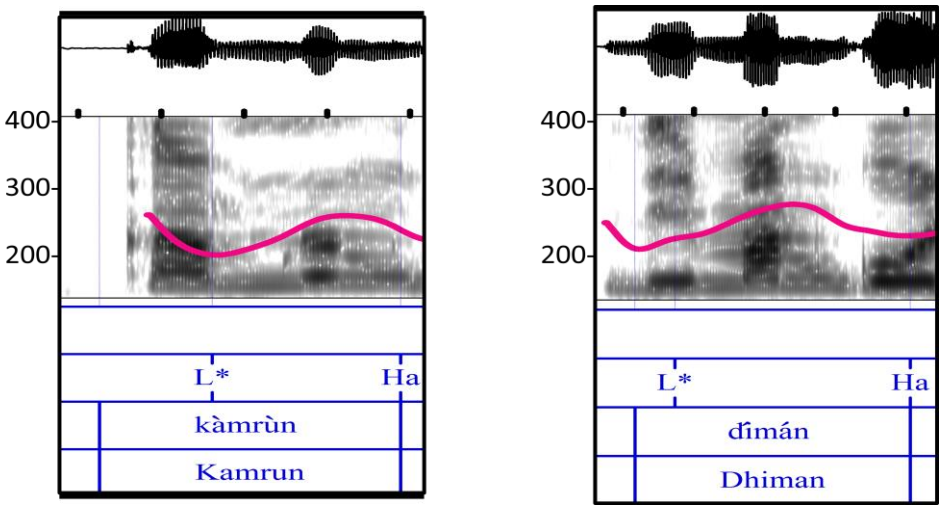

Figure 2.11: pitch track of IP initial APs with L underlying tone (left panel) and H underlying tone (right panel) from near-identical sentences produced by a female speaker

The Figure above (**Figure 2.11**) shows the pitch contour of two different initial APs. Despite the differences in the underlying tonal specification of the words, the surface pitch of the two are scaled with similar f0 values. The average f0 of the AP containing an underlying H tone word is 254 Hz (right panel), while the AP with an underlying L toned

[2] All syllables in lexical roots carry the same tone regardless of the position of aspiration loss, making the roots Tone Bearing Units (TBU). This gives us words with LL or HH tonal specifications. Although, most nominal suffixes in Sylheti are underlyingly toneless, they obtain the opposite tone of the root due to tonal polarity, leading to formation of words with LH, HL, LLH, HHL tonal structure (Gope, 2016).

word has an average f0 of 252 (left panel).[3] We interpret these results as evidence that the intonational tone of the AP at left IP boundaries completely overrides the lexical tones.

In the sentence medial positions, we find an interesting interplay of lexical and intonational tones. While the intonational repeated rising contours are still seen to stretch across APs, the lexical tones of the component words manifest by means of differences in f0 scaling. Specifically, the pitch contour on APs composed of words with underlying H tones is a more upstepped rise compared to that of an underlying L tone word (**Figure 2.12** and **Figure 2.13**). This reveals a co-existence of lexical tones and intonational pitch accents, distinguishing Sylheti from most South Asian languages. Furthermore, as we argue below, the manifestation of lexical tone as a difference in pitch accent scaling distinguishes Sylheti from languages like Swedish where lexical tone surfaces through pitch accent type (Myrberg, 2010), or Tokyo Japanese, where lexical tone surfaces through pitch accent location and presence/absence; (Beckman and Pierrehumbert, 1986; Pierrehumbert and Beckman, 1988). The retention of underlying tonal specifications is demonstrated here to compare the declaratives differing only in terms of the lexical tone specification of the direct object. There is a clear distinction between the f0 scalings of the words /gó̜ri/ 'watch' vs. /bari/ 'house,' in the two sentences produced by a female speaker during **(Figure 2.12),** indicating that along with the superimposition of intonational tones on sentence medial words, the lexical contrast is also projected. In terms of f0 values, the average of the AP containing the word /gó̜ri/ is 242 Hz, while that of the AP containing /bari/ is 211 Hz. A similar occurrence can be seen in **Figure 2.13,** where /zal/ ‘net’ has a mean f0 of 196 Hz, and the AP constituted by /d̪án/ ‘rice paddy’ has a mean f0 of 231 Hz. The upward scaling in APs with H lexical roots, therefore, results in a shallow rising contour.

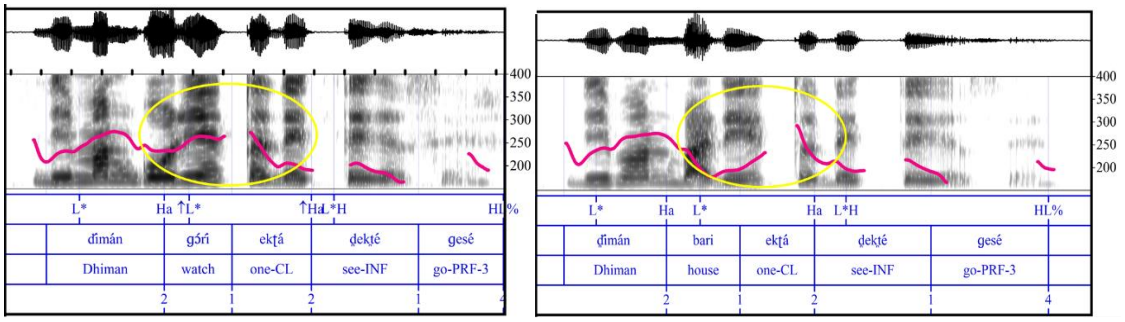

Figure 2.12: f0 contours of the sentences [ɖíman gɔ́ri ekʈá ɖekʈé gesé], 'Dhiman has gone to see a watch*'* (left panel), and, [ɖímán barí ekʈá ɖekʈé gesé], 'Dhiman has gone to see a house*'*, (right panel) produced by a female speaker.

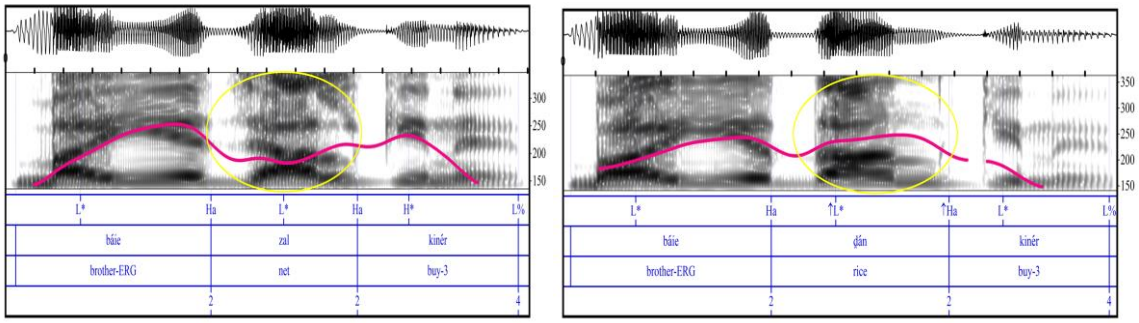

Figure 2.13: f0 contours of the sentences [báie zal kinér], '(my) brother is buying a net' (left panel), and, [báie ɖán kinér], '(my) brother is buying' paddy, (right panel) produced by a female speaker

Although this striking difference of f0 scaling is clear in surface contours of lexical roots with underlying contrastive tones, the H tone suffixes, however, do not induce a higher scaling.

2.3.5. intermediate phrase (ip)and Intonational phrase (IP)

The marker for **intermediate phrases (ip**) is a perceived juncture that is longer than the ones present between APs. In our corpus, ips are found in list utterances as well as in alternative questions with two verb phrases. In **Figure 2.14**, each non-final element in the list utterance constitutes individual ips, identified through a perceived break.

