JOURNAL OF SOUTH ASIAN LINGUISTICS

Volume 10: Special issue on prosody August 2020

Journal of South Asian Linguistics

CSLI PUBLICATIONS

Contents

Editorial statement

I am thrilled to present to you volume 10 of the *Journal of South Asian Linguistics*, which is in many ways very special. Firstly, it is our first volume dedicated entirely to prosody, and as such is also the first themed volume for our journal. Second, all four papers in this volume are invited contributions, solicited directly from researchers who were in turn invited to present their work at the one-day workshop on the prosody of South Asian languages, held on 18 June 2018, in coordination with the 34th South Asian Languages Analysis Roundtable (SALA-34), at the University of Konstanz. Lastly, this is the first complete volume to come out during the tenure of our current editorial board.

The first paper, by Saba Urooj, Benazir Mumtaz, and Sarmad Hussain (Al-Khawarizmi Institute of Computer Science), presents the first large-scale model of Urdu intonation, built under the framework of Autosegmental-Metrical Theory using a diverse set of utterances. The authors propose the Accentual Phrase (AP) as the principal building blocks of Urdu intonation, in line with what we are learning about many languages of South Asia and beyond. The paper's special focus on case markers, postpositions, compounds, and coordinative constructions also reveals details that can shed light on similar structures often left unexplored in other languages of the region.

The second paper, by Hans Henrich Hock (University of Illinois at Urbana-Champaign), takes on a very different lens, demonstrating how many seemingly unrelated phenomena in Hindi, Kashmiri, and Vedic Sanskrit can be characterized as stemming from the intersection of two domains of prosodic weakening: verbhood and utterance-finality. Here, rather than modeling the prosody itself, the author deftly applies our understanding of prosodic structure and phenomena towards a unified explanation for syntactic phenomena such as negation inversion, copula deletion, and V2 order.

The third paper, by Frank Kügler (Goethe University Frankfurt), uses both production and perceptual data to demonstrate that in Hindi, focus is not marked through phonetic or phonological adjustments to the focused element itself, rather it is marked primarily and consistently through the compression of post-focal material, in effect the lowering of the ceiling of the speaker's pitch register. He argues that because of its central role in languages like Hindi, pitch register must be proposed as a third phonological dimension for focus realization, alongside prosodic enhancement and changes to prosodic phrasing.

The fourth and final paper, by Caroline Féry (University of Frankfurt) and Gisbert Fanselow (University of Potsdam), brings us back to a comparative view, this time comparing the prosody of languages from the Indo-Aryan (Hindi), Dravidian (Tamil, Malayalam), and (notably) Tibeto-Burman (Bodo, Meithei) families. The authors focus on how discontinuous nominal phrases (NPs) are phrased and accentuated, depending on whether or not the discontinuity preserves c-commanding relations. They argue that due to their rampant scrambling and phrasal intonational properties, Indian languages behave unlike better-described intonation languages such as English or German in terms of the prosodic characteristics of their discontinuous NPs.

All four papers do an incredible job of highlighting the prosody of South Asian languages, making comparisons with better-described languages such as English, German, and Mandarin, to demonstrate both how South Asian prosody stands apart from other regions while still sharing many of the same building blocks of intonation. These papers also showcase the diversity within the region, calling for more attention to this long-overlooked area of prosodic typology.

Finally, I would like to thank my colleagues on the editorial team, Emily Manetta (University of Vermont) and Mythili Menon (Wichita State University), our senior advisers Miriam Butt (University of Konstanz) and Rajesh Bhatt (University of Massachusetts at Amherst), Sebastian Sulger (University of Konstanz) for setting up our website, and Sebastian Danisch (KIM at the University of Konstanz) for maintaining it. Special thanks also are due to the anonymous reviewers, and to Miriam Butt and Farhat Jabeen (University of Konstanz) for organizing the workshop that inspired this special volume.

I look forward to many more volumes, both regular and themed, of our now twelve-year-old journal, which has maintained its high standards while increasing accessibility through our diverse board, our wide range of contributors, and our open-access online platform. I especially invite those readers who are exploring the prosody of South Asian languages to consider *JSAL* as the venue for their ongoing and future work.

Sameer ud Dowla Khan, Reed College

Urdu Intonation

SABA UROOJ, *Al-Khawarizmi Institute of Computer Science*
BENAZIR MUMTAZ, *Al-Khawarizmi Institute of Computer Science*
SARMAD HUSSAIN, *Al-Khawarizmi Institute of Computer Science*

Received JULY 2018; Revised OCTOBER 2019

ABSTRACT

The current study is an analysis of an Urdu speech corpus using a Tone and Break Indices (ToBI) transcription system to develop a model of Urdu intonation. The analysis indicates that Urdu has three pitch accents (L*, L*+H, H*) and boundary tones associated to two phrase types: accentual phrase (AP) boundaries (Ha, La) and intonational phrase (IP) boundaries (L%, H%, LH%). The AP is a pitch bearing unit on a single word, or more than one word in the context of (a) *izāfat*, (b) conjunctive *vāo*, (c) case markers, (d) complex postpositions, and (e) complex verbs. Moreover, this study also investigates the tonal structure of declarative, interrogative (wh-questions, yes/no-questions), and imperative (semi-honorific, polite honorific) sentences in neutral focus context using 50 utterances produced by ten speakers. Results indicate that (i) all declarative sentences consist of a series of APs, represented as (aL) L* (H) Ha, except the sentence final AP, represented as (H*) L%. (ii) wh-questions are different from their corresponding declaratives in terms of pitch range and the final boundary tone; (iii) imperatives are different form their corresponding declaratives in terms of final boundary tone.

1 Introduction

This paper investigates the tonal inventory and tonal structure of declarative, interrogative, and imperative sentences of Urdu. Intonational studies make use of different types of speech data, including read speech, spontaneous speech, retelling of a story (Grabe, 1997), dialogue games (Krahmer & Swerts, 2001), and map tasks (Grice & Savino, 2003). The current study uses read speech as a baseline for three reasons. First, the work is being used for training a speech synthesis system, which uses speech articulated in the same mode (Kiruthiga & Krishnamoorthy, 2012). Second, read speech is more clearly articulated (Face, 2003), providing more consistent data for a baseline study. Finally, this method allows for designing the data to capture the necessary diversity of sentence structures.

In this paper the types of pitch accents, the tonal structure of a prosodic phrase and the interaction between various prosodic phrases for Urdu will be explored. The organization of the paper is as follows. Section 2 provides an overview of the intonation of selected South Asian languages (SALs). The methodological details concerning the collection, transcription and analysis of the speech corpus are given in Section 3. Sections 4 and 5 present the prosodic structure and the sentence specific prosody of Urdu respectively. In Section 6 conclusion and future work are presented.

2 Literature review

The existing literature on the intonation of SALs reports the repeated rising contour (RRC) as the most characteristic unit of SAL prosody. However, there are some areas where SALs significantly differ, such as the surface realization of RRC, the placement of prominence, the contribution in syllable weight, and the number of tonal targets (Khan, 2016). The following sections summarize work on the intonation of four SALs: Hindi, Bengali, Tamil, and lastly Urdu.

2.1 SAL prosody

The intonation of Hindi, a language with substantial overlap with Urdu outside the higher registers of the lexicon and orthography (Masica, 1993), has been analyzed in multiple studies (Moore 1965, Harnsberger 1994, Dyrud 2001, Féry 2010). Early work on Hindi utterances indicates three levels

of phrasing (Moore, 1965), and there exists a hierarchical relationship among these levels where the *foot* is at the lowest level and consists of one or more syllables. The second level is *measure*, which distinguishes focused element from other phrases of the sentence. On highest level is the *sentence*, which aligns with the complete utterance. Moore (1965) shows foot as a domain in which pitch rises from beginning to end. Similar to this, later work by Harnsberger (1994) notes that non-final content words in Hindi bear a rising pitch accent and a phrase boundary (cf. also Moore 1965, Patil et al. 2008). Harnsberger (1994) examines phrase-internal tones in SOV sentences, finding a repetition of LH tones except for the final verb, which takes the boundary tone of the sentence i.e., L% or H% depending upon sentence type. Harnsberger (1994) also noted that the low part of the rising contour is a low pitch accent (transcribed L*, where the * represents lexical stress) and the high part is either the trailing tone from the pitch accent (L*H) or a high boundary tone (H_P), where subscript P represents a phrase boundary lower than the intonation phrase. Sengar & Mannell (2012) later argued that Hindi intonation includes tones on three kinds of prosodic phrases: the accentual phrase (AP), the intermediate phrase (ip), and the intonational phrase (IP), with L*+H as the default pitch accent for Hindi. Dyrud (2001) provides evidence that both pitch and duration show significant correlation with the presence of a stressed syllable.

Stress in Bengali coincides with word-initial syllable, which bears the L* tone. Generally, in Bengali, if a stressed content word is not followed by a prosodic break, it displays a rising contour from the L* of stressed syllable to the edge of the word (Hayes & Lahiri, 1991). Prosodic phrasing in Indian Bengali is shown to exhibit two levels, the phonological phrase (P-phrase) and intonational phrase (I-phrase) (Hayes & Lahiri, 1991). Khan (2014) expanded the two-level hierarchy of Hayes & Lahiri (1991) into a three-level one for Bangladeshi Bengali, introducing the IP, ip, and AP. Khan shows that an ip can be identified by an ip-boundary tone, lengthening of the final syllable, optional pitch reset, and/or optional pause following the ip-final word.

Research on intonation in Tamil, a SAL of the Dravidian language family, suggests two levels of prosodic phrasing: AP and IP. Keane's (2007) analysis indicates that the first syllable in each content morpheme bears the lexical stress and a L* tone followed by a rise in f0 towards the end of the AP. That rise may refer to a boundary tone or may be a part of the bitonal pitch accent L*+H. Content words bear this rise consistently while function words, personal pronouns and demonstrative adjectives lack this rise. The final word of a declarative shows a different pattern where f0 falls abruptly and then declines steadily, due to the presence of a low IP boundary tone.

2.2 Urdu prosody

Stress in Urdu depends on syllable weight, and has been explored by Hussain (2004), who proposed an Urdu phonological stress-marking algorithm. This algorithm classifies Urdu syllables as either monomoraic, bimoraic, or trimoraic. Given these definitions, the algorithm states that a trimoraic (i.e., super-heavy) syllable in final position is stressed; if the final syllable is not trimoraic, then the rightmost non-final bimoraic (i.e., heavy) or trimoraic (i.e., super-heavy) syllable is stressed; and if all syllables are monomoraic (i.e., light), the penultimate syllable is stressed. The stressed syllable attracts a pitch accent. If all the syllables are light, any syllable in a word can get a pitch accent independent of syllable status as stressed or unstressed.

Previous work on the intonation of Urdu (Jabeen, 2010) discusses aspects of the tonal inventory, focusing on the pitch pattern of declarative sentences. Jabeen (2010) reports that the basic pitch pattern of Urdu declarative is L H L-L%. There are other pitch contours such as H L-L% and L L-L%, which can be considered variations of the basic L H L-L% contour. Jabeen et al. (2015) report that transitivity of verbs does not affect the basic pitch pattern of declarative sentences, but the syntactic structure (SOV vs. SVO orders) in broad vs. contrastive focus context can affect the prosodic realization of verb focus in declarative sentences. The tonal pattern is described in terms of the fundamental frequency (f0) contour (rising, falling) and the placement of f0 peaks.

3 Methodology

This section includes the details of data collection for the experiments conducted to formulate the proposed intonation model for Urdu. In Experiment 1, 15 declaratives (see Table 1) were recorded from 13 speakers (7M, 6F) to validate the tonal structure of Urdu across speakers. These speakers,

ranging in age between 25–45 years with 14–16 years of education, are from Lahore, Pakistan, and use Urdu at home and outside. They can also usually understand the regional language Punjabi.

Table 1: Target sentences for Experiment 1

#	Orthography and gloss	Romanization and IPA transcription
1	میرے ہوش و حواس جواب دے رہے ہیں۔ 'My senses are failing.'	*Mere hoš o havās javāb de rahe hãĩ.* /ˈmeːreː ˈhoːʃ oː həˈvɑːs ʤəˈvɑːb ɖeː ˈrəheː hæ̃ː/
2	بچپن سے انہوں نے اب تک پرکشش زندگی گزاری۔ 'From childhood, he has lived a luxurious life.'	*Bacpan se unhõ ne ab tak purkašiš zindagī guzārī.* /ˈbəʧpən seː ˈʊnhõː neː əb ʈək ˈpʊrkəʃɪʃ ˈzɪnɖəgiː gʊˈzɑːriː/
3	اس کے پاؤں تھک گئے۔ 'His feet got tired.'	*Us ke pāõ thak gae.* /ˈʊs keː ˈpɑːõː ˈʈʰək gəeː/
4	درویش کمرے سے نکل گیا۔ 'The saint went out of the room.'	*Darveš kamre se nikal gêā.* /ɖərˈveːʃ ˈkəmreː seː ˈnɪkəl ˈgæɑː/
5	نوجوانوں کو بھی پہرے کی ذمہ داری سونپی گئی۔ 'The young people were also given the responsibility of guarding.'	*Naujavānõ ko bhī pêhre kī zimmedārī saumpī gaī.* /nɔːʤəˈvɑːnõː koː bʱiː ˈpæhreː kiː zɪmmeːˈɖɑːriː ˈsõːmpiː ˈgəiː/
6	اوکاڑہ ہمارا آبائی شہر ہے۔ 'Okara is our hometown.'	*Okāṛā hamārā ābāī šêhêr hai.* /oːˈkɑːɽɑː həˈmɑːrɑː ɑːˈbɑːiː ˈʃæhər hæː/
7	بشارت منت سماجت کرتے رہے۔ 'Basharat kept on insisting.'	*Bašārat minnat samājat karte rahe.* /bəˈʃɑːrəʈ ˈmɪnnəʈ səˈmɑːʤəʈ ˈkərʈeː ˈrəheː/
8	اخوت نے مسلمانوں کو ایک قوم بنا دیا ہے۔ 'Brotherhood has made Muslims a nation.'	*Axuvvat ne musalmānõ ko ek qaum banā diyā hai.* /əˈxʊvvəʈ neː mʊsəlˈmɑːnõː koː eːk ˈqɔːm ˈbənɑː ˈɖɪjɑː hæː/
9	نماز تو وہ ہے جو ہم پڑھ رہے ہیں۔ 'The prayer is what we are offering.'	*Namāz to vo hai jo han paṛ rahe hãĩ.* /nəˈmɑːz ʈoː voː hæː ʤoː həm pəɽ ˈrəheː hæ̃ː/
10	دعا میں گڑگڑانے والے کو اللہ محبوب رکھتا ہے۔ 'God loves the one who laments in prayer.'	*Duā mẽ giṛgiṛāne vāle ko Allāh mêhbūb rakhtā hai.* /ˈɖʊaː mẽː gɪɽgɪˈɽɑːneː ˈvɑːleː koː əlˈlɑːh mæhˈbuːb ˈrəkʰʈɑː hæː/
11	محبت ان کے لیے کافی ہے۔ 'Love is enough for them.'	*Muhabbat un ke lie kāfī hai.* /mʊˈhəbbəʈ ˈʊn keː ˈlɪeː ˈkɑːfiː hæː/
12	ہماری ذمہ داریوں میں غیر معمولی اضافہ ہو گیا ہے۔ 'Our responsibilities have grown significantly.'	*Hamārī zimmedāriõ mẽ ğair māmūlī izāfā ho gêā hai.* /həˈmɑːriː zɪmmeːˈɖɑːrɪõː mẽː ɣæːr mɑːˈmuːliː ɪˈzɑːfɑː hoː ˈgæɑː hæː/
13	ان کے ذہن بدل گئے۔ 'Their minds changed.'	*Un ke zêhên badal gae.* /ˈʊn keː ˈzæhæn ˈbəɖəl ˈgəeː/
14	ان کے دل کی شریانیں خراب ہو گئی تھیں۔ 'The arteries of his heart were damaged.'	*Un ke dil kī šaryānẽ xarāb ho gaī thĩ.* /ʊn keː ɖɪl kiː ʃərˈjɑːnẽː xəˈrɑːb hoː ˈgəiː ʈʰĩː/
15	عذرا نے نعیم کے منہ پر پانی چھڑکا۔ 'Azra sprayed water on Naeem's face.'	*Azrā ne Naīm ke mũ par pānī chiṛkā.* /ˈəzrɑː neː ˈnəiːm keː mũː pər ˈpɑːniː ˈʧʰɪɽkɑː/

For Experiment 2, versions of one sentence were recorded in five different pragmatic contexts (shown in Table 2) from ten speakers (6F, 4M) to determine prosodic differences by sentence type.

Table 2. Target sentences for Experiment 2

#	Target sentence	Utterance type
1	نایاب نے دوکاندار سے لیموں مانگا۔ *Nāyāb ne dukāndār se līmũ̃ māṅgā.* /naːˈjaːb neː ɖuˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgaː/ 'Nayab asked for limes from the shopkeeper.'	Declarative
2	کس نے دوکاندار سے لیموں مانگا؟ *Kis ne dukāndār se līmũ̃ māṅgā?* /ˈkɪs neː ɖuˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgaː/ 'Who asked for limes from the shopkeeper?'	Wh-question
3	کیا نایاب نے دوکاندار سے لیموں مانگا؟ *Kyā Nāyāb ne dukāndār se līmũ̃ māṅgā?* /kjaː naːˈjaːb neː ɖuˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgaː/ 'Did Nayab ask for limes from the shop-keeper?'	Yes/no question
4	دوکاندار سے لیموں مانگو۔ *Dukāndār se līmũ̃ māṅgo.* /ɖuˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgoː/ 'Ask for limes from the shopkeeper.'	Semi-honorific imperative
5	دوکاندار سے لیموں مانگیے۔ *Dukāndār se līmũ̃ māṅgie.* /ɖuˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgɪeː/ 'Please ask for limes from the shopkeeper.'	Polite honorific imperative

The recording sessions took place in an anechoic chamber. Before each session, the speakers were familiarized with the sentences. The speakers were instructed to maintain normal tempo and speaking intensity. The sentences with mispronounced words were re-recorded, repeating the preceding sentence to neutralize any boundary effects.

The recordings from Experiment 1 and Experiment 2 are annotated with tones, break indices (BI), and word boundaries by three trained linguists. A sample of annotated Urdu speech is given below in Figure 1. All examples in the current paper are shown with the blue f0 contour aligning with two labeling tiers: the tone tier with labels for pitch accents and boundary tone, the word tier with the IPA transcription of words along with their boundaries. Below these are English glosses for each word, followed by a sentence-level translation.

The boundary tones and pitch accents are verified using stylized pitch contour using Praat. A total of 20% of speech files are tagged by two different linguists to check inter-annotator accuracy, controlled to be at least 95% in agreement across the annotators using an automated testing tool.

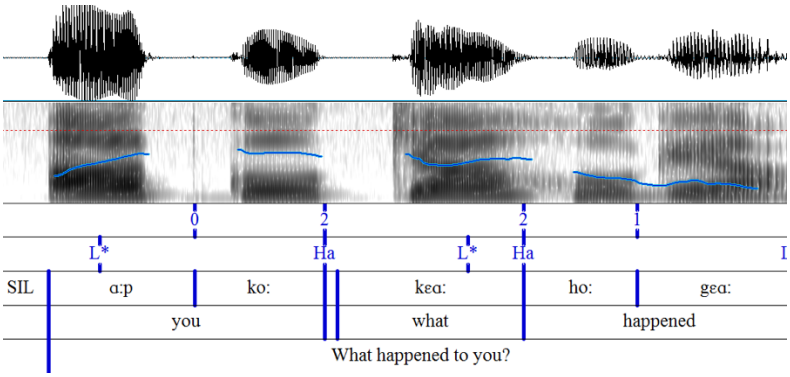

Figure 1. Example of an annotated utterance of Urdu

4 Model of Urdu intonation

The data collected in the experiments presented above provides evidence that the prosodic system of Urdu is composed of three basic pitch accents – low (L*), high (H*) and rising (L*+H) – and boundary tones associated with two prosodic phrases above the word level: AP and IP. Though tone is considered one of the most prominent cues for determining boundaries within phrases, there are also non-tonal cues which can cue the presence of phrases. These include pauses and phrase-final lengthening of segments.

4.1 The accentual phrase (AP)

The basic phrase of Urdu prosody is the AP, which is composed of a pitch accent and AP boundary tone. Pitch accents are those tones that attach to the stressed syllables (Hussain, 1997), and can be either low L*, high H*, or rising L*+H. A high AP boundary tone Ha appears at the right edge of each AP. By using an AP boundary tone, the speaker separates adjacent APs.

The sequence of L* Ha is the default tonal pattern of Urdu AP (covering 65% of the tonal patterns in Experiment 1 as shown in Table 4 and 55% of the tonal patterns in Experiment 2 as shown in Table 5) when AP consists of two syllables as shown in Figure 2 where L* accent is the low pitch accent. The pitch track falls abruptly or stays low in the accented syllable. This pitch pattern of rising contours consisting of either a lexical word or a small phrase is also found in Bengali (Khan, 2014), Korean (Jun, 2005), Tamil (Keane, 2007), and Hindi (Patil et al., 2008).

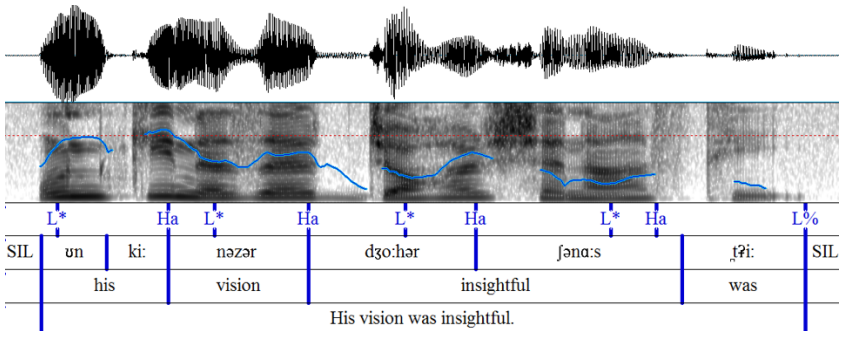

Figure 2. The words *un kī* /ˈʊn kiː/ '3HON.OBL GEN-F', *nazar* /ˈnəzər/ 'vision', and each word in the compound *johar šanās* /ˈdʒoːhər ʃəˈnɑːs/ 'insightful' bear rising APs composed of L* and Ha.

Whenever the AP covers more than two syllables, L* Ha is not the default pattern. When the stressed syllable is non-initial, we see an L target at the beginning of the AP that persists until the L*. This first L target is AP-initial low tone, labeled as aL and is shown on the word *namāz* /nəˈmɑːz/ 'prayer' in Figure 3. AP-initial tones are also found in Korean (Jun, 2004).

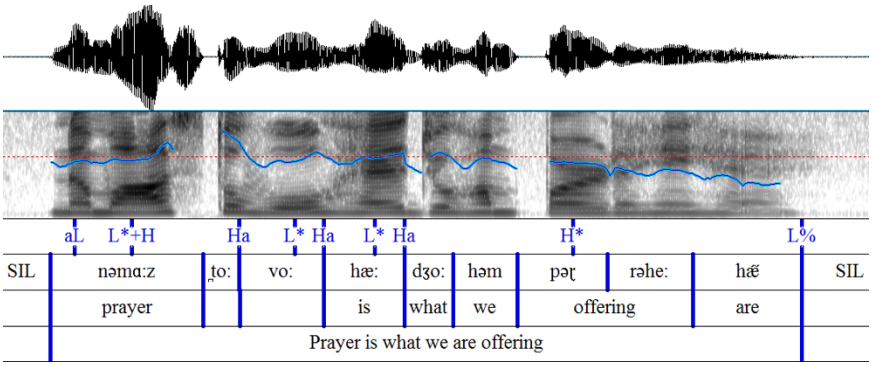

Figure 3. AP-initial aL tone on *namāz* /nəˈmɑːz/ ‘prayer’

This L tone appears on the initial unstressed syllable of an AP, as shown on the words *namāz* /nəˈmɑːz/ ‘prayer’ in Figure 3 and *giṛgiṛāne* /gɪɽgɪˈɽɑːneː/ ‘lament’ in Figure 4.

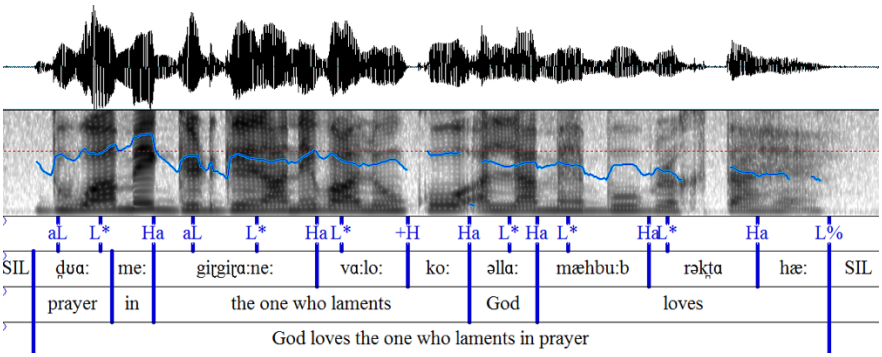

Figure 4. Two examples of the AP-initial aL tone

In many cases the H target is reached earlier than the edge of AP, and then sustained until the AP boundary, suggesting two H targets. This is seen in *Azrā ne* /ˈəzrɑː neː/ ‘Azra ERG’ in Figure 5. Given that the first H target is always immediately post-accentual, and the second H target is always at the AP edge, the data suggest that the first H target is part of a bitonal pitch accent (L*+H) followed by the Ha target for the AP’s right edge. This trailing high (+H), the unstarred half of the pitch accent, often appears when a noun is followed by a case marker or complex postposition.

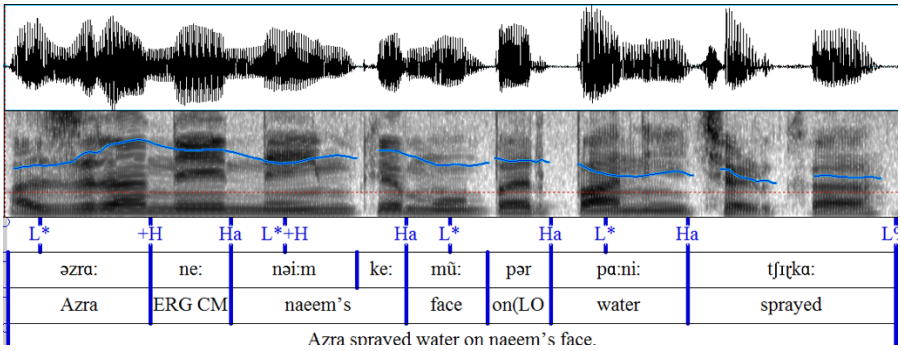

Figure 5. Two examples of a bitonal pitch accent L*+H followed by Ha

Hence, the fully realized AP tonal pattern is (aL) L*(+H) Ha, where the first and third targets are dropped when there is not sufficient duration to bear those tones. An example of the fully realized pattern is provided in Figure 6.

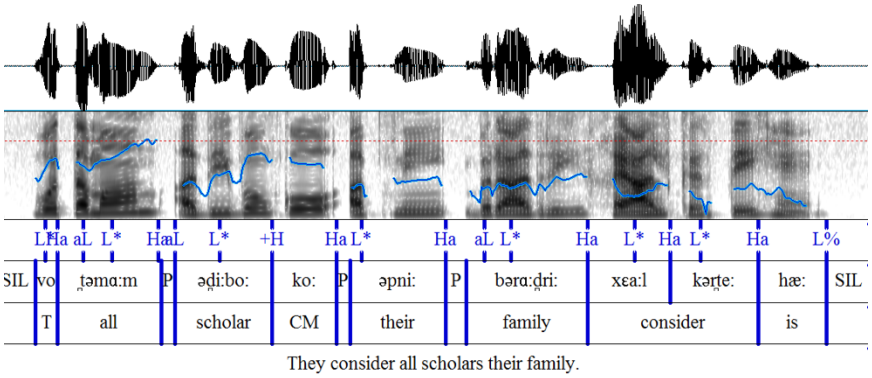

Figure 6. Fully realized AP tonal pattern in *adībo ko* /əˈɖi:bo: ko:/ 'scholars ACC'

The less common AP tonal pattern involves the pitch accent H* La which occurs sentence-finally when the verb is stressed. Due to its typically sentence-final position, this tonal pattern usually does not have AP boundary realized, as any AP boundary tone posited would be overridden by the boundary of a higher prosodic domain i.e., the IP boundary. (Few examples of H* La in non-final position were found in the corpus.) This pitch accent is shown on the word *thak* /ʈʰək/ 'tired' in Figure 7. In this example H* (La) pattern occurs in the sentence with a complex predicate. It is possible that the H* lodges itself on the first member of a complex predicate. However, the relationship between the H* and complex predicates is still undetermined and needs further investigation.