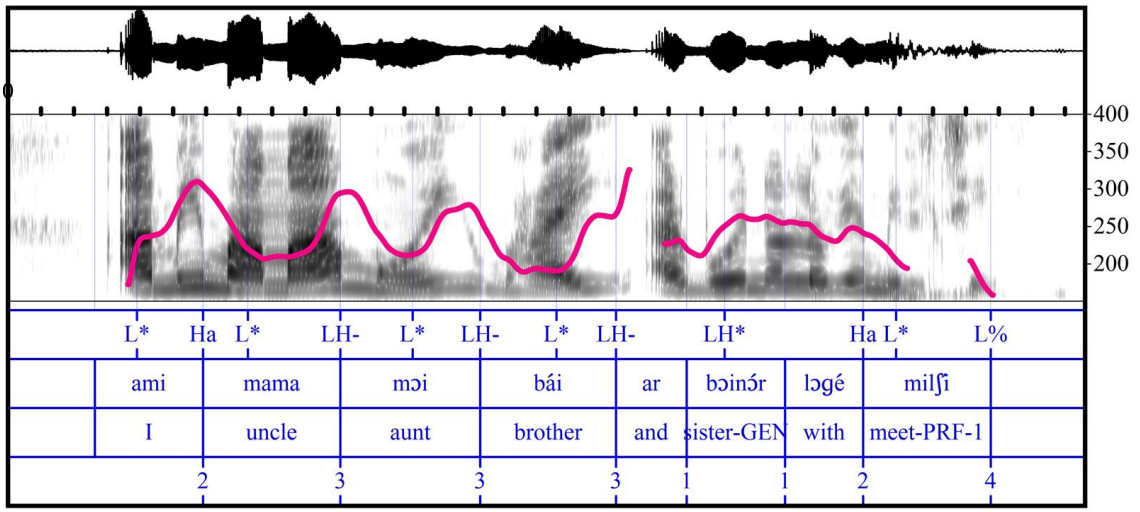

Figure 2.14: f0 contour of the sentence [ami mama, məi, bái, ar bɔinɔ́r ləgé milʃĩ], 'I met my uncle, aunt, brother, and sister', produced by a female speaker

Figure 2.15 shows an ip containing a verb phrase- the first alternative element of an alternative question.

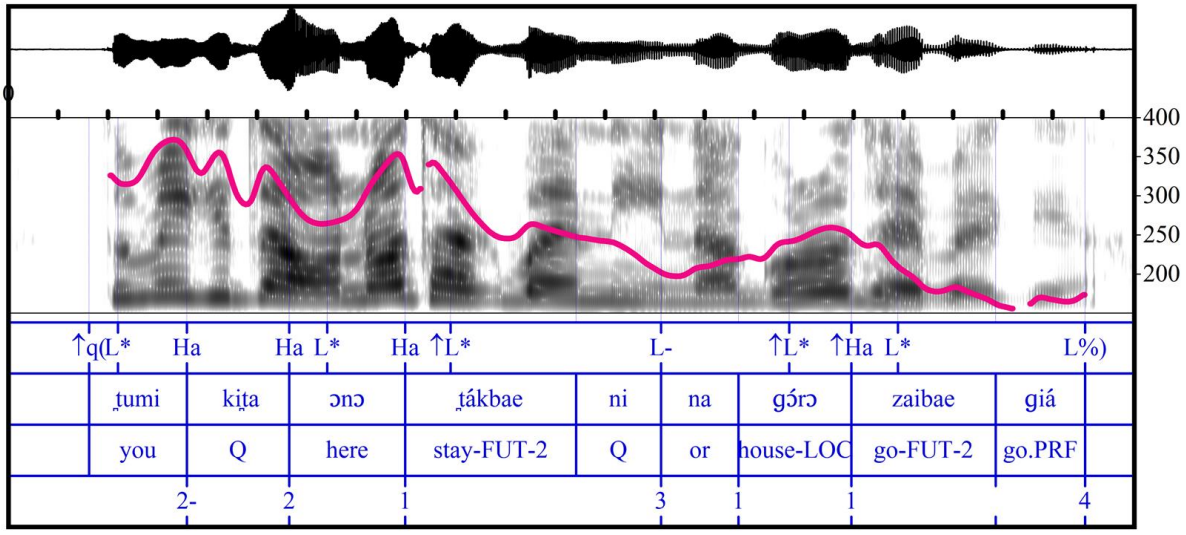

Figure 2.15: f0 contour of the sentence [t̪umi kiṭa ɔnɔ takbáe ní na gɔ́rɔ zaibái gía], 'Will you stay there or go back home? ' produced by a female speaker

The highest prosodic unit in Syheti is the **Intonational Phrase (IP),** which often corresponds to clauses. They are marked with distinct right boundary tones. The pitch contour of utterances falls to an f0 minimum (L%) or an f0 maximum (H%) at the end of each IP. Additional intonational cues present at IP boundaries include pauses following the right boundaries and pitch resets at left boundaries. Unlike the simple sentences where all medial tones undergo phonetic downstepping of AP-level H tones in

compound and complex constructions, a non-initial clause exhibits a resetting of pitch to restart a downtrend pattern. Three IP boundary tones are present in the language-

(I) L% found at the right edges of IPs consisting of complete sentences and subordinate clauses of complex sentences.

(II) H% present at right edges of subordinate clauses in compound sentences.

(III) HL% wherein the f0 undergoes rising before the realization of the final f0 minimum, they are operational mostly at the right edges of some sentence final IPs.

The IP boundary tones mostly override the underlying lexical tones of the words, as their tonal targets and scaling remain unaffected by the underlying tonal specification of the syllables.

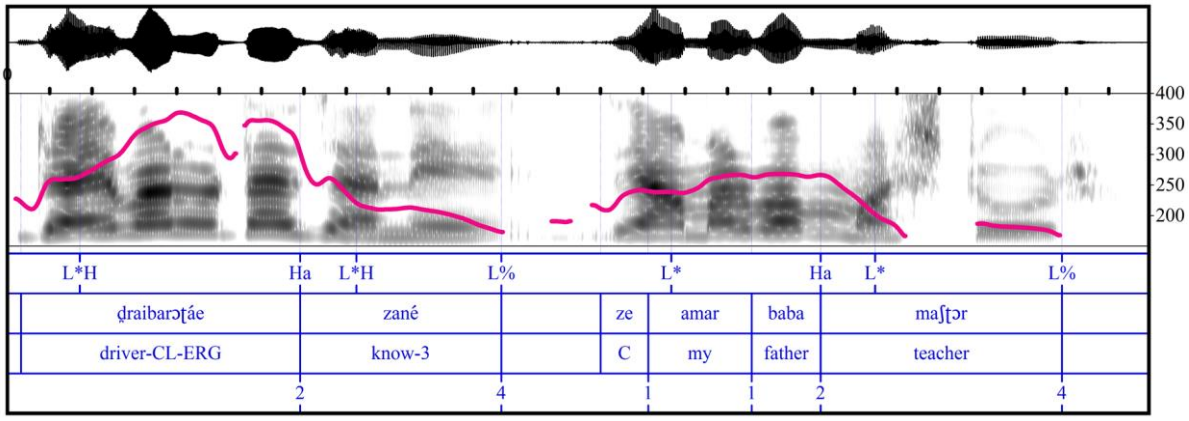

Figure 2.16: f0 contour of the sentence [ɖraibar ɔ́ʈáe zané ze amar baba maʃʈɔr] 'the driver knows that my father is a teacher', produced by a female speaker

In **Figure 2.16,** The main clause [ɖraibar ɔ́ʈáe zané] has an L% boundary tone at its right edge. The longer pause between the two clauses is indicated using break index 4 at the bottom tier. The beginning of the dependent clause undergoes pitch reset, blocking the downstepping effects of the previous clause. In **Figure 2.17**, the initial subordinate clause ends with an H% right boundary tone followed by a pause.

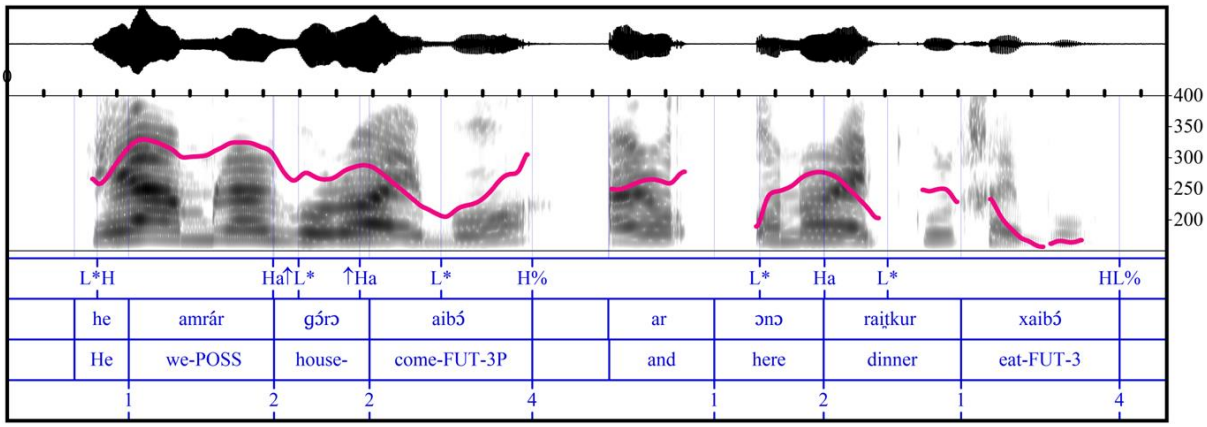

Figure 2.17: f0 contour of the sentence [he amrár gɔ́rɔ aibɔ́ ar ɔnɔ raiṭ̪kur xaibɔ́] 'he will come to my house and have dinner here' produced by a female speaker

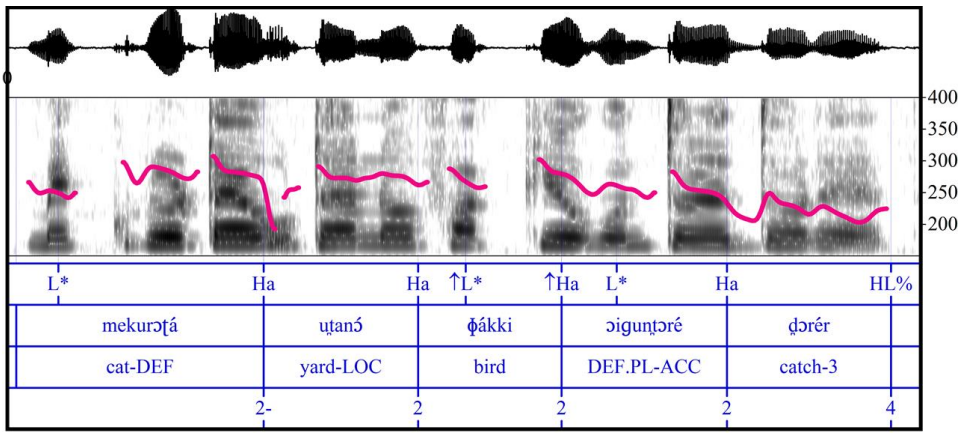

Figure 2.18: f0 contour of the sentence [mekurɔ ʈá uṭ̪anɔ́ ɸákkí ɔigunṭ̪ɔré ɖɔrér] 'the cat is chasing the birds in the courtyard', produced by a female speaker

Figure 2.18 shows a complete utterance in Sylheti, constituting one prosodic unit, i.e., the IP, ending with the falling contour L% at the right boundary.

2.4. Intonational Prosodic Units in Chokri: Results and Analysis

2.4.1. Prosodic Word (PrWd)

The use of structural prosody for the division of longer utterances into smaller units is limited in Chokri. Studies on lexical tone in the language report that contrastive tones in the language are docked to the final syllable of the lexical roots, while the non-final syllables are specified with a default mid tone. The affixes, on the other hand, have their own tonal specifications, which are not subject to any phonological change upon affixation (Tetseo and Gope, 2021) (see **Figure 2.19** and **2.20**). Both roots and affixes, therefore, constitute separate Prosodic Words (PrWd). The criterion for PrWd in the

language is as follows: *the unit must have only one syllable that represents and maintains tonal distinctions*. This also entails that PrWd in the language does not correspond to syntactic or grammatical words. No tonal process applies beyond the domain of PrWd.

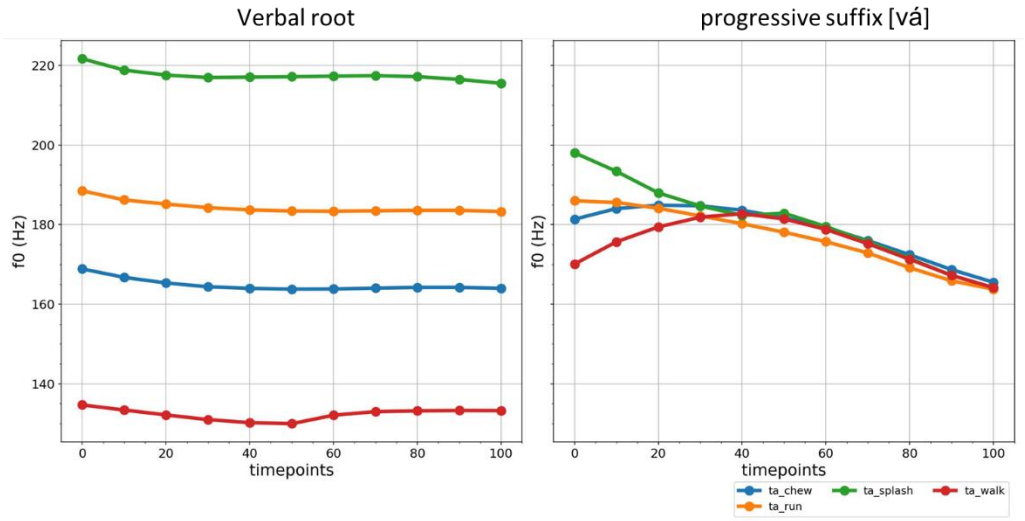

Figure 2.19: time-normalized average f0 of the minimal quadruplet-[tā̋] 'splash' (in green), [tá] 'run' (in orange), [tā] 'chew'(in blue), [tà] 'walk' (in red) affixed with the verbal progressive suffix [vá] (right panel). The progressive aspect marking suffix maintains its underlying H tone regardless of the tone of the root it attaches to.