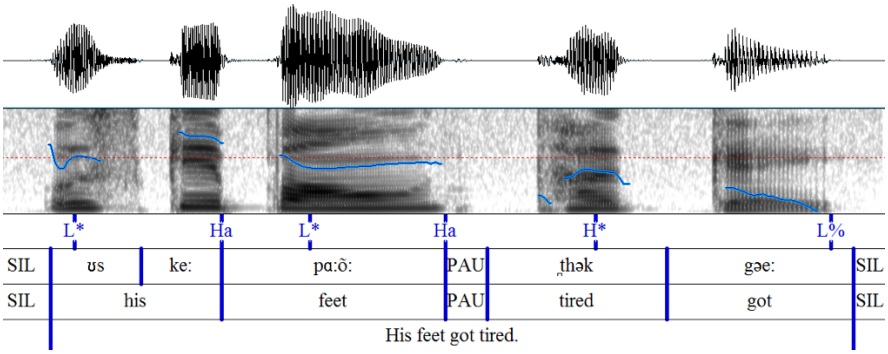

Figure 7. The less common AP tonal pattern H* L%

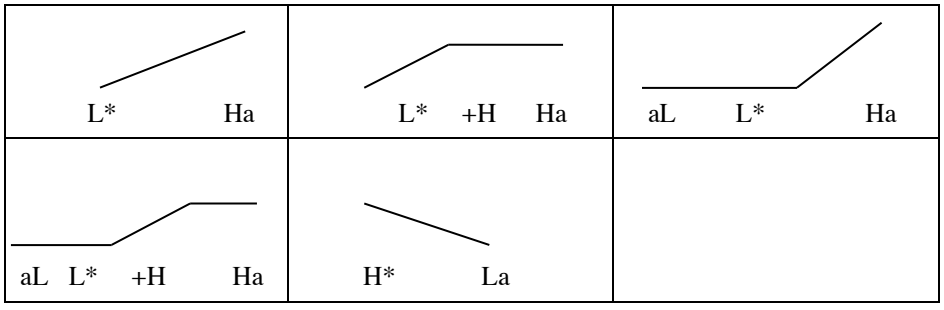

Figure 8. Schematic f0 contours of five type of APs

Schematic f0 contours of five AP patterns and corresponding tone labels are shown in Figure 8.

Like Hindi (Sengar & Mannell, 2012), Korean (Jun, 2005) and Bengali (Khan, 2014), Urdu can have single word APs and multiword APs. For example, in Figure 2, *un* /ʊn/ '3HON.OBL' does not bear a rising contour of its own suggesting that it is phrased with the case marker. This tendency of phrasing with preceding or following words is mostly restricted to some function words (case markers, postpositions, pronouns, and auxiliaries discussed in detail below) as content words seem to bear a rise of their own as can be seen on the words *hamārī* /həˈmɑːr-iː/ '1PL.GEN-F' and *sosāiṭī* /soːˈsɑːɪʈiː/ 'society' in Figure 16. The tendency may also be a matter a speaker choice. Table 3 is an example of a declarative spoken by 13 speakers. The ergative case marker *ne* /neː/ is phrased with *Azrā* /ˈəzrɑː/ 'Azra' by 12 speakers while one speaker has treated these words separately by assigning L* Ha tone on both the words. The postposition *par* /pər/ 'on' is phrased with *mu̐* /mũː/ 'face' by 11 speakers while two speakers have treated these words separately by assigning L*Ha tone on both the words. The contour of the most frequent tonal pattern is given in Figure 5.

Table 3. Interspeaker tonal variation in declarative *Azrā ne Naīm ke mu̐ par pānī chiṛkā.* /ˈəzrɑː neː ˈnəiːm keː mũː pər ˈpɑːniː ˈʧʰɪɽkɑː/ 'Azra sprayed water on Naeem's face'. '—' represents a lack of AP-level tones.

APs	**AP-1**		**AP-2**		**AP-3**		**AP-4**	**AP-5**	**IP**
Words	*Azrā* ˈəzrɑː	*ne* neː	*Naīm* ˈnəiːm	*ke* keː	*mu̐* mũː	*par* pər	*pānī* ˈpɑːniː	*chiṛkā* ˈʧʰɪɽkɑː	
Most frequent pattern	L*+H	Ha	L*+H	Ha	L*	Ha	L* Ha	—	L%
	7		8		4		12	8	13
Speaker-specific variations from most frequent pattern									
SP1					L*+H	Ha			
SP2					L*+H	Ha			
SP3	aL L*	Ha	L*	Ha				L*	
SP4	L* Ha	L* Ha	L* Ha	aL	L* Ha			H*	
SP5	L*	Ha	L*	Ha					
SP6	H*	La			L* Ha	L%			
SP7					L* Ha	L* Ha		H*	
SP8								H*	
SP9									
SP10					L* Ha	L* Ha			
SP11			L* Ha	aL	L*+Ha	Ha		L*	
SP12	H*	La			H*	La	no accent		
SP13	L* Ha	L%	L* Ha	aL	L* Ha	L%			

The analysis of Urdu data reveals five contexts of multiword APs: *izāfat* (4.1.1), conjunctive *vāo* (4.1.2), case markers (4.1.3), complex postpositions (4.1.4), and complex verbs (4.1.5).

4.1.1 Izāfat

Urdu uses a special noun modifier construction known as *izāfat* (also known by its Persian form *ezâfe*), which uses the morpheme *e* /eː/ to link multiple nouns, e.g. *hāl e dil* /hɑːl eː dɪl/ 'condition of heart' and to link a noun with an attributive adjective, e.g. *cašm e nam* /ʧəʃm eː nəm/ 'wet eye' (Chandra & Kumar, 2013). *Izāfat* originates from a relative clause construction and is analyzed synchronically as either a morphological affix (Samvelian, 2007) or a clitic (Butt & King, 2008). Butt & King (2008) argue that *izāfat* licenses a dependency relation between the head noun and a modifier to the right of that NP. Syntactically, *izāfat* is part of the modifying construction. However, prosodically *izāfat* is incorporated into the head noun to its left.

Our findings also align with Butt & King's (2008) findings. In case of *izāfat*, we have observed

that the preceding noun/adjective of *izāfat* carries low pitch accent and H boundary tone aligns itself with *izāfat e* /eː/ rather than aligning with the linked word. In this way, *izāfat* joins with the preceding word and makes one AP. The following noun/adjective is part of a subsequent AP. Prosodically one AP ends on *izāfat* and the new AP starts after *izāfat* as shown in Figure 9.

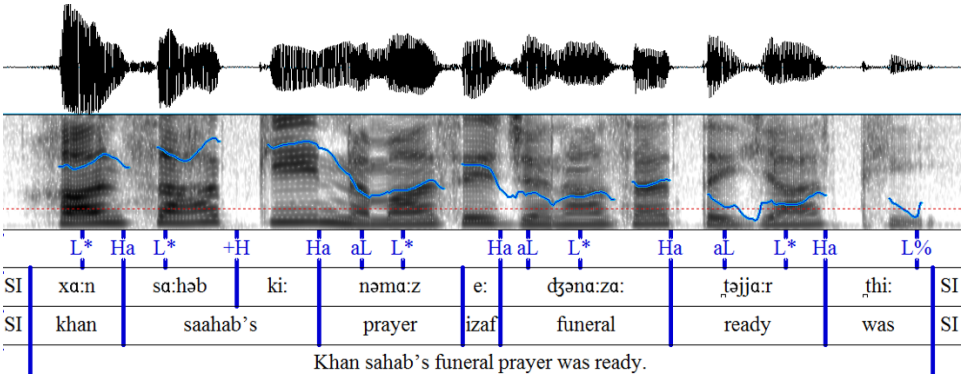

Figure 9. *Izāfat e* /eː/ joining with the preceding word *namāz* /nəˈmɑːz/ 'prayer'

4.1.2 Conjunctive *vāo*

In Urdu, the conjunctive *vāo* 'and' is pronounced as *o* /oː/ is used to conjoin two nouns (Mangrio, 2016). Prosodically, the noun preceding noun the conjunctive *vāo* carries a low pitch accent and H boundary tone aligns itself with conjunctive *vāo* rather than aligning with the noun. In this way, conjunctive *vāo* joins with the preceding noun *hoš* /hoːʃ/ 'consciousness' to form one AP. The conjoined noun *havās* /həˈvɑːs/ 'senses' then forms a separate AP as shown in Figure 10.

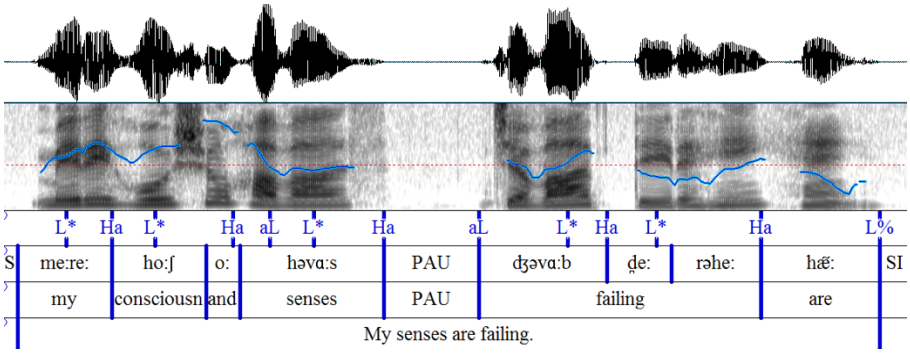

Figure 10. Conjunctive *vāo o* /oː/ phrased with the preceding noun *hoš* /hoːʃ/ 'consciousness'

4.1.3 Case markers

Prosodically, case markers (e.g., *ne* /neː/ 'ERG', *ko* /koː/ 'ACC', *ke* /keː/ 'GEN', *mẽ* /mẽː/ 'LOC') behave differently when they are attached with monosyllabic pronouns versus when they are attached with polysyllabic nouns. When a monosyllabic pronoun is followed by a case marker, the intervening word boundary is prosodically elided. In Figure 7, the Ha boundary tone is realized on the case marker *ke* /keː/ 'GEN' instead of the pronoun *us* /ʊs/ '3SG.OBL' indicating two words are uttered as one AP. The pronoun and case marker are sometimes also joined orthographically.

For a polysyllabic noun followed by a case marker, however, the H target is often reached earlier than the AP edge, and then prolonged through the boundary, suggesting two H targets. This is clearly seen on the phrase *Āgrā mẽ* /ˈɑːgrɑː mẽː/ 'in Agra' in Figure 11. Hence, in case of both nouns and pronouns followed by case markers, the case markers are merged with the preceding noun/pronoun forming one AP. But pitch realization is different in both the cases as shown in Figure 7 for pronouns and Figure 11 for nouns.

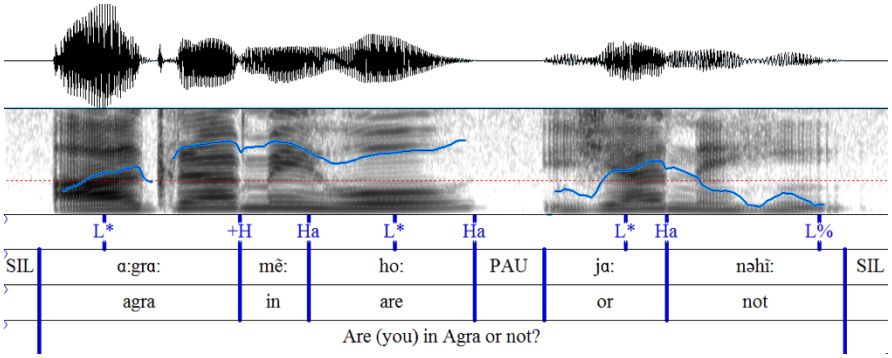

Figure 11. Noun and case marker phrased as one AP

4.1.4 Complex postpositions

Complex postpositions, which involve a case marker followed by a postposition, generally form a single AP apart from the noun. For example, the case marker *ke* /keː/ 'GEN' takes low initial AP boundary aL, and a L* can be observed on first syllable of the postposition *liye* /ˈlɪjeː/ 'for' as shown in Figure 12. The case marker and postposition are sometimes also joined orthographically.

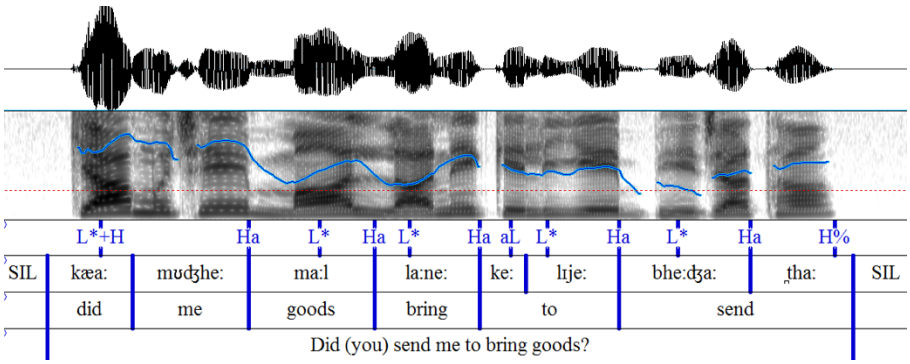

Figure 12. Complex postposition *ke liye* /keː ˈlɪjeː/ 'GEN for' forming one AP

4.1.5 Verb auxiliaries

A complex verb forms a single AP. For example, the main verb *de* /ɖeː/ 'give' takes L* and the following auxiliary *rahe* /ˈrəheː/ 'PRG' takes the Ha boundary in Figure 10.