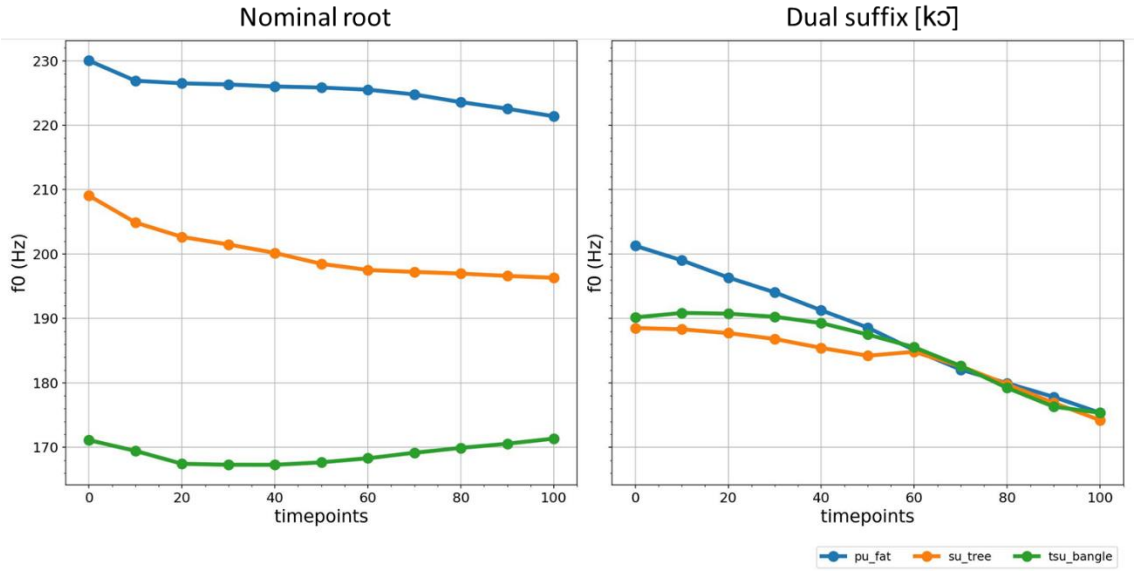

Figure 2.20: time-normalized average f0 of three nominal roots with three contrastive tones - [tsə̀] *'bangle'* (in green) [pə̋] 'fat' (in blue) and [sə́] '*tree*' (in orange) with the dual suffix [kɔ̄] (right panel) maintains its underlying tone regardless of the tone of the root it attaches to.

2.4.2. intermediate phrase(ip) and intonational phrase (IP)

Through our analysis of longer utterances in Chokri, we identified two layers of the prosodic hierarchy that are marked with intonational properties- the Intonational Phrase (IP) and the intermediate phrase (ip). Phrasal tones are not present in the language to indicate prosodic boundaries. As seen in the majority of the tonal languages, Chokri also does not have pitch accents as markers of prosodic units. The use of phrasal tones at unit boundaries is also absent. The underlying lexical tonal specifications of all the syllables are retained on the surface pitch, regardless of their positions. Instead, the language makes use of features other than pitch to encode information about prosodic constituency.

The Intonational Phrase (IP) primarily corresponds to complete sentences or main clauses. There are no discernible boundary tones on either edge of the IPs. The primary feature employed for distinguishing a speech unit as an IP is **final lengthening**. In all the tokens considered for this study, the final syllable of the IPs is produced with a longer duration than that of non-final syllables. Moreover, the IPs are also the domains of downtrends in Chokri, where syllables occurring at a later part of the utterance have lower f0 values compared to previously occurring syllables specified for the same tone.

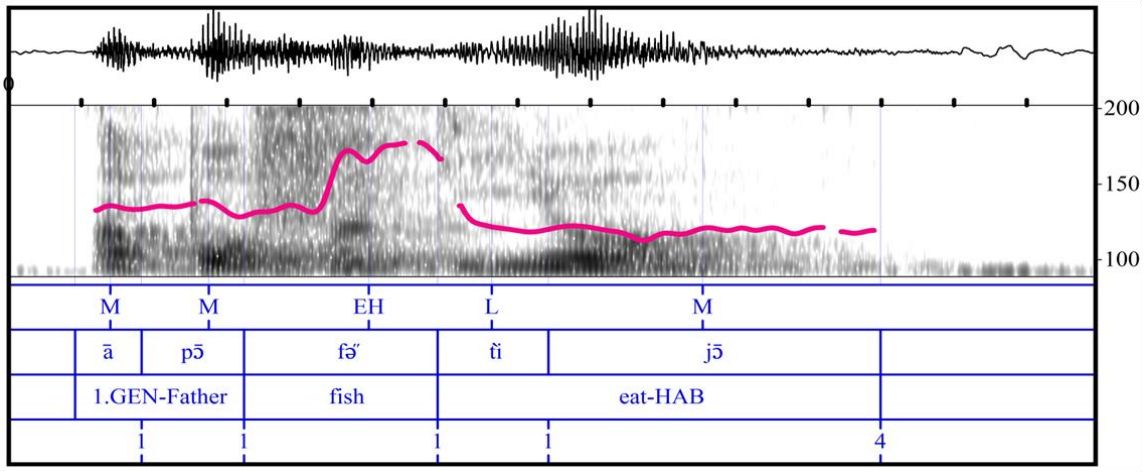

Figure 2.21: f0 contour of the declarative sentence [ā-pɔ̄ fə̋ tì jɔ̄] ‘my father eats fish’ produced by a male speaker

Figure 2.21 shows an utterance that constitutes an IP in Chokri. The duration of the final syllable is evidently longer than that of the non-final syllable; it has a mean duration of 456 ms, whereas the previously occurring syllables have a mean duration of 162 ms. It

also shows the downtrend effect on f0 levels in the pitch contour. The initial M tone has an f0 of 127 Hz; in the second syllable with the same tone, f0 is 125 Hz, and the final M that occurs later in the utterance, the f0 drops down to 115 Hz. The initial drop in pitch is due to the effect of declination, which causes a steeper fall at the very beginning of all utterances, yet it does not move beyond its tonal space (see 3.3.3). The pitch does not fall to an f0 minimum, nor does it rise at the right boundary; the last tone is instead maintained till the very end of the utterance. This confirms the absence of an IP boundary tone. **Figure 2.22** shows the time normalized average f0 values of all tokens of the sentence, representing both declination and stability of lexical tones.

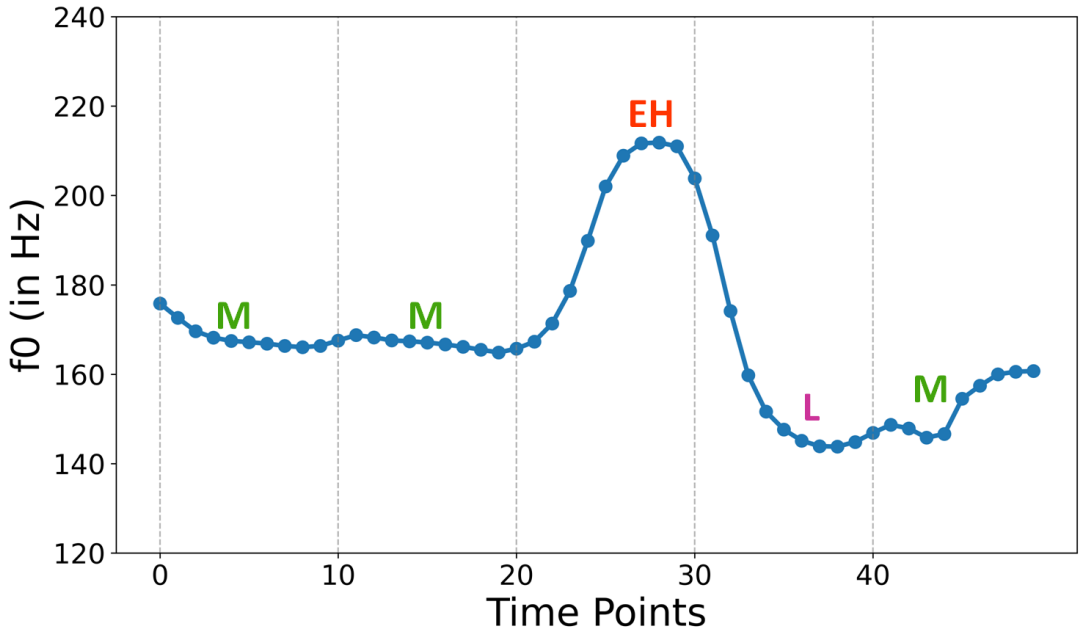

Figure 2.22: time normalized average f0 of the sentence [ā-pɔ̄ fə̌ tì jɔ̄] 'my father eats fish' of all tokens by the five speakers. The lexical tone of each syllable is labeled with tonal notations. Each 10 points in the x-axis corresponds to one syllable.