4.2 The intonation phrase (IP)

The IP in Urdu is a group of APs, typically spanning over a clause or a sentence. The IP is the highest unit in Urdu prosodic hierarchy and is marked by the presence of final lengthening, pause

and one of the three boundary tones; low (L%), high (H%) and a rising (LH%). The selection of final boundary tone is dependent on the particular sentence type.

The most frequent IP boundary tone is L% which is found at the end of declarative sentences. This tone is realized as falling pitch in the IP-final syllable as shown on the word *thī* /t̪ʰiː/ 'was' in Figure 2. The second IP boundary tone is H% is found most frequently at the end of yes/no questions (8/10 as shown in Table 9) and less frequently at the end of semi-honorific imperatives and wh-questions (6/10 as shown in Table 10 and 5/10 as shown in Table 8 respectively). This tone is realized as sharply rising f0 in the IP-final syllable and is placed at the end of the phrase, as shown in Figure 13.

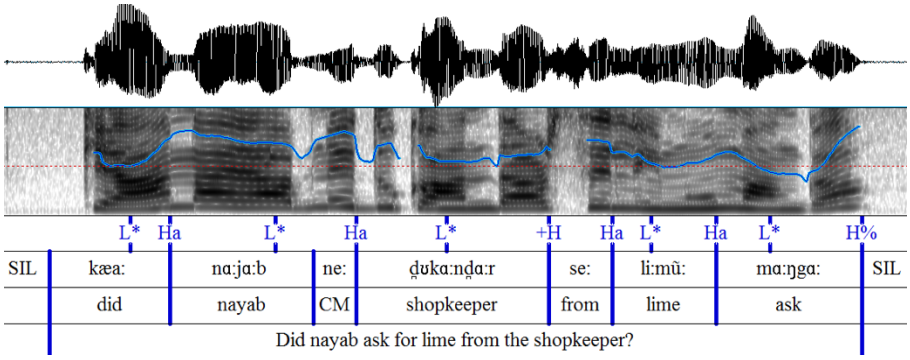

Figure 13. High IP boundary tone H%

The third IP boundary tone is LH%, found most frequently at the end of polite honorific imperatives (6/10 as shown in Table 11) and less frequently at the end of wh-questions (2/10 as shown in Table 8). LH% and H% boundary tones are different; H% tone depicts an abrupt rise in the final syllable while LH% boundary tone starts with sustained low pitch following a rise as shown in Figure 14.

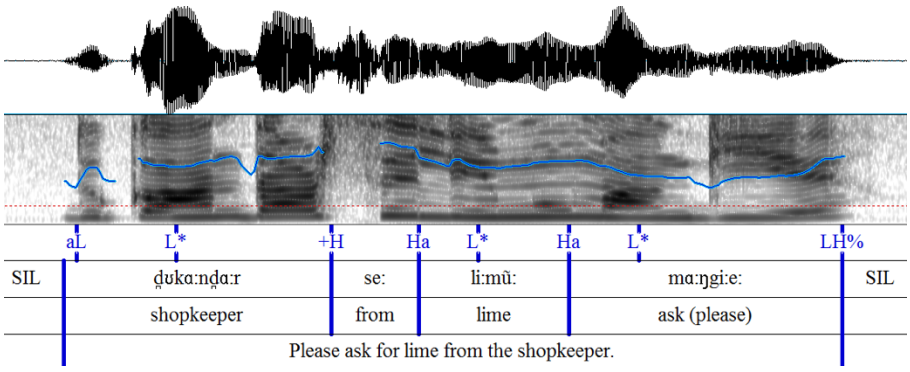

Figure 14. Rising IP boundary tone LH%

Schematic f0 contours of three types of IP boundary tone realizations are shown in Figure 15. The vertical line shown in each contour marks the beginning of the IP-final syllable.

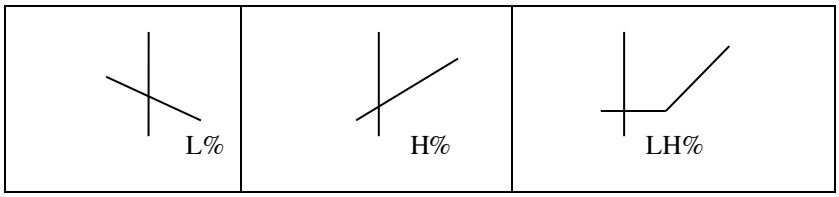

Figure 15: Schematic f0 contours of three types of IP boundary tones

Tables 4 and 5 show the occurrence frequency of all the accentual phrase patterns and boundary tones in Experiment 1 and Experiment 2 respectively. It is evident from the results that L*Ha is the default tonal pattern covering 65% of all the tonal patterns in Experiment 1 as shown in Table 4 and 55% of all the tonal patterns in Experiment 2 as shown in Table 5. H* (La) is the least frequent tonal patterns as they cover 2.97%/2.38% of all the tonal patterns and 6.25%/2.5% of all the tonal patterns in the two experiments respectively. Further L% is the most frequent IP boundary tone covering 84% of all the tonal patterns in Experiment 1. L% is not frequent in experiment two because the data of experiment two contains different sentence types: wh-questions, Yes/no questions and imperatives, which use different boundary tones for IP.

Table 4. Tone counts from Experiment 1

AP tones	**Count (Total = 839)**	**Percentage**
L* Ha	547	65.0%
L*+H Ha	107	12.0%
aL L* Ha	113	13.0%
aL L*+H Ha	27	3.2%
H* La	20	2.3%
H*	25	2.9%
IP tone	**Count (Total = 179)**	**Percentage**
L%	152	84.0%
H%	8	4.4%
LH%	19	10.6%

Table 5. Tone counts from Experiment 2

AP tones	**Count (Total = 160)**	**Percentage**
L*Ha	88	55.0%
L*+H Ha	28	17.5%
aL L*Ha	18	11.2%
aL L*+H Ha	12	7.5%
H* La	4	2.5%
H*	10	6.2%
IP tone	**Count (Total = 45)**	**Percentage**
L%	16	35.0%
H%	18	40.0%
LH%	11	24.4%

4.3 Other phrase types

The analysis of data suggests that Urdu has two prosodic phrases (i.e., AP, IP), with no ip in the data analyzed for the current study. Khan (2014) says that in Bangladeshi Bengali a high ip H-

boundary tone reaches a higher pitch than the corresponding high AP boundary tones which phonologically determines the presence of ips. This is not observed in case of Urdu as shown in Figure 16. There is pause and final lengthening on the word *vāqif* /ˈvɑːqɪf/ 'familiar' but the pitch height at the right edge is lower from the height of other high AP boundary tones in the sentence indicating absence of an ip in Urdu. In future, focus realization and its impact on the intonation patterns will be studied with reference to ips.

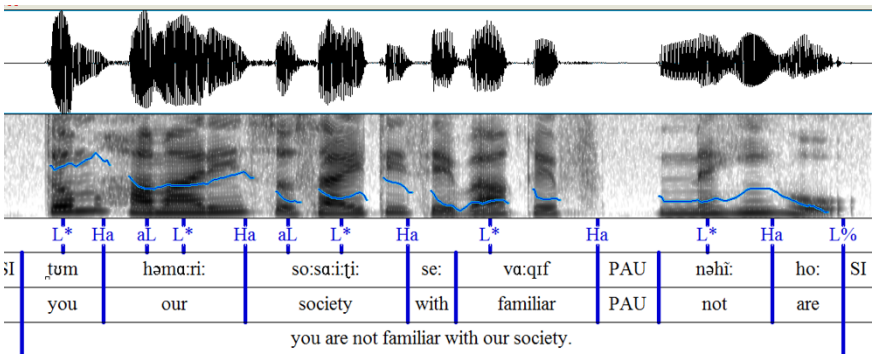

Figure 16. Absence of pitch reset

5 Sentence-specific prosody

To determine the differences in prosody for the various types of sentences, the sentence *Nāyāb ne dukāndār se līmū̃ māṅgā* /nɑːˈjɑːb neː ˌɖʊˈkɑːnˌɖɑːr seː ˈliːmũː ˈmɑːŋgɑː/ 'Nayab asked for limes from the shopkeeper' or its relevant variations are recorded in five different syntactic contexts from ten speakers. These contexts included declaratives, wh-questions, yes/no questions, semi-honorific imperatives, and polite honorific imperatives.

5.1 Declaratives

All declarative sentences consist of a series of APs with rising f0 contour within each AP: (aL) L* (H) Ha; except the sentence final AP which has a (high) falling f0 contour: (H)L%. The same trend is reported for Bengali (Khan, 2014), French (Jun & Fougeron, 2002), and Korean (Jun, 2005). Table 6 is an example of a declarative sentence spoken by ten speakers. The most frequent tonal pattern is written on the top row with corresponding APs over which those tones are realized. The speaker-wise variations from that most frequent pattern are given in the rows below. The pitch range of each speaker for this particular sentence is given in the last column. The table shows that eight speakers have shown final lowering of f0 suggesting the L% IP boundary. The pitch track of most frequent contour for declarative sentence is given in Figure 17.

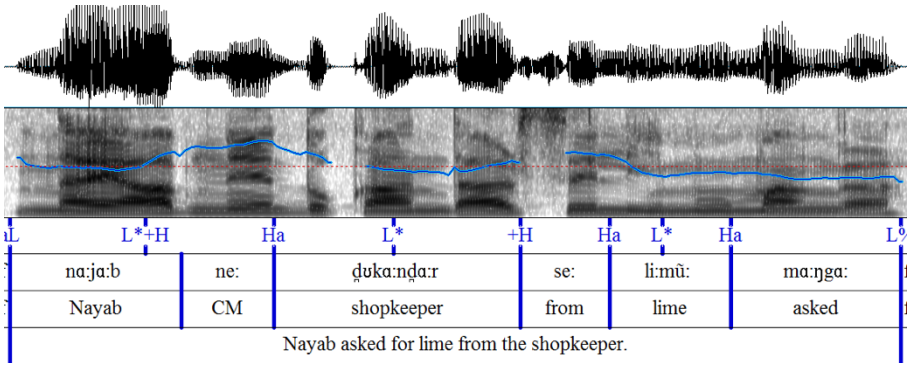

Figure 17. Pitch track of most frequent contour for declarative sentence

Table 6. Interspeaker tonal variation in declarative *Nāyāb ne dukāndār se līmū̃ māṅgā* /nɑːˈjɑːb neː ɖ̪ʊˈkɑːnɖ̪ɑːr seː ˈliːmũː ˈmɑːŋgɑː/ 'Nayab asked for lime from the shopkeeper'. '—' represents a lack of AP-level tones.

APs	**AP-1**		**AP-2**		**AP-3**	**AP-4**	**IP**	**Range (Hz)**
Words	*Nāyāb* nɑːˈjɑːb	*ne* neː	*dukāndār* ɖ̪ʊˈkɑːnɖ̪ɑːr	*se* seː	*līmū̃* ˈliːmũː	*māṅgā* ˈmɑːŋgɑː		
Most frequent pattern[1]	aL L*+H	Ha	L*+H	Ha	L* Ha	—	L%	
	4		4		7	5	8	
Speaker-specific variations from most frequent pattern								
SP1						L*		265:175
SP2	aL L* Ha	L* Ha	L*	Ha				266:182
SP3			aL L*	Ha		H*		233:172
SP4			aL L*	Ha	—	L*	LH%	306:195
SP5			L* Ha L*	Ha	—	H*		141:89
SP6	aL L*+H	Ha				H*		121:88
SP7	aL L*+H	Ha	L* Ha L*+H	Ha				356:186
SP8	aL L*+H	Ha						378:218
SP9	aL L*+H	Ha						190:121
SP10	aL L* Ha	L* Ha	L*Ha	L* Ha	—		LH%	265:192

5.2 Wh-questions

In Urdu, the default position for the wh-phrase is at the beginning of the sentence. The wh-phrase is then followed by indirect object and direct object (of the verb) while the verb is in sentence-final position. These questions are distinguished from their corresponding declaratives in terms of final boundary tone and pitch range of the question word. The measurement of f0 range for the question word shows differences between wh-sentences and corresponding declarative sentences. Pitch range differences in semitone are given in Table 7.

Table 7. Pitch range difference for AP containing wh-word *kis ne* /ˈkɪs neː/ 'who ERG'

Speaker	**f0 max:min (Hz)**	**f0 range difference (semitones)**
Female speakers		
SP1	357:270	4.83
SP2	306:255	3.15
SP4	343:312	1.63
SP7	410:321	4.23
SP8	373:297	3.94
SP10	310:250	3.72
Male speakers		
SP3	250:223	1.97
SP5	134:124	1.34
SP6	131:119	1.66
SP9	191:174	1.61

[1] On the first AP *Nāyāb ne* /nɑːˈjɑːb neː/ 'Nayab ERG', the most frequent tonal patterns were aL L* Ha and aL L*+H Ha, spoken by equal no of speakers i.e. 4 and 4 respectively. However, aL L* Ha is written in the top row (of most frequent tones) because this pattern is simpler than the other.

Out of ten speakers, seven (6F, 1M) showed a wider pitch range in wh-sentences than in the declaratives. Figure 18 shows one such pair, uttered by the same speaker. A similar phenomenon is reported for Tamil as well (Keane, 2007).

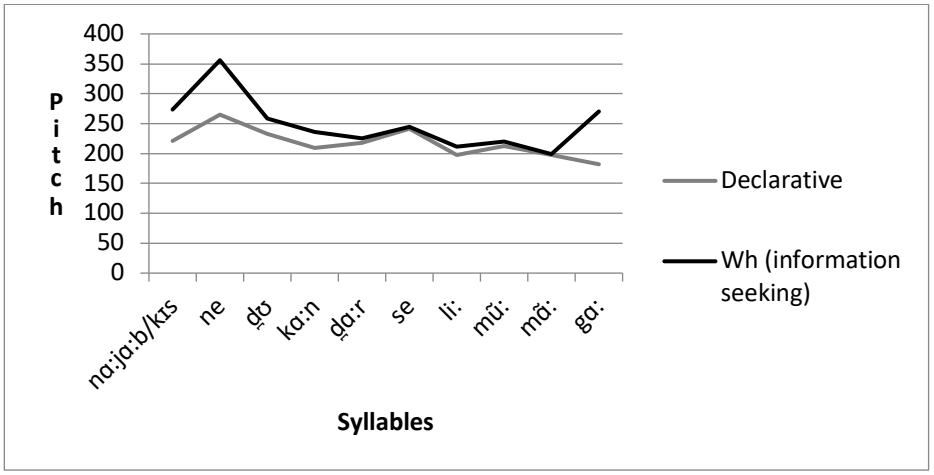

Figure 18. Declarative vs. wh-question uttered by the same speaker

Apart from pitch range, final boundary tone also exhibits differences between wh-sentences and their corresponding declaratives. Table 8 shows that out of ten speakers, seven showed rising tone at the end of the sentence. Of those seven, five showed H% boundary tone and two showed LH% boundary tone. The remaining three speakers did not show H% or LH% boundary tones, yet they treated the sentence-final verb differently from that of their corresponding declaratives by showing H* pitch accent. The pitch track of most frequent contour for wh-questions is given in Figure 19.