Similarly, in **Figure 2.23**, the final syllable [vɛ́] has a considerably longer duration than the others. The retention of lexical tone specification is evident in the changes at the pitch level for every syllable. The first two syllables with an M tone have an average f0 of 125 Hz, while in the H toned syllable, the average f0 is 130.2. The EH tone syllables have an average of 151 Hz. The discrete tone levels maintained throughout each syllable are further evident in **Figure 2.24,** which draws syllable-wise time-normalized average f0 of all the tokens of this sentence. In the **Figure 2.25** too, the underlying tone can be identified from the distinct pitch heights in each syllable with different lexical tones.

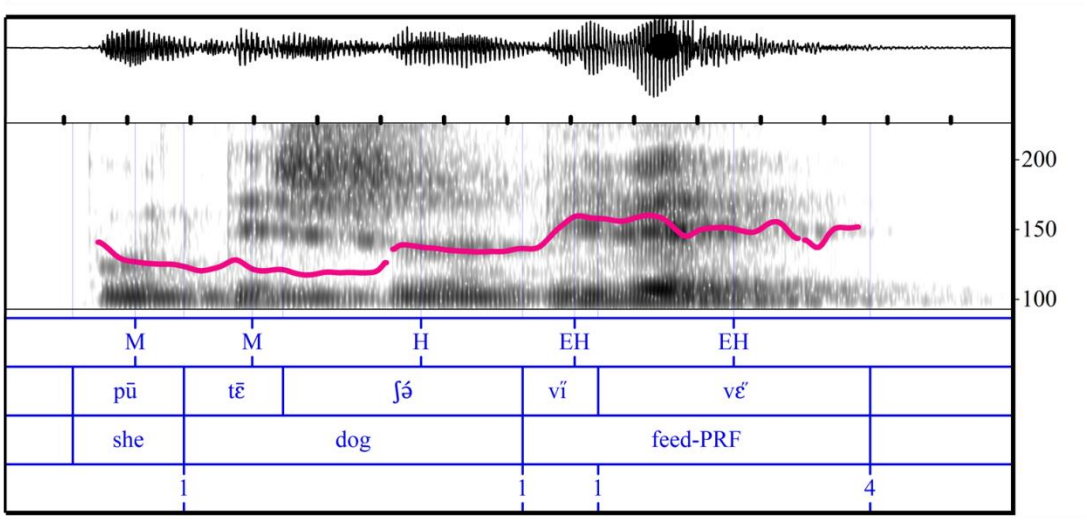

Figure 2.23: f0 contour of the declarative sentence [pū tēʃə́ vǐ vɛ̋] 'she fed the dog' produced by a male speaker.

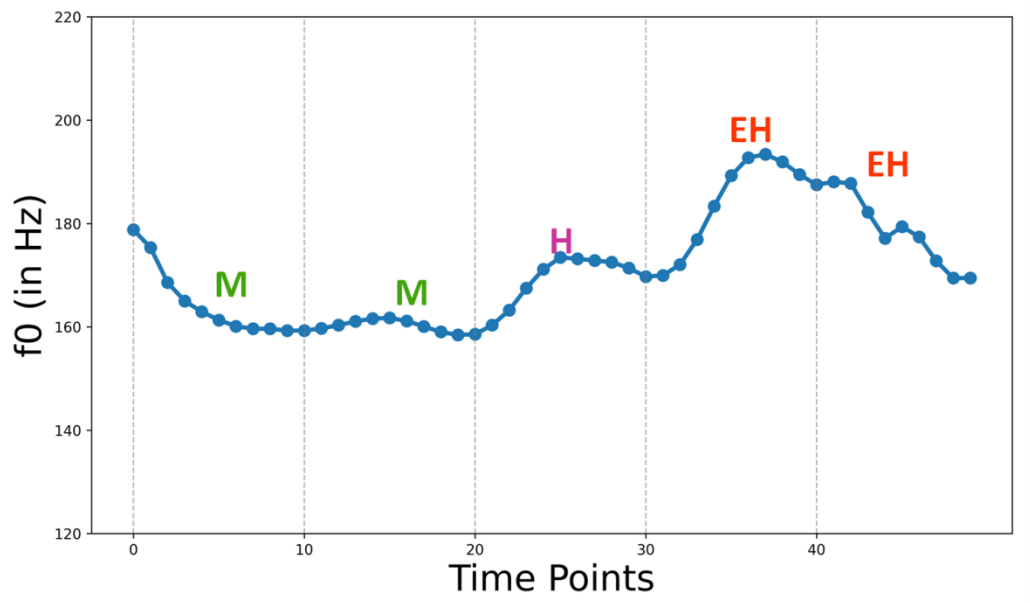

Figure 2.24: f0 contour of the declarative sentence [pū tēʃə́ vǐ vɛ̋] 'she fed the dog' of all tokens by the five speakers.

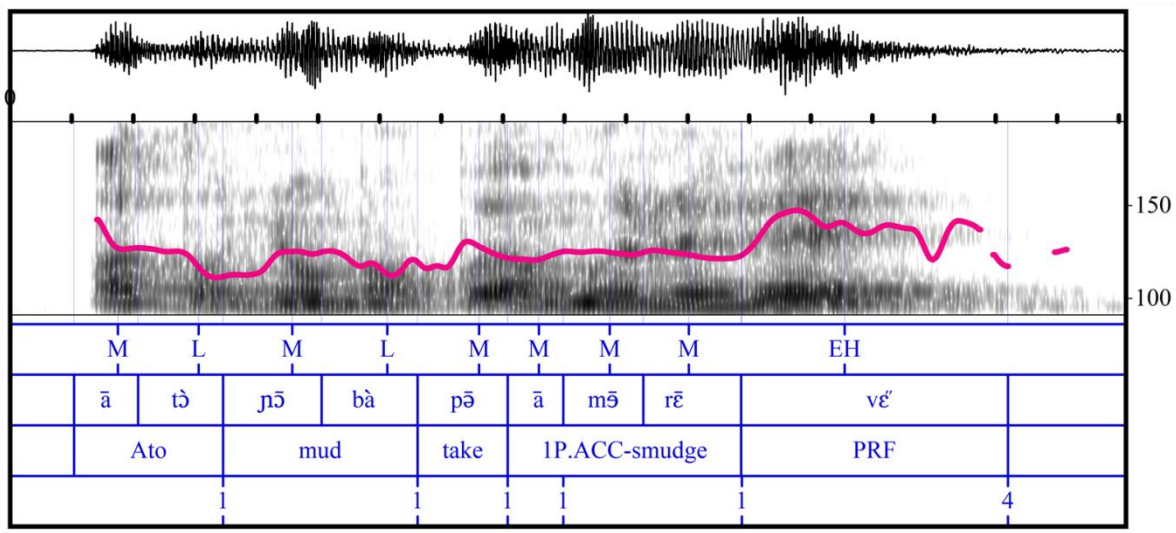

Figure 2.25: f0 contour of the declarative sentence [ātɔ̀ ɲɔ̄bà pə̄ ā-mə̄rɛ̄ vɛ́] 'Ato smudged me with mud' produced by a male speaker.

To confirm that the difference in duration observed between final and non-final syllables of IP is statistically significant, a repeated measures one-way ANOVA was conducted in Python. The duration data for both types of syllables for all the speakers was arranged in a long format. After loading the data using the `pandas` library, the 'speaker' and 'Syllable_Position' columns were converted to categorical variables. We used the AnovaRM function from the statsmodels library, with 'f0' as the dependent variable, 'speaker' as the subject identifier, and 'Syllable_Position' as the within-subject factor. The results of the repeated measures ANOVA shows that there is a statistically significant difference in 'duration' between 'non-final' and 'final' syllables of IPs of around 200 ms ($F(1.0, 5.0) = 44.95$, $p < 0.05$), (see **Figure 2.26**).