Table 8. Interspeaker tonal variation in wh-question *kis ne dukāndār se līmū̃ māṅgā* /ˈkɪs neː ɖʊˈkɑːnɖɑːr seː ˈliːmũː ˈmɑːŋgɑː/ 'Who asked for lime from the shopkeeper?' '—' represents a lack of AP-level tones.

APs	**AP-1**		**AP-2**		**AP-3**	**AP-4**	**IP**	**Max:Min (Hz)**
Words	*kis* ˈkɪs	*ne* neː	*dukāndār* ɖʊˈkɑːnɖɑːr	*se* seː	*līmū̃* ˈliːmũː	*māṅgā* ˈmɑːŋgɑː		
Most frequent pattern	L*	Ha	L*	Ha	L*Ha	L*	H%	
	7		5		7	5	5	
Speaker-specific variations from most frequent pattern								
SP1								355:195
SP2						H*	L%	301:188
SP3						H*	L%	225:170
SP4	L* Ha	L* Ha					LH%	341:200
SP5					—	H*	L%	134:89
SP6			L*+H	Ha		—		128:93
SP7			L*+H	Ha	—			410:222
SP8			L*+H	Ha				373:205
SP9	L* Ha	L* Ha	L*+H	Ha				190:121
SP10	L* Ha	L* Ha	L*+H	Ha	—	—	LH%	303:171

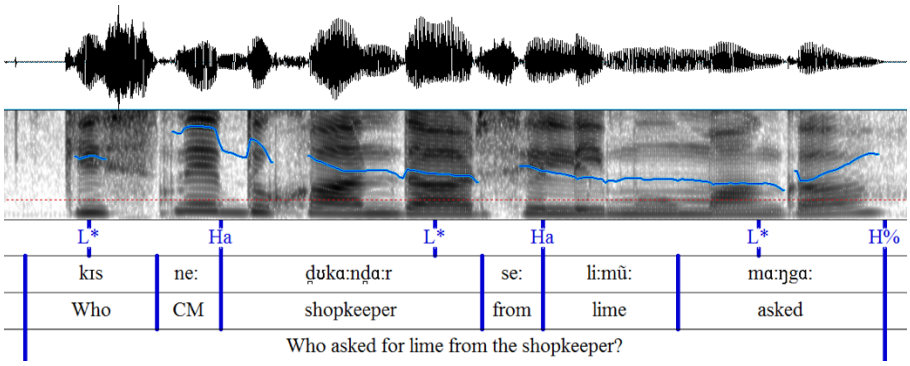

Figure 19. Pitch track of most frequent contour for wh-questions

5.3 Yes/no questions

Yes/no questions in Urdu are formed by adding an initial *kyā* /kjaː/ (lit. 'what'). Yes/no questions are distinguished prosodically in terms of IP tone. Table 9 shows that out of ten speakers, nine showed a rising tone (eight by showing H% boundary tone and one by showing LH% boundary tone). Only one speaker did not show H% or LH% boundary tone, treating the sentence final verb differently from its corresponding declarative by showing H* pitch accent on it. The pitch track of most frequent contour for yes/no questions is given in Figure 20.

Table 9. Interspeaker tonal variation in yes/no question *kyā Nāyāb ne dukāndār se līmū̃ māṅgā?* /kjaː naːˈjaːb neː ɖuˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgaː/ 'Did Nayab ask for lime from the shopkeeper?' '—' represents a lack of AP-level tones.

APs	**AP-1**	**AP-2**		**AP-3**		**AP-4**	**AP-5**	**IP**	**Max:Min (Hz)**
Words	*kyā* kjaː	*Nāyāb* naːˈjaːb	*ne* neː	*dukāndār* ɖuˈkaːnɖaːr	*se* seː	*līmū̃* ˈliːmũː	*māṅgā* ˈmaːŋgaː		
Most frequent pattern	L* Ha	L*	Ha	L*+H	Ha	L* Ha	L*	H%	
	10	5		6		7	8	8	
Speaker-specific variations from most frequent pattern									
SP1									314:205
SP2				L*	Ha	—			275:198
SP3		aL L*	Ha	aL L*	Ha		—		234:196
SP4		aL L*	Ha	L*	Ha			LH%	369:215
SP5				L*	Ha	—	H*	L%	126:89
SP6		L*+H	Ha						113:95
SP7									389:235
SP8		L*+H	Ha						347:243
SP9									181:120
SP10		L*+H	Ha			—			290:207

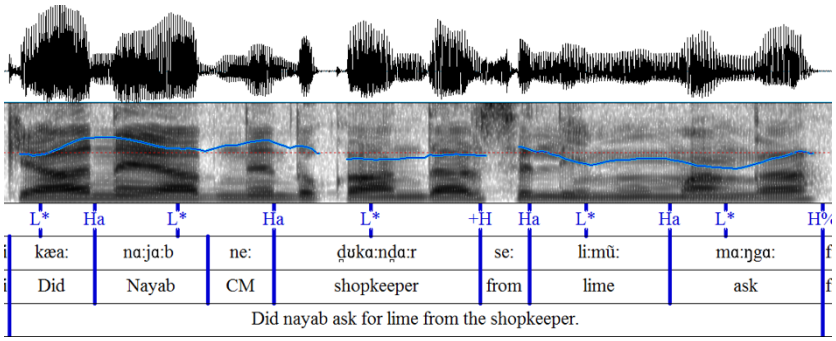

Figure 20. Pitch track of most frequent contour for Yes/no questions

5.4 Imperative sentences

In Urdu, imperatives have four forms: (i) non-honorific (e.g. *de* /ɖeː/ 'give'), used with the singular non-honorific pronoun *tū* /ʈuː/, (ii) semi-honorific (e.g. *do* /ɖoː/), used with the plural and/or semi-honorific pronoun *tum* /ʈʊm/, (iii) subjunctive honorific (e.g. *dẽ* /ɖẽː/), used with the honorific pronoun *āp* /ɑːp/, and (iv) polite honorific (e.g. *dījiye* /ˈɖiːdʒɪjeː/), also used with the honorific pronoun *āp* /ɑːp/, cf. Koul (2008) and Platts (1874). For this study, we have selected the semi-honorific imperatives and polite honorific imperatives. Semi-honorific imperatives differ from their corresponding declaratives in terms of final boundary tone. Figure 21 shows the most frequent contour for semi-honorific imperatives.

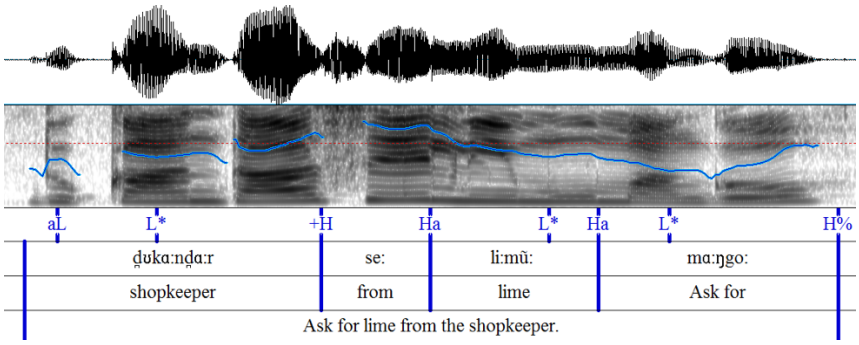

Figure 21. Pitch track of most frequent contour for semi-honorific imperatives

Table 10 indicates that out of ten speakers, seven showed rising pitch (six by using H% and one by using LH%) on the verb. The remaining three did not show this trend, yet two of the three treated the imperative verb differently from corresponding declarative by using a H* pitch accent. Polite honorific imperatives differ from their corresponding declaratives in terms of IP boundary tone. In polite honorific imperatives, the LH% outnumbers the H% seen in semi-honorific imperatives. The pitch track of most frequent contour for polite honorific imperatives is given in Figure 22.

Table 10. Speaker variation in semi-honorific imperative *dukāndār se līmṹ māŋgo* /ɖʊˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgoː/ ‘Ask for lime from the shopkeeper.’ ‘—’ represents a lack of AP-level tones.

APs	AP-1		AP-2	AP-3	IP	Max:Min (Hz)
Words	*dukāndār* ɖʊˈkaːnɖaːr	*se* seː	*līmṹ* ˈliːmũː	*māṅgo* ˈmaːŋgoː		
Most frequent pattern	aL L*+H	Ha	L* Ha	L*	H%	
	7		9	8	6	
Speaker-specific variations from most frequent pattern						
SP1	aL L*	Ha				274:192
SP2	aL L*	Ha				257:205
SP3	aL L*	Ha		H*	L%	229:158
SP4						339:208
SP5			—	H*	L%	132:90
SP6						117:92
SP7					L%	362:189
SP8						342:219
SP9						168:122
SP10					LH%	293:194

Table 11 indicates that out of ten speakers, eight speakers showed a rising pitch, six by showing LH% boundary tone, and two by showing H% IP boundary on the final verb. These counts are reversed in case of semi-honorific imperatives. Remaining two speakers did not show this trend, yet they treated the command verb differently from declaratives by showing H* pitch accent.

Table 11. Speaker variation in polite honorific imperative *dukāndār se līmṹ māṅgie* /ɖʊˈkaːnɖaːr seː ˈliːmũː ˈmaːŋgɪeː/ ‘Please ask for lime from the shopkeeper’. ‘—’ represents a lack of AP-level tones.

APs	AP-1		AP-2	AP-3	IP	Max:Min (Hz)
Words	*dukāndār* ɖʊˈkaːnɖaːr	*se* seː	*līmṹ* ˈliːmũː	*māṅgie* ˈmaːŋgɪeː		
Most frequent pattern	aL L*+H	Ha	L* Ha	L*	LH%	
	7		8	8	6	
Speaker wise differences in tones						
SP1	aL L*	Ha		aL H*	L%	262:192
SP2	aL L*	Ha				270:205
SP3	aL L*	Ha				230:186
SP4						364:198
SP5			—	H*	L%	123:90
SP6					H%	112:95
SP7						360:220
SP8					H%	332:227
SP9						164:121
SP10			—			261:168

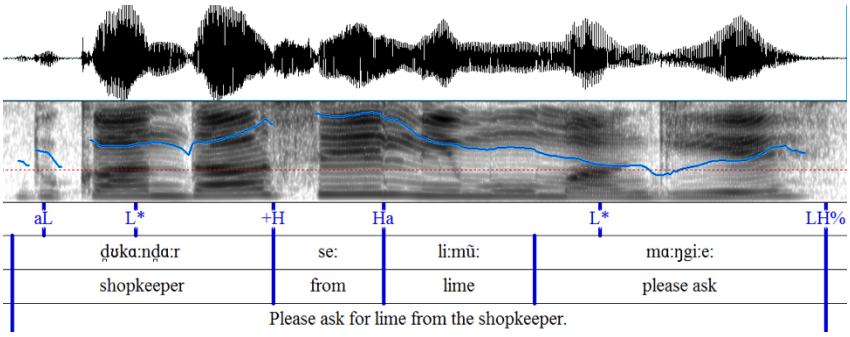

Figure 22. Pitch track of most frequent contour for polite honorific imperatives

6 Conclusion and future work

The corpus analysis suggests that Urdu has two phrase levels: the AP and IP, which correspond to the phrase levels found in Tamil (Keane, 2007) and Indian Bengali (Hayes & Lahiri, 1991). The IP in Urdu corresponds roughly to the sentence level, and the AP corresponds roughly to one or more words. This phrase-level correspondence of APs and IPs has also been confirmed for Hindi (Sengar & Mannell, 2012) and Bengali (Khan, 2014). The present study also explores the contexts where AP contains more than one word, with *izāfat*, conjunctive *vāo*, case markers, postpositions, and verb auxiliaries. It is found that Urdu tonal inventory includes the pitch accents, AP tones, and IP tones in Table 13.

Table 12. Urdu tonal inventory

AP-initial tone	Pitch accents	AP-final tones	IP-final tones
aL	L* L*+H H*	Ha La	L% LH% H%

The intonation model developed through the experiments was used to annotate a corpus of 1285 sentences. This corpus was extracted from an already existing text corpus used for the development of Text to Speech system (TTS) (Habib et al. 2014). The larger corpus includes sentences that were carefully chosen to include sonorant consonants and vowels to facilitate pitch analysis. Different sentence types selected include 951 declarative sentences, 183 interrogative sentences (Yes/no questions, wh-questions) and 151 imperative sentences. This corpus will be used to develop natural sounding Urdu TTS.

The study also investigates the pitch pattern of declarative, interrogative, and imperative sentences. Results indicate that (i) declarative sentences consist of a series of APs with rising f0 contours within each: (aL) L* (H) Ha, except the sentence-final AP which has a falling f0 contour: (H*) L% (ii) wh-questions differ from their corresponding declaratives in terms of pitch range of the question word and the IP-final boundary tone; (iii) imperatives differ from their corresponding declaratives in terms of IP-final boundary tone. In the future, focus realization and its impact on the intonation patterns in Urdu will be studied in more detail.

References

Butt, M., & King, T. H. (2008). Urdu *ezafe* and the morphology–syntax interface. *Proceedings of the 8th Conference on Lexical Functional Grammar.*

Chandra, P., & Kumar, R. (2013). Urdu Possession: An instance of *ezafe*. *Indian Linguistics* 74, 101–109.