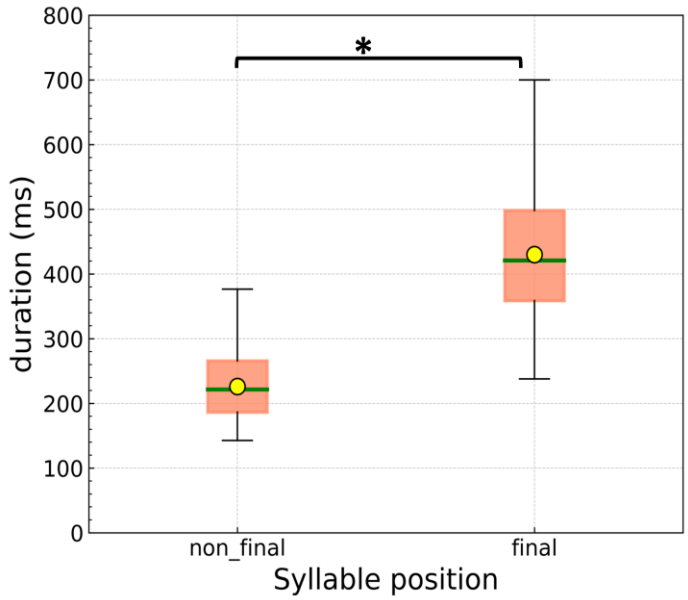

Figure 2.26: Comparison of average duration of IP final syllables and non-final syllables in Chokri. The asterisk * marks the significance of the difference between the two.

Another intonationally marked prosodic unit in Chokri is the intermediate phrase (ip). The ips correspond to embedded clauses in the language. Similar to IPs, ips also do not involve any type of phrasal tones. They are indicated only through an **optional pause** that follows its right boundary. Unlike IPs, ip right boundaries do not exhibit lengthening. In our production experiment, pauses at the boundaries of constituent clauses were consistently present in data from two speakers (one male and one female). In the other three speakers, some iterations contained the pause, and others did not. This suggests that the embedded clauses are not always treated as individual prosodic units (ip) by the speakers. Nevertheless, a prosodic feature optionally employed at this specific location by some of the speakers, pause can be identified as a marker of ip right boundaries. This optional pause is shown in the **Figures 2.28** and **2.30**. The syllable occurring at the ip boundaries does not undergo lengthening. Other than that, no boundary tone and pitch reset are observed.

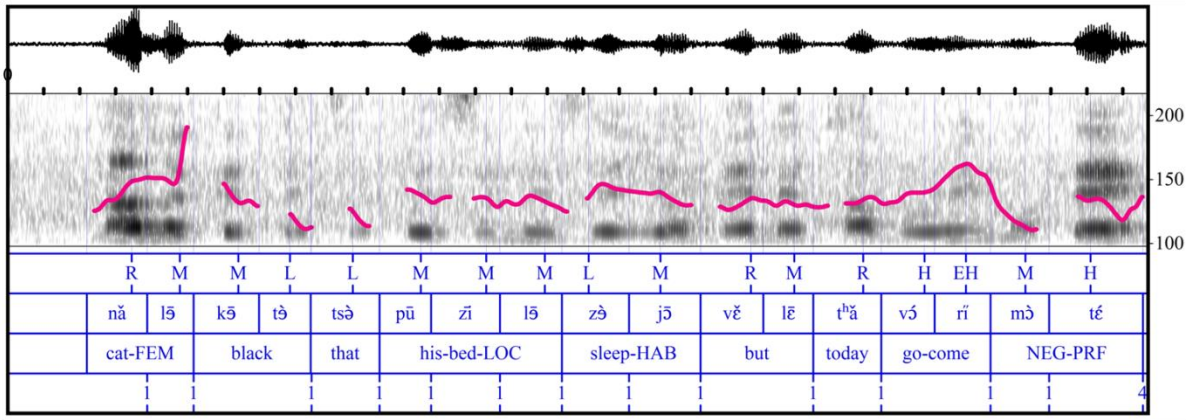

Figure 2.27: Pitch track of the sentence [nălɔ̄ kɔ̄tə̀ tsə̀ pū-zī-lɔ̄ zə̀-jɔ̄ vɛ̌lē tʰă vɔ́rǐ mɔ̀-tɛ] 'The black cat sleeps in his bed but it did not come today' produced by a male speaker without pause between the two clauses.

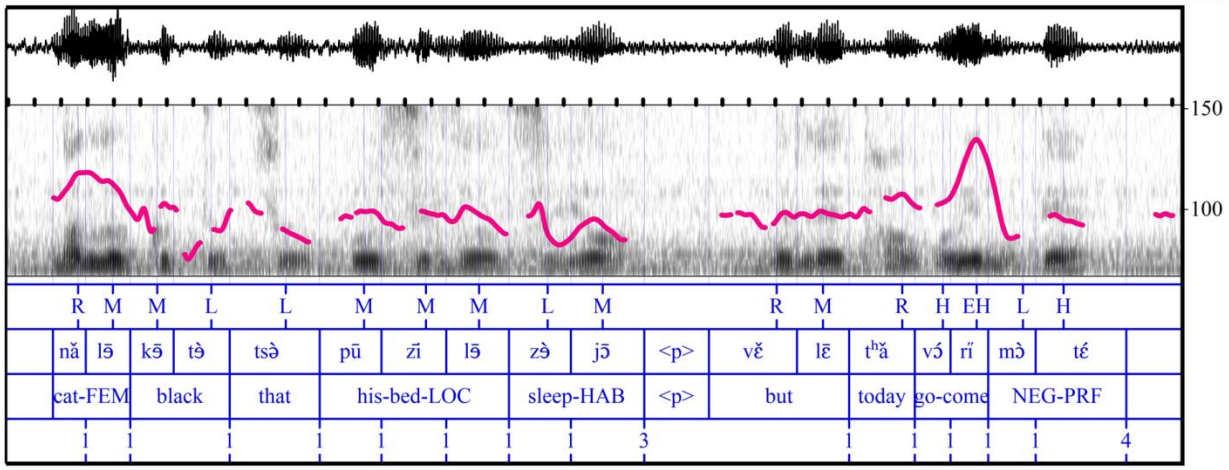

Figure 2.28: Pitch track of the sentence [nălɔ̄ kɔ̄tə̀ tsə̀ pū-zī-lɔ̄ zə̀-jɔ̄ vɛ̌lē tʰă vɔ́rǐ mɔ̀-tɛ] 'The black cat sleeps in his bed, but it did not come today' produced by a male speaker.

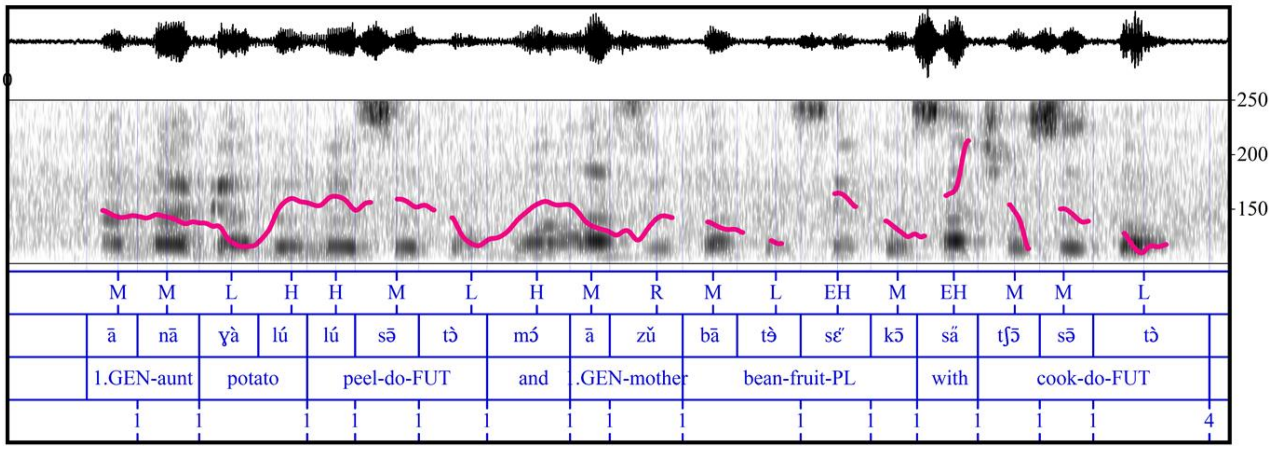

Figure 2.29: Pitch track of the sentence [ā-nā gàlú lɘ́sə̄-tɔ̀ mɔ́ ā-zǔ bātə̀seɛ̋ sā̋ ʧɔ̄sɔ́-tɔ̀] 'Aunt will cut the potatoes and mother will fry them with beans' produced by a male speaker.