Dyrud., L. O. (2001). *Hindi-Urdu: Stress accent or non-stress accent?* Doctoral dissertation. University of North Dakota.

Face, T. L. (2003). Intonation in Spanish declaratives: difference between lab speech and spontaneous speech. *Catalan Journal of Linguistics* 2, 115–131.

Féry, C. (2010). Indian Languages as Intonational Phrase Languages. In Hasnain, I., & Chaudhury, S. (eds.) *Festschrift to honour Ramakant Agnihotri*. Delhi: Aakar Publisher.

Grabe, E. (1997). Comparative intonational phonology: English and German. *Intonation: Theory, Models and Applications,* 157–160.

Grice, M., & Savino, M. (2003). Map tasks in Italian: Asking questions about given, accessible and new information. *Catalan journal of linguistics* 2, 153–180.

Habib, W., Basit, R. H., Hussain, S., & Adeeba, F. (2014). Design of Speech Corpus for Open Domain Urdu Text to Speech System Using Greedy Algorithm. *Proceedings of Conference on Language and Technology 2014 (CLT14).* Karachi.

Harnsberger, J. D. (1994). *Towards an intonational phonology of Hindi*. Gainesville: University of Florida.

Hayes, B., & Lahiri, A. (1991). Bengali intonational phonology. *Natural Language & Linguistic Theory* 9(1), 47–96.

Hussain, S. (1997). *Phonetic correlates of lexical stress in Urdu.* Doctoral dissertation, Northwestern University.

Hussain, S. (2004). Letter to-sound conversion for Urdu text-to-speech system. *Proceedings of the workshop on computational approaches to Arabic script-based languages*, 74–79. Association for Computational Linguistics.

Jabeen, F. (2010). *An Acoustic Study of the Influence of Urdu on the Intonation Patterns of English in Pakistan.* Doctoral dissertation, Government College University.

Jabeen, F., Bögel, T., & Butt, M. (2015). Immediately postverbal questions in urdu. *INTERSPEECH*, (pp. 943-947). Dresden.

Jun, S.-A. (2004). Intonational phonology of Seoul Korean revisited. *14th Japanese/Korean Linguistics conference.* Arizona.

Jun, S.-A. (2005). Korean intonational phonology and prosodic transcription. *Prosodic typology: The phonology of intonation and phrasing vol. 1*. Oxford University Press.

Jun, S.-A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus 14 (1)*, 147-172.

Keane, E. (2007). Distribution and alignment of F0 contours in Tamil. In *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken.

Khan, S. D. (2014). The intonational phonology of Bangladeshi Standard Bengali. In Jun, S.-A. (ed.) *Prosodic typology: The phonology of intonation and phrasing vol. 2.* Oxford University Press.

Khan, S. D. (2016). The intonation of South Asian languages. *Proceedings of Formal Approaches to South Asian Languages 6*, 23–36.

Kiruthiga, S., & Krishnamoorthy, K. (2012). Annotating Speech Corpus for Prosody Modeling in Indian Language Text to Speech Systems. *International Journal of Computer Science Issues* 9.

Koul, O. N. (2008). *Modern Hindi Grammar.* Dunwoody.

Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech communication*, 391–405.

Mangrio, R. A. (2016). *The Morphology of Loanwords in Urdu: The Persian, Arabic and English Strands.* Cambridge Scholars Publishing.

Masica, C. P. (1993). *The Indo-Aryan Languages.* Cambridge University Press.

Moore, R. R. (1965). *A study of Hindi intonation.* Doctoral dissertation, University of Michigan.

Patil., U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., & Vasishth, S. (2008). Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics* 1.

Platts, J. T. (1874). *A Grammar of the Hindustani or Urdu language.* London: WH Allen.

Samvelian, P. (2007). A (phrasal) affix analysis of the Persian *ezafe*. *Journal of Linguistics* 43, 605–645.

Sengar, A., & Mannell, R. (2012). A preliminary study of Hindi intonation. In *Proceedings of Speech and Science Technology.* Sydney: Australasian Speech Science and Technology Association.

On some effects of utterance finality, with special consideration of South Asian languages

HANS HENRICH HOCK, *University of Illinois at Urbana-Champaign*

Received JANUARY 2019; Revised OCTOBER 2019

ABSTRACT

Pitch lowering, avoidance of prosodic prominence, and segmental reductions in utterance-final position are well known crosslinguistic tendencies. In verb-final languages the prosodic effects of Utterance Finality intersect with an independent, crosslinguistic tendency of verbs to receive relatively weak prominence within larger prosodic domains. As a consequence, verbs in SOV languages are special targets for the effects of Utterance Finality. After providing crosslinguistic illustrations of these effects I focus on a number of phenomena in South Asian languages which can be explained in terms of the intersection between Utterance Finality and Verb Finality. These include the relative order of negation and verb and the apparent optionality of 'be'-deletion in Hindi, the difference in verb accentuation between main and dependent clauses in Vedic, and (possibly) the fact that Kashmiri *ki/zi*-clauses, unlike relative clauses, have V2, rather than verb-final order.

1 Introduction

As a historical process, the crosslinguistically common phenomenon of final devoicing (Hock 1986/1991: 88, 92–93, 95–96) has been known at least since the 19th century, and as a synchronic phenomenon it has figured prominently in both Praguian and generative phonology. Although its domain is commonly considered the word, it has also been suggested that it originates in a larger prosodic domain, i.e. the utterance. In Hock (1999) I have presented arguments and evidence for the view that utterance-final position is indeed the most likely point of departure, not only for final devoicing but for many other processes or phenomena, both segmental and suprasegmental, that are commonly referred to as word-final; that word-final processes or phenomena generally result from extension of utterance-final developments; and that utterance-final changes can, by similar extensions, give rise to other, even more far-reaching developments.

This paper focuses on Utterance Finality and its effects on verbs in verb-final languages, with special attention to South Asian languages. Phenomena to be considered include apocope and accent retraction in finite verbs and their consequences, as well as apparent optionality in the order of negation and verb and in the deletability of 'be' in Hindi.

Section 2 establishes the conceptual framework for the rest of the paper, arguing for the utterance as the relevant domain for the issues to be discussed, rather than the word or the sentence. Section 3 addresses the prosodic properties of utterance finality and their interaction with the tendency for verbs to be less prominent in context than nouns or NPs, resulting in the fact that verbs in SOV languages are subject to special segmental and suprasegmental changes. Section 4 illustrates synchronic consequences in Hindi of the interaction of utterance finality and the prosodic weakness of verbs. Section 5 addresses historical developments in South Asian languages that can be explained by this interaction. Finally, Section 6 presents conclusions and discusses further implications.

2 Word finality, sentence finality, utterance finality, and the relation between syntax and phrasal prosody

Summarizing arguments in Hock (1999), this section establishes a conceptual framework for the rest of the paper by arguing for Utterance Finality as a more insightful prosodic motivation than word or sentence finality for phenomena such as final devoicing or accent retraction. In addition, following earlier work such as Selkirk (1984), Nespor & Vogel (1986), it accepts the position that there is no one-to-one correspondence between syntax and prosodic phrasing.

2.1 Word finality vs. utterance finality

Opinions have varied as to whether words are the proper focus of phonology (both segmental and suprasegmental) or larger domains such as the sentence or the utterance. The Praguians opted for a *phonologie du mot* and so did many generativists. Exceptions include Vennemann (1974), Hooper (1976), and Hyman (1977) who account for final devoicing or neutralization as an extension of utterance finality effects. The situation is similar in phonetics. Ohala (1993) claimed that the "domain of change is overwhelmingly the word or possibly phrases which occur so often that they could also be said to be lexicalized." In contrast, Keating (1988) argued that "final devoicing of obstruents can be motivated physically by aerodynamic considerations, but only for utterance-final position; languages that employ devoicing rules in word- or syllable-final positions are no longer responding only to physical considerations."

While notions such as word-final devoicing have a place in descriptive phonology, the utterance-based proposals of Keating, Vennemann, Hooper, and Hyman explain how such processes come about.[1] Word finality has no clear phonetic correlate and thus provides no phonetic motivation for devoicing or neutralization. Utterance Finality, by contrast, does provide such a motivation. At the same time, words are potentially minimal utterances and most utterances end in a full word. Utterance Finality and word finality, therefore, may coincide, and this coincidence can make extension from utterance-final to word-final position possible.

In most cases, this perspective is only an explanation in principle, but there is at least one case that clearly establishes the extension of an originally utterance-final development to word-final position, and beyond. In two important papers, which unfortunately have not received the visibility they deserve, Becker (1977, 1979) draws on dialectal data to argue that the well-known phenomenon of Serbian-Croatian (or "Bosnian-Croatian-Montenegrin-Serbian") accent retraction originated in utterance-final position. The geographically most peripheral, also otherwise conservative varieties (dialects of Čakavian) show no accent shift at all (1a); neighboring Čakavian dialects exhibit the change only utterance-finally, as a shift of high pitch from the final to the penultimate mora (1b); in the standard (Štokavian) language, by contrast, the change has been extended to general word-final position (1c) and has been further extended as generalized accent retraction (1d–e). (Interestingly, as (1e) shows, accent retraction may lead to new contour tones on short vowels, with ` indicating a rising contour and ̏ a falling contour.)[2]

(1) Serbian-Croatian accent retraction (Becker 1977, 1979)

a.	Čakavian 1	*krãl^{y}*	=	[kraál^{y}]		'king'
b.	Čakavian 2	*krâl^{y}*	=	[kráal^{y}]	/ ___ ##	
	vs.	*krãl^{y}*	=	[kraál^{y}]	elsewhere	
c.	Štokavian	*krâl^{y}*	=	[kráal^{y}]	/ ___ #	
d.	Štokavian	*lopàta*	>	*lòpata*		'shovel'
e.	Štokavian	*vodá*	>	*vòda*		'water' (N sg.)
	vs.	*vódu*	>	*vȍdu*		'water' (A sg.)

2.2 Sentence finality vs. utterance finality and the relation between syntax and phrasal prosody

Not only does utterance finality generally coincide with the end of a full word, in many cases it also coincides with the end of a full sentence, and so it may be tempting to define prosodic finality in terms of sentence boundaries. Doing so, however, ignores the rich literature, initiated by Selkirk (1984) and Nespor & Vogel (1986), showing that there is no one-to-one correlation between syntax and phrasal prosody.

Especially interesting is the finding of Vogel (1986) that sandhi phenomena like English *r*-linking (e.g., *the idea[r] is*) can apply across clause boundaries, given the right prosodic phrasing and discourse conditions. Vogel concludes:

---

[1] See also Hock (1986/1991: 80), Crowley (1992: 55), and Trask (1996: 60).

[2] Here and elsewhere # indicates word boundary, ## utterance boundary.

> In prosodic terms, such sandhi rules apply within the phonological utterance, the largest domain in the phonological hierarchy, and the one that may include more than one sentence, depending on certain aspects of the discourse structure. (Vogel 1986: 63)

Further, Hock & Dutta (2010, 2013) provide experimental evidence that English utterance-final vocatives generally are prosodically incorporated into the preceding structure, in spite of the fact that vocatives are syntactically separate from that structure.

3 Utterance Finality, prosodic weakness of verbs, and SOV

This section addresses the prosodic effects of Utterance Finality and its interaction in verb-final languages with the fact that verbs tend to receive less prominence in context than nouns.

3.1 The prosodic weakness of utterance-final position and its common effects

The initial step of example (1) above has parallels in many other languages. As early as 1917, Bloomfield noted that in Tagalog, "… an accent on the last syllable of a sentence often entirely loses its pitch-rise." Cheng & Kisseberth (1979: 34–35) posit a rule of Phrase-Final Lowering for Makua and justify it as "an expected accentual phenomenon – lowering of pitch at the end of an utterance." More far-reaching developments are found in Huichol, where utterance-final constituents lose their underlying tones and exhibit only the pitch properties of the sentence intonation (Grimes 1959); see (2) and (3), where the a-parts illustrate the finality effect, while the b-parts give corresponding forms with their normal tones (accents mark tone; numerals, final pitch contour).

(2) Huichol (Grimes 1959)
- a. *yaawi+kámʌ́+**maa^{3}na^{l}**#*
  'Look! There's a coyote.'
- b. *hutăa+ríeka+tá **mána**+ pairéiku+tua^{3}ni^{1}!*
  'She hauled him back there a second time.'

(3) Huichol (Grimes 1959)
- a. ***yaa^{4}wi^{1}***
  'A coyote!' (uttered with surprise)
- b. ***yaawi**+kámʌ́+maa^{3}na^{l}#*
  'Look! There's a coyote.'

A plausible explanation for these and similar changes[3] is that in unmarked, declarative utterances, the final position has the lowest pitch of the intonational curve; see Figure 1. In fact, Pike (1948: 28) observed that utterance-final pitch is often realized at a much lower level than otherwise expected. See also Hyman (1977). Liberman (1975), and following him, Pierrehumbert (1980) and Beckman & Pierrehumbert (1986) have captured that insight by positing a L(ow) boundary tone at the right edge of Intonational Phrases (IPs), i.e. of utterances.

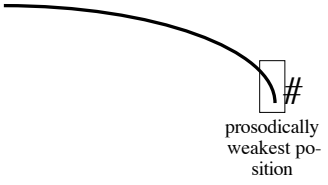

Figure 1. Falling intonation and utterance-final prosodic weakness

Utterance-final position, thus, is the prosodically weakest position, with the smallest acoustic space

[3] See Hock (1999) for other examples.

to carry prominence or to make tonal distinctions. Many tone languages deal with this situation by merely reducing the acoustic space between the contrasting tones; see Shen (1990) and Herman (1996) among others. Other languages evidently react by loss of underlying tone or accent distinctions (e.g., Huichol) or by retraction of prominence to the left (e.g., Makua and Serbo-Croatian). As argued in Hock (1999), the conflict between utterance-final prosodic weakness and prosodic prominence can also account for Hyman's (1977) observation that crosslinguistically word-initial and word-final prominence are quite common, prominence on the penult is also common, but prominence on the "pen-ant" (the second syllable) is quite rare: Penult accent can be accounted for as the result of accent retraction from word-final position, originating in utterance-final position, in response to its prosodic weakness.

3.2 Utterance Finality and its interaction with the prosodic weakness of verbs in SOV languages

In verb-final languages the prosodic effects of Utterance Finality intersect with an independent, crosslinguistic tendency of verbs to receive relatively weak prominence within larger prosodic domains; see already Mathesius (1911) and see also the syntactically oriented account of Kratzer & Selkirk (2007). Not surprisingly, therefore, Ladd (1996) states that final verbs in SOV languages crosslinguistically tend to have reduced prominence or lose their prominence. Let us refer to this phenomenon as Verb Finality.

There are, however, potential counterexamples to this generalization that deserve discussion, one of which is mentioned by Ladd and comes from Bangla (Bengali); the other comes from German. For Bangla, the data in Hayes & Lahiri (1991) suggest major prominence on the final verb. Ladd therefore considers the language an exception to the crosslinguistic tendency of Verb Finality.

However, an experimental study by Hock & Dutta (2006) shows that Bangla does in fact conform to the crosslinguistic tendency. Consider for instance the utterance *Maya Malar bari ḍhuke jae* 'Maya enters Mala's house' in Figure 2, where prominence trails off on the verb *ḍhuke jae*, and where the final syllable of the verb is accompanied by creaky voice.

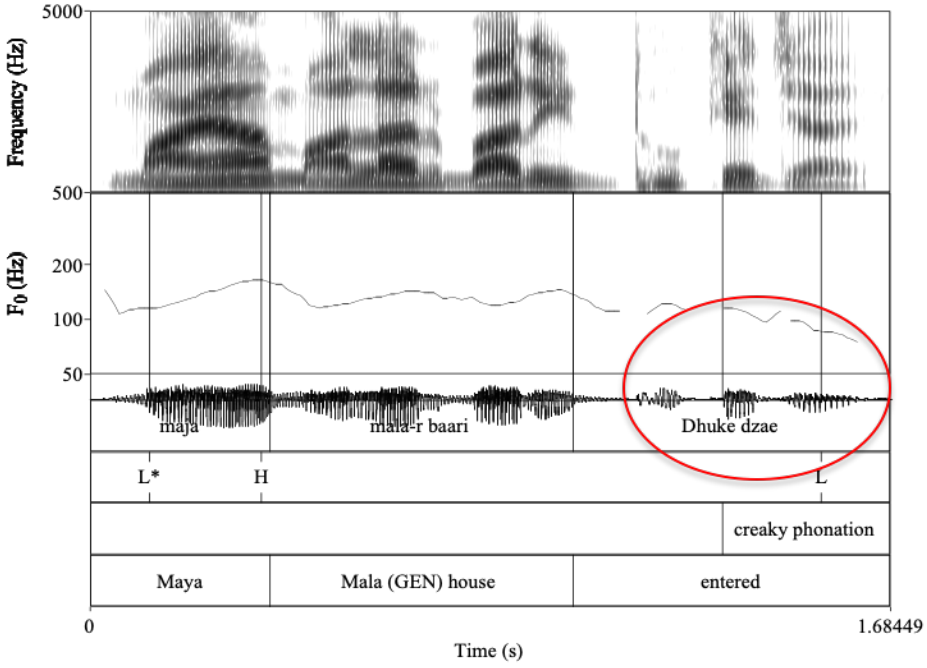

Figure 2. Utterance-finality effect on final verb in Bangla

In one instance, a participant in the experiment produced the sentence *Šêmoli uthonṭa dhobe* 'Shyamoli will clean the courtyard' with the final prominence predicted by Hayes & Lahiri (1991);

Figure 3. But when the utterance was played back with the request to rate whether it sounded acceptable, (s)he said "no" and when asked to speak the sentence in a more natural fashion produced the version in Figure 4, thus confirming that final prominence is marked.

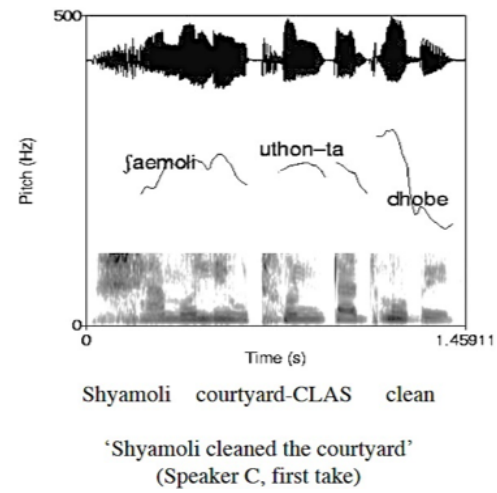

Figure 3. Utterance with final prominence

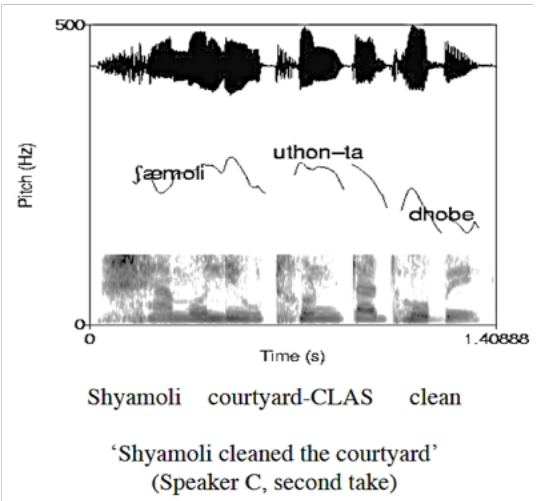

Figure 4. Utterance with self-correction

A possible explanation of the final prominence in Hayes and Lahiri's (1991) data may be that it is an artifact of their experiment, in which every phonological phrase was placed under focus, leading to a highly marked intonational pattern.

A second possible problem was suggested by Manfred Krifka (p.c. 2004), who argued that in German sentences such as (4), the final participle *gesehen* 'seen' carries prominence (represented with the acute accent), rather than the preceding definite object, *den Mann* 'the man'; see (4a). This judgment contrasts with mine (and that of other speakers of German) who would normally have highest prominence on the object (4b). See also Kratzer & Selkirk (2007).[4] Interestingly, Krifka and I agree on placing highest prominence on the object if it is indefinite (4c). The solution seems to be that in Krifka's variety of German, definiteness is interpreted as indicating that the object is given information, and hence the following participle, providing new information, receives prominence. On the other hand, an indefinite object does not convey old information and hence receives prominence. By contrast, my variety does not make this distinction; rather, the subject is construed as old information, and the entire rest of the utterance as new information. Note however that, given the right pragmatic context, I can produce a version in which *gesehen* is treated as new information; see 4(d). But as the example shows, in that case I have prominence on both *den Mann* and on *gesehen*, and the latter accent is downstepped (represented with the "!" diacritic).

(4) German

a.	*Er*	*hat*	*den*	*Mann*	*geséhen*	(Krifka)
b.	*Er*	*hat*	*den*	*Mánn*	*gesehen*	(Hock)
c.	*Er*	*hat*	*einen*	*Mánn*	*gesehen*	(Krifka and Hock)
d.	*Er*	*hat*	*den*	*Mánn*	*!geséhen*	(Hock alternative)
	he	has	the	man	seen	

'He has seen the man.'

[4] Kratzer and Selkirk's intuitions, however, differ from mine in structures such as (i), with intransitive "unaccusatives or eventive unergatives" in "all-new" utterances. Kratzer and Selkirk consider a. to display the correct accentuation; but I find a. to be marked, with *Bayern* carrying focus accent, and only b. strikes me as a natural out-of-the-blue utterance.

(i) a. *Ich hab' gerade im Radio gehört, dass der König von Báyern ertrunken ist*
b. *Ich hab' gerade im Radio gehört, dass der König von Bayern ertrúnken ist*
'I have just heard on the radio that the king of Bavaria has drowned.'

A possible case similar to (4) may be found in the Bangla data of Dutta & Hock (2006).[5] In Figure 5, the final verb *dao* 'give.IMP' has higher prominence than expected, especially considering that the initial consonant is voiced and hence would be expected to lower F_0. Since the preceding pronoun is marked as definite (by the case marker *-ke*), it is possible to interpret this utterance as an example of prominence both on the pronoun and on the verb, but with downstep on the latter.

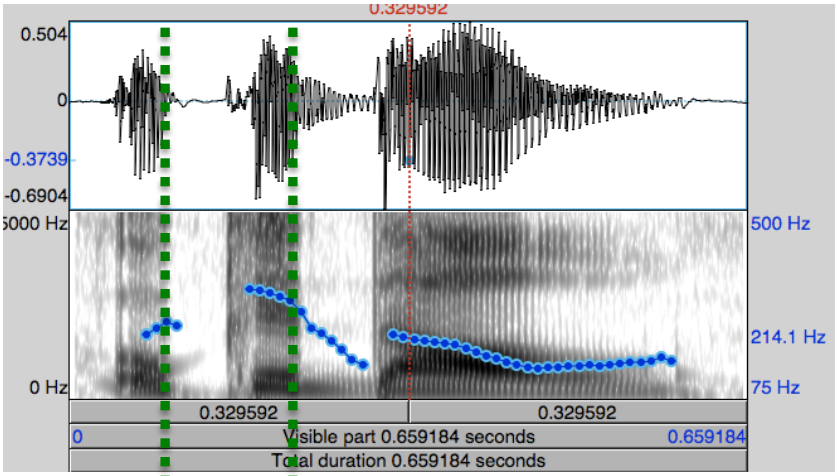

Figure 5. Final downstep

It thus seems safe at this point to conclude that, with some minor, pragmatic or discourse-conditioned exceptions, Verb Finality holds crosslinguistically for SOV languages.

3.3 (Un)markedness and similar issues

The preceding conclusion still requires modification. First, as mentioned earlier, the highly reduced acoustic space responsible for the Utterance-Finality effect holds true only for unmarked, declarative statements. Yes-no questions, which crosslinguistically favor a rising intonation, do not have the same characteristics. If, then, phenomena such as loss or accent retraction also occur in questions,[6] that must be attributed to extension from unmarked, declarative statements.

Further, the accent reduction or loss on finite verbs in SOV languages seems to be generally limited to transitive structures with overt complements; intransitives usually do not seem to synchronically show the same effects. Consider the Hindi example in (5), where (5b) is acceptable only with strong contrastive accent on *maiṁ* 'I'. If, then, historical changes affect all verbs in an SOV language, irrespective of whether they are transitive or not, that must be attributed to extension from transitive to intransitive constructions.

(5) Hindi

a.	*maiṁ*	*gáyā*
b.	*?maíṁ*	*gayā*
	I	went
	'I went.'	

[5] I am grateful to Aditi Lahiri for pointing out this problem at the 2nd Workshop on Tone and Intonation: Models, Computation and Evaluation, February 2013, at EFLU, Hyderabad. The following is based on discussion between Indranil Dutta and me following the Workshop.

[6] An example would be the following from Modern Persian which has the accent retraction discussed in Section 3.4 below, even though it does not have falling intonation.

hanúz	*harekat*	*ná-kard-i?*
still	leave	NEG-do.PST-2SG

'Haven't you left yet?'

3.4 Some diachronic effects of Verb Finality in SOV languages

The hypothesis that Verb Finality can lead to special developments targeting verbs in SOV languages turns out to be highly productive, in that it provides a straightforward sound-change analysis for what otherwise would have to be an analysis in terms of the problematic notion of grammatically conditioned sound change.

On the segmental level, Harms (1964, 1990) argued that the apocope of *-e* in Finnish finite verbs, but not in other morphological categories, can be attributed to the fact that the original Finnish word order was SOV. Similarly, in a variety of early Indo-European languages, final *-i* underwent more extensive or earlier apocope in finite verbs than in other morphological categories; see Hock (2012). Thus, in Latin nouns, apocope takes place only if the *-i* is preceded by more than one syllable (6a), while in verbs there is no such restriction (6b). In both Finnish and early Indo-European, it might be tempting to claim that the special developments, targeting finite verbs, are simply grammatically conditioned; but what would be the grammatical motivations that single out finite verbs for special treatment? A prosodic account, taking note of the fact that these are (or have been) SOV languages can provide a principled prosodic explanation in terms of the Verb Finality effect.

(6) Latin

a.	**animali*	>	*animal*	'animated'
vs.	**mari*	>	*mare*	'sea'
	**pedi*	>	*pede*	'by foot'
b.	**esti*	>	*est*	'is'
	**eyti*	>	*it*	'goes'
	**weniti*	>	*uenit*	'comes'

On the suprasegmental level, Modern Persian offers an interesting case of accent retraction on finite verbs, whose original prosodic motivation is still discernible.

As is well known, the default accent of Modern Persian is word-final. Final finite verbs of main clauses, however, retract the accent as far to the left as possible, even onto the preceding complement; see (7a–c); the only complication is that negation may block further retraction (7d). Finite verbs of preposed dependent clauses, by contrast, do not retract the accent (7e). (The accent on main-clause *níst* is conditioned by the negation *n(á)*.) So far, everything looks like it is completely prosodically conditioned: utterance-final verbs undergo accent retraction (with further retraction blocked by negation), verbs in other contexts do not. Example (7f) may at first sight conform to this generalization, in that the postposed dependent clause, whose verb is utterance-final, shows accent retraction. Note, however, that the verb of the preceding main clause also shows retraction, even though it is not utterance-final. Evidently, some kind of grammaticalization has taken place, with accent retraction generalized in main clauses, no matter whether they are utterance-final or not. For further details see Lazard (1957, 1989) and Windfuhr (1987).

(7) Persian

a.	*ráft-am*			'I went'
b.	*bé-rav-am*			'I would go'
c.	*kā́r mi-kon-am*			'I always do the work'
d.	*ná-raft-am*			'I did not go'
e.	[*àgar*	*be-rav-ád*]$_{DC}$		'If he goes, there will be no-one left.'
	if	go.3SG.SBJV		
	[*kàs-i*	*digàr*	*níst*]$_{MC}$	
	anybody	left.over	NEG.be.3SG.PRS	
f.	[*hàqq-aš*	*ín*	*ast*]$_{MC}$	'The truth of it is that I do not have money.'
	truth-of.it	this	be.3SG.PRS	
	[*ke*	*púl*	*nà-dār-am*]$_{DC}$	
	that	money	NEG-have.PRS-1SG	

Again, only a prosodic account in terms of Utterance Finality and Verb Finality is able to account for the different behavior of finite verbs in main clauses and *ke*-clauses vs. preposed dependent

clauses. At the same time, as in all other cases, we have to accept extensions that go beyond the original triggering context — in this case, extension of accent retraction to main clauses with final finite verb,[7] no matter whether they follow or precede dependent clauses.