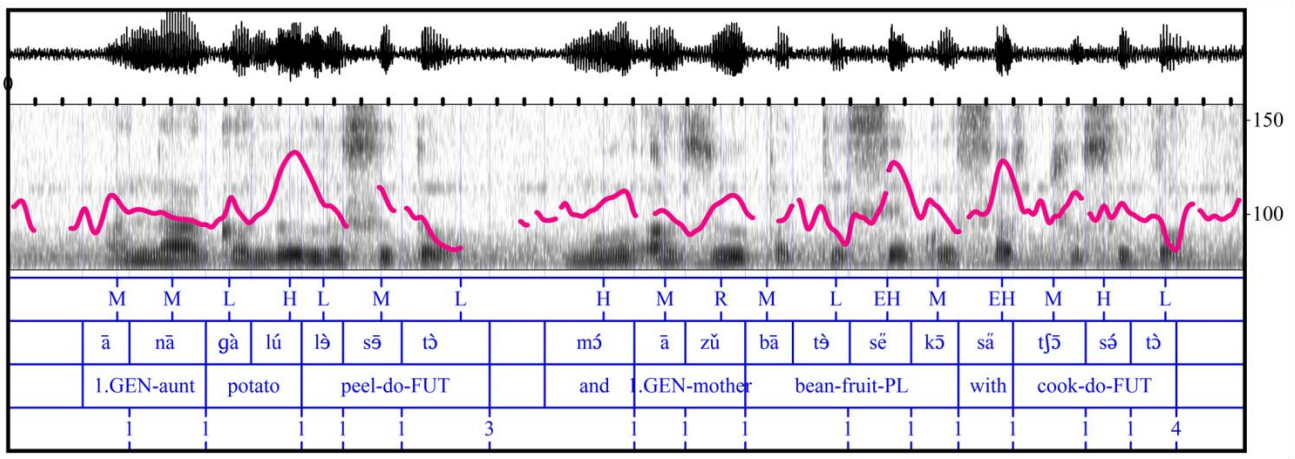

Figure 2.30: Pitch track of the sentence [ā-nā gàlú lə́sə̄-tɔ̀ mɔ́ ā-zŭ bātə̀seɛ̋ sā̋ ʧṓsə́-tɔ̀] 'Aunt will cut the potatoes and mother will fry them with beans.' produced by a male speaker.

2.5. Discussion and Conclusion

Our study on the prosodic structures in Sylheti and Chokri reveals how intonational functions of prosodic domain marking in languages from opposing points of the tonal complexity spectrum. The experimental evidence confirms that Sylheti intonational phonology, in many ways, exhibits affinity to those of its related non-tonal languages, including Kolkata Standard Bengali (Hayes and Lahiri, 1991) and Bangladeshi Standard Bengali (Khan, 2014). This is evident in intonationally marked prosodic units that are characterized by post-lexical tonal events, i.e., pitch accents and boundary tones. The APs, in particular, are reflective of the intonational pattern of 'repeated rising' shared by many South Asian languages (Fery, 2010).

The inventory of intonemes in Sylheti is shown below:

Pitch accents:

L*	default pitch accent (low alternative)
L*H	default pitch accent (rising alternative)
H*	question word pitch accent

Lexical scaling:

↑ H	raising of AP's H due to the presence of IP-medial, non-word-final lexical

AP boundary tones:

Ha	all non-IP-final APs

IP boundary tones:

H%	first half of incomplete sentences
HL%, L%	all other sentence types

The presence of pitch accents in a language where syllables are specified for lexical tone is a phenomenon that is scarce across available literature on intonation in tone language. The most striking revelation of this study is that both lexical and post-lexical tones participate in making up the surface pitch of utterances. Post-lexical tones play a primary role in the language and are not overridden by the lexical tones; lexical tones, however, manifest themselves as adjustment of intonational tones in certain locations. Although intonational tones are superimposed, at least in the sentence medial positions, lexical tones are still recoverable. Despite sharing many intonational features with its related languages, the difference in the scaling of phrasal tones induced by lexical tone immediately distinguishes Sylheti from them. The observed interaction between tone and intonation in Sylheti is reminiscent of Clement's (1979) 'Tone Levels Model': "… a tone level frame does not constitute a set of absolute acoustic parameters. Rather, it is subject to modification as a result of the intonational processes that apply to it. The identity of the frame itself, however, is not affected by these modifications." A similar instance of the presence of phrasal tones in domains smaller than clauses has been reported in Burmese, where the consecutive short phrases are marked by an initial peak on the first major/non-reduced syllable (Ozerov, 2017).

Similarly, in the projection of the phrase final H in Northern Khammu, f0 is always higher for the word with a lexical High tone than for that with a Low tone (Karlsson *et al.,* 2012). In a language like Sylheti, the representation of lexical tones cannot be separated from the intonational pitch specifications, at least not when they are part of longer utterances, instead of occurring in isolation. Changes in pitch range or register in which lexical tones are realized for intonational purposes have been well attested with data from several languages. However, as we showed based on our data, Sylheti demonstrates a sort of inverse of the same: instead of intonation serving as a scaling factor on lexical tone, Sylheti lexical tones serve as a scaling factor on the intonational contour. An AM account for the same targets hosting lexical and post-lexical pitch information would involve separate descriptions and rules for 'sequential tonal specification', i.e., the intonational tones in this case and 'pitch register distinctions' or scaling differences for indicating lexical tones. To accommodate tonal south Asian

languages like Sylheti in a unified ToBI system like the InTraSAL (Khan, 2019), crucial information like retention of lexical tone distinction in an intonational contour has to be transcribed. We propose adding a ↑ symbol before the pitch accent and boundary tones that are realized with a raised f0 scaling due to the effect of the underlying H tone).

On the other hand, the results of our analysis indicate a restricted use of intonational features for distinguishing prosodic units in Chokri. The absence of phrasal tones (both pitch accents and boundary tones) and pitch reset suggests that Chokri primarily uses non-pitch prosodic features like duration and pause for information about prosodic units. The underlying lexical tones dominate the intonational properties, possibly due to the complex and rich tonal properties in the language. However, the IP as the domain of downtrends indicates an interaction between lexical tones and intonational pitch effects used for marking domain boundary. The absence of boundary tones has also been reported in Mambila, a Bantoid language with four contrastive lexical tones (Connel, 2017). The lack of categorical intonation features in these languages with rich tonal inventories implies that the lexical tones are given a high functional load and are not obliterated through the superimposition of intonemes. The prosodic strategies- final lengthening and pause encode the distinction of speech units in Chokri. Another important finding of this study is the non-correspondence between grammatical words and PrWd in the language, as tones are lexically specified not only for roots but also for the prefixes and suffixes. Cross-linguistically, the Prosodic Stems (PStem) are reported in several languages where prefixes have prosodic features that are not related to or influenced by the Pstem, i.e., root+suffixes (Downing and Kadenge, 2020). In Chokri, however, both suffixes and prefixes are non-cohering affixes and can make up individual PrWd domains. This further highlights the stability and high functional load of lexical tones in the language that explains the limited use of intonational pitch categories.

References

Beckman, M. E., & Pierrehumbert, J. B. (1985). Synthesizing Japanese intonation using a downstep model. *The Journal of the Acoustical Society of America, 77*(S1), S38. https://doi.org/10.1121/1.2022308

Beckman, M., Hirschberg, J., & Shattuck-Hufnagel, S. (2004). The original ToBI system and the evolution of the ToBI framework. In *Prosody: Theory and experiments* (pp. 9–28). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199249633.003.0002

Bennett, R., & Elfner, E. (2019). The syntax–prosody interface. *Annual Review of Linguistics, 5*, 151–171. https://doi.org/10.1146/annurev-linguistics-011718-012503

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* [Computer program] (Version 6.4.07). Retrieved March 17, 2024, from http://www.praat.org/

Cahill, M. (2017). Kɔnni Intonation. In L. Downing & A. Rialland (Ed.), *Intonation in African Tone Languages* (pp. 53-88). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110503524-003

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1–52). MIT Press.

Clements, G. N. (1979). The description of terraced-level tone languages. *Language, 55*(3), 536–558. https://doi.org/10.2307/412167

Connell, B. (2017). Tone and intonation in Mambila. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 231–266). De Gruyter Mouton. https://doi.org/10.1515/9783110503524-013

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Harvard University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge University Press.

Dobashi, Y. (2003). *Phonological phrasing and syntactic derivation* (Doctoral dissertation). Cornell University.

Downing, L., & Kadenge, M. (2020). Re-placing PStem in the prosodic hierarchy. *The Linguistic Review, 37*(3), 433–461. https://doi.org/10.1515/tlr-2020-2050

Downing, L., & Rialland, A. (Eds.). (2017). *Intonation in African tone languages*. De Gruyter Mouton. https://doi.org/10.1515/9783110503524

Elordieta, G. (2008). An overview of theories of the syntax–phonology interface. *International Journal of Basque Linguistics and Philology, 42*, 209–286.

Elvira García, W. (2022). *Create pictures with tiers v.6* [Praat script].

Fery, C. (2010). The intonation of Indian languages: An areal phenomenon. In *Festschrift for Rama Agnihotri* (pp. 288–312).

Frota, S. (2012). Prosodic structure, constituents, and their implementation. In A. Cohn, C. Fougeron, & M. Huffman (Eds.), *The Oxford handbook of laboratory phonology* (pp. 255–265). Oxford University Press.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology, 15*, 411–458. https://doi.org/10.1016/0010-0285(83)90011-9

Gope, A. (2016). *The phonetics and phonology of Sylheti tonogenesis* (Doctoral dissertation). Indian Institute of Technology, Guwahati.

Mahanta, S., & Gope, A. (2018). Tonal polarity in Sylheti in the context of noun faithfulness. *Language Sciences, 69*, 80–97. https://doi.org/10.1016/j.langsci.2018.06.010

Gussenhoven, C. (2015). *Suprasegmentals*. Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.52024-8

Hansson, P. A. (2003). *Prosodic phrasing in spontaneous Swedish* (Doctoral dissertation). Lund University.

Hartmann, K., & Zimmermann, M. (2007). In place — out of place? Focus strategies in Hausa. In *Linguistik aktuell* (pp. 365–403). https://doi.org/10.1075/la.100.20har

Hayes, B., & Lahiri, A. (1991). Bengali intonational phonology. *Natural Language & Linguistic Theory, 9*, 47–96. https://doi.org/10.1007/BF00133326

Inkelas, S. (1990). *Prosodic constituency in the lexicon*. Garland Publishing.

Jun, S. A. (Ed.). (2005). *Prosodic typology: The phonology of intonation and phrasing*. Oxford University Press.

Kaisse, E. M. (1985). *Connected speech: The interaction of syntax and phonology*. Academic Press.

Karlsson, A., House, D., Svantesson, J.-O., & Tayanin, D. (2010). Influence of lexical tones on intonation in Kammu. In *Proceedings of Interspeech 2010* (pp. 1740–1743).

Keane, E. (2014). The intonational phonology of Tamil. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 375–399). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199567300.003.0008

Khan, S. D. (2014). The intonational phonology of Bangladeshi Standard Bengali. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 49–78). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199567300.003.0004

Khan, S. D. (2019). InTraSAL: An intonational model for South Asian languages. In *ICPhS Satellite Meeting on the Intonational Phonology of Typologically Rare or Understudied Languages* (pp. 1–20). University of Melbourne.

Khan, S. D. (to appear). The intonational Phonology of South Asian Languages. In Jun S.A, & Khan, SD (eds). Prosodic Typology III. Oxford University Press.

Kügler, F. (2015). Phonological phrasing and ATR vowel harmony in Akan. *Phonology, 32*(1), 177–204. http://www.jstor.org/stable/43865846

Kula, N. C., & Bickmore, L. S. (2015). Phrasal phonology in Copperbelt Bemba. *Phonology, 32*(1), 147–176. http://www.jstor.org/stable/43865845

Ladd, R. D. (2008). *Intonational phonology*. Cambridge University Press.

Makasso, E., Hamlaoui, F., & Lee, S. (2017). Aspects of the intonational phonology of Bàsàá. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 167–194). De Gruyter Mouton. https://doi.org/10.1515/9783110503524-010

Gope, A. (2016). *The phonetics and phonology of Sylheti tonogenesis* (Doctoral dissertation). Indian Institute of Technology, Guwahati.

Mahanta, S., & Gope, A. (2018). Tonal polarity in Sylheti in the context of noun faithfulness. *Language Sciences, 69,* 80–97. https://doi.org/10.1016/j.langsci.2018.06.010

Myrberg, S. (2010). *The intonational phonology of Swedish* (Doctoral dissertation). Stockholm University.

Nespor, M., & Vogel, I. (2007). *Prosodic phonology: With a new foreword*. De Gruyter Mouton. https://doi.org/10.1515/9783110977790

Ozerov, P. (2017). *Tone, intonation, and phrasing in Burmese*. Retrieved from https://www.academia.edu/16774194/Tone_intonation_and_phrasing_in_Burmese

Patin, C. (2017). Tone and intonation in Shingazida. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 99–120). De Gruyter Mouton. https://doi.org/10.1515/9783110503524-007

Pierrehumbert, J., & Beckman, M. E. (1988). *Japanese tone structure*. Linguistic Inquiry Monographs, 15.

Raychoudhury, P. (2023). Prosodic Aspects of Sylheti. PhD Dissertation. IIT Guwahati.

Seidl, A. (2001). *Minimal indirect reference: A theory of the syntax–phonology interface*. Routledge.

Selkirk, E. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim (Ed.), *Nordic prosody* (2nd ed., pp. 268–271). TAPIR.

Selkirk, E. (1988). Constraints on minor phrase formation in Japanese. In *Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society* (pp. 316–336). Chicago Linguistic Society.

Selkirk, E. (2011). The syntax-phonology interface. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (pp. 435–484). Blackwell Publishing.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). ToBI: A standard scheme for labeling prosody. In *Proceedings of the 2nd International Conference on Spoken Language Processing* (pp. 867–879). Banff.

Tetseo, S., & Gope, A. (2021). Tonal contrast and grammatical tone in Chokri. [Paper presentation]. 43rd International Conference of Linguistic Society of India, CIIL-Mysore.

Truckenbrodt, H. (1995). *Phonological phrases: Their relation to syntax, focus, and prominence* (Doctoral dissertation). MIT.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry, 30*(2), 219–256. https://doi.org/10.1162/ling.1999.30.2.219

Wagner, M. (2005). *Prosody and recursion* (Doctoral dissertation). MIT.

Xu, Y. (2013). ProsodyPro — A tool for large-scale systematic prosody analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (pp. 1–5).