4 Synchronic consequences of Verb Finality in Hindi

In this section I address two Hindi phenomena for which Utterance- and Verb-Finality prosodic accounts offer attractive explanations. These are the deletability of the verb 'be' and the placement of negation.

4.1 Hindi 'be'-deletion and Utterance Finality

The verb 'be' is frequently deleted in Hindi negative sentences. The common wisdom on 'be'-deletion is that it is an optional process, with the possible exception of existential 'be', which may be stable.[8]

However, examples such as (8) vs. (9) suggest that 'be'-deletion is not entirely optional, in that 'be' can be deleted only if the preceding element is prosodically non-prominent, and not if the preceding element is prominent. Put differently, deletion is possible only if the output leaves behind an utterance-final prosodic slope or declination; this Slope principle requires that the final lexical element have lower prominence than the penultimate one, as shown in (8). If deletion violates the Slope principle, as in (9), then it is blocked.

(8) Hindi

a.	*vah kitāb*	*nahĩṁ*	*acchī*	*hai*
	that book	not	good	is
	'That book is not good.'			
b.	*vah kitāb*	*nahĩṁ*	*acchī*	Ø

(9) Hindi

a.	*vah kitāb*	*acchī*	*nahĩṁ*	*hai*
	that book	good	not	is
	'That book is not good.'			
b.	*vah kitāb*	*acchī*	*nahĩṁ*	*?Ø

Further support for the Slope principle comes from utterances like (10) in which *bilkul* 'really' is

[7] If main-clause verbs are not clause-final, accent retraction apparently does not occur, as in colloquial *mi-ra-vám tehrún* 'I am going to Tehran'.

[8] The question whether existential 'be' is in fact stable deserves further investigation. In utterances such as (i–ii) below, the fact that 'be' cannot be deleted may simply be a consequence of the fact that the only thing that can precede is the negation, and therefore (i) follows the same pattern as (8).

i.	*yahāṁ par*	*acchī kitābeṁ*	*nahĩṁ*	*haiṁ*	
	here	good books	not	are	
	'There are no good books here.'				
ii.	*yahāṁ par*	*acchī kitābeṁ*	*nahĩṁ*	*Ø	

A decision whether this is the correct analysis, or whether existential 'be' is really stable, would depend on the grammaticality of (iv), in which *bilkul* 'really' is prominent, and *nahīṁ* 'not' is not. Example (10) might suggest that it is, but 'be' here is not the existential verb. This issue clearly deserves further testing.

iii.	*yahāṁ par*	*acchī kitābeṁ*	*bílkul*	*nahīṁ*	*haiṁ*	
	here	good books	really	not	are	
'There really aren't any good books here.'						
iv.	*yahāṁ par*	*acchī kitābeṁ*	*bílkul*	*nahīṁ*	??Ø	

prominent, and not the negation (and where an additional element, *pasand* 'liked', follows *nahīṁ* 'not'). As (10b) shows, because of the prominence on *bilkul*, the following *nahīṁ pasand* forms a sufficient downward slope such that 'be'-deletion is permitted.

(10) Hindi

a.	*yah*	*mujhe*	*bílkul*	*nahīṁ*	*pasand*	*hai*
	this	to me	really	not	liked	is
	'I really don't like this.'					
b.	*yah*	*mujhe*	*bílkul*	*nahīṁ*	*pasand*	Ø[9]

Examples like these suggest that the apparent optionality of 'be'-deletion is in fact governed by the prosodic principle of Utterance Finality: Deletion is permitted only if what remains has a prosodic Slope and thus avoids prosodic prominence in utterance-final position.

4.2 The order of negation and finite verb in Hindi

While in the preceding case a prosodic surface filter could possibly be invoked (because 'be'-deletion may be considered a "surfacey" phenomenon, close to the syntax/phonology interface), the situation is different for the relative order of negation and finite verb in expressions of the type (11) and (12). For, unlike 'be'-deletion, the different order in (11) vs. (12) would, in current transformational approaches, have to be accounted for by syntactic movement, presumably to different left-peripheral landing sites, whatever these sites may be.

(11) Hindi

a.	*vah*	*kām*	*nahī́ṁ*	*kar-t-ā*
	he	work	not	do-IPFV-MSG
	'He doesn't do the work.'			
b.	*?*vah*	*kām*	*nahīṁ*	*kár-t-ā*
	'He doesn't do the work.'			

(12) Hindi

a.	*?*vah*	*kām*	*kar-t-ā*	*nahī́ṁ*
	he	work	do-IPFV-MSG	not
	'He doesn't do the work.'			
b.	*vah*	*kām*	*kár-t-ā*	*nahīṁ*
	'He doesn't do the work.'			

Significantly, however, the permissible (or felicitous) orders – (11a) and (12b) – are precisely those which conform to the prosodic Slope principle and thus conform to the expectations of Utterance Finality; the unacceptable orders violate the principle.

If we wanted to account for this fact syntactically, we would have to generate both surface orders without any restrictions through some kind of movement (which would violate the principle that movement must be conditioned), with a (prosodic) surface filter weeding out structures that violate the Slope principle. However, as Maling & Zaenen (1981) observe, surface filters are difficult to constrain and therefore are in principle excessively powerful. Moreover, even if a filter account were to be accepted, it would support the claim that prosody plays a larger role than commonly assumed in what ordinarily is considered Hindi syntax.

An alternative would be a purely prosodic account which directly correlates prominence and final Slope – prosodic phenomena par excellence – with word order, presumably in the syntax/prosody interface.

Until recently I myself expressed a fair amount of qualms about this proposal. My concern was that such a direct correlation of prosody and word order comes at a price, for it requires the assumption that at least some aspects of word order are not syntactically, but prosodically determined; put differently, that there is (or can be) prosodic movement. True, in Hock (1996) I had offered evidence

[9] This is an utterance that my wife used in exasperation when our dog was misbehaving.

and arguments that second-position (P2) ordering of clitics can – or even must – be attributed to prosodic factors. But neither the finite verb nor the negation in (11) and (12) are clitics. Attributing the word order preferences in (11) and (12) directly to prosodic considerations may therefore not be any better than invoking a prosodic filter; for it may be argued that like the prosodic filter approach, it is difficult or even impossible to constrain. Most important, we do not (as yet) have a well-developed theory of phrasal prosody and prosodic movement.

Recent publications suggest that these qualms may be excessive, and that there is increasing evidence for prosodic movement, at least as confined to the edges of prosodic domains, including utterance-initial and utterance-final position (the domains of P2 and of Utterance or Verb Finality). See e.g. Agbayani & Golston (2010), Agbayani, Golston, & Hederer (2011).

5 South Asian developments that can be attributed to Verb Finality

This section presents two phenomena in South Asian languages that are explainable in terms of Verb Finality, one (in Vedic) that involves accentuation, a second, more speculative one (in Kashmiri) involving word order. In both cases the Modern Persian accent retraction (section 3.4) provides a likely parallel.

5.1 Vedic verb accentuation

In Vedic Sanskrit, finite verbs are unaccented in main clauses (unless initial in the clause or the poetic line[10]), but accented in dependent structures (13).

(13) Vedic Sanskrit

[*tásmai*	*víśaḥ*	*svayám*	*evā́*	*namante*]MC
CP.DAT.MSG	people.NOM.FPL	self	PCLE	bow.PRS.3PL
[*yásmin*	*brahmā́*	*pū́rva*	*éti*]DC	
RP.LOC.MSG	brahmin.NOM.MSG	first.NOM.MSG	go.PRS.3SG	

'Even the common people bow to him for whom the brahmin goes first.' (RV 4.50.8)

A comprehensive, prosodically motivated account for this difference was first proposed by Klein (1992) with reference to Hock (1986/1991). According to Klein, non-accentuation of MC verbs arose in unmarked, or canonical constructions with the MC verb in sentence-final position, and with resolution of the conflict between the high pitch of the verb accent and the low pitch of sentence-final falling intonation through loss of verb accent.

Klein's account is further developed by Hock (2014), who argues that the main-clause accent loss results from an earlier stage of accent retraction in canonical utterance-final position and that this retraction also accounts for the well-known accent retraction of Greek finite verbs. The major ingredients of the account for Sanskrit are as follows: Given canonical SOV word order (14i) and canonical clause order DC + MC (14ii), the verb of the main clause occurs in utterance-final position, while the DC verb does not (14iii). It is the MC verb, therefore, that is subject to the Verb Finality effect and (via accent retraction) loses its prosodic prominence (14iv). Up to this point, the developments are remarkably similar to those of Modern Persian. In both languages the Verb Finality effect applies in MCs, no matter whether utterance-final or not. The developments differ in their treatment of DCs. Modern Persian postposed complement clauses have accent retraction, but preposed DCs do not. In the prehistory of Vedic, grammaticalization extends lack of accent retraction to all DCs, whether postposed or preposed, and thus leads to a complete polarization of main and dependent clauses; see (14v). A parallel for this polarization can be found in German, Dutch, and Frisian, where V2 gets generalized in main clauses, verb-final order in dependent clauses. In both cases, the polarization introduces a secondary feature that further distinguishes dependent clauses from main clauses, beyond the complementizer or relative pronominal.

[10] Other exceptions, motivated by poetic prosody or by discourse considerations, are discussed in Klein (1992), Hock (1996).

(14) i.	Original canonical word order:	S O	V #		
ii.	Original canonical clause order:	RC	MC		
iii.	Hence:	[S O V]$_{DC}$		[S O V]$_{MC}$ ##	
iv.	Finality-conditioned accent loss:	[S O V]$_{DC}$		[S O V]$_{MC}$ ##	
				[-accent]	
v.	Polarization:	DC	:	MC	
		[verb +accent]		[verb -accent]	

As in other cases, Utterance and Verb Finality (plus further extensions) provides a prosodically motivated explanation for Vedic non-accentuation of verbs in main clauses, while notions such as grammatical conditioning fail to provide a meaningful account.

5.2 Kashmiri word order

As is well known, Kashmiri resembles German, Dutch, and Frisian by exhibiting an innovated second-position order of finite verbs in main clauses. There are some interesting differences such as the placement of interrogatives into pre-finite-verb position in Kashmiri, but not in German and related languages; see e.g. Bhatt (1999). These details, however, need not concern us here.

A more remarkable difference is that Kashmiri, unlike its European counterparts, does not retain verb-final order in all dependent clauses, but only in relative clauses; complement *ki/zi*-clauses, by contrast, have the same V2 order as main clauses. See the examples in (15) vs. (16).

(15) Kashmiri (adapted from Koul 2003: 918–919)

[*yɔs*	*kūr*	*tse*	*pasand*	***chay***]$_{RC}$	
RP	girl	you.DAT	pleasing	be.PRS.DAT.2SG	
[*sɔ*	*kūr*	***cha***	*me*	*ti*	*pasand*]$_{MC}$
CP	girl	be.PRS.DAT.1SG	I.DAT	also	pleasing

'The girl that you like I like too.'

(16) Kashmiri (adapted from Koul 2003: 918–919)

[*me*	***chu***	*patā*]$_{MC}$		
I.DAT	be.PRS.DAT.1MSG	known		
[*ki/zi*	*təm'*	***h'ot***	*nov kōṭh*	*bāzri*]$_{DC}$
that	he.AG	buy.PST.3SG	new coat	bazaar

'I know that he bought a new coat in the bazaar.'

While the V2 in the complement clause can be syntactically accounted for by assuming that *ki/zi*-structures are obligatorily double-Comp constructions, the obligatoriness of the phenomenon in *ki/zi*-structures and its obligatory absence in relative clauses remain unexplained; but given its assumptions and constraints that may be the best synchronic analysis that current transformational syntax can offer.

From the historical perspective, an alternative – even if speculative – prosodic analysis is possible. As a starting point, note the parallelism between Kashmiri verb position in (15) and (16) and Modern Persian verb accentuation in (7e) and (7f), repeated as (17a) and (17b), respectively, in modified form.[11]

(17) Persian

a.

[*àgar*	***be-rav-ád***]$_{DC}$	
if	go.3SG.SBJV	
[*kàs-i*	*digàr*	***níst***]$_{MC}$
anybody	left.over	NEG.be.3SG.PRS

'If he goes, there will be no-one left.'

[11] Note that Kashmiri relative clauses are canonically left-peripheral and follow their main clauses much less frequently than in other Indo-Aryan languages (Rakesh Bhatt, p.c. 2009).

b. [*hàqq-aš* *ín* ***ast***]$_{MC}$
truth-of.it this be.3SG.PRS
[*ke* *púl* ***nà-dār-am***]$_{DC}$
that money NEG-have.PRS-1SG
'The truth of it is that I do not have money.'

If we assume that at an earlier stage Kashmiri had a similar pattern as Modern Persian, with accent retraction on final verbs in main clauses and right-peripheral complement clauses, but not in left-peripheral dependent clauses, the following (speculative) account becomes possible.

As in Germanic (Hock 1982) the change to V2 may have been initiated by the movement of clitic 'be' (*ch-* < Apabhr. *acchaï*), followed by other light verbs. Now, as Modern Persian examples like (7c) and the pre-Vedic developments in (14) show, accent retraction on utterance-final verbs may lead to complete accent loss. A similar development in Kashmiri would have led to utterance-final verbs becoming unaccented and hence, in effect, prosodically light verbs, thus being included in the shift to V2. This would account for the change to V2 not only in main clauses (where it was subsequently generalized to all occurrences) but also in right-peripheral complement clauses. Left-peripheral relative clauses, by contrast, would fail to undergo the change. See the summary in (18).

(18) i.	Original canonical word order:	S O V #	
ii.	Original canonical clause orders:	DC	MC
		MC	*ki/zi*-clause
iii.	Hence:	[S O V]$_{RC}$	[S O V]$_{MC}$ ##
		[S O V]$_{MC}$	[S O V]$_{ki/zi}$ ##
iv.	Finality-conditioned accent loss:	[S O V]$_{RC}$	[S O V]$_{MC}$ ## [-accent]
		[S O V]$_{MC}$	[S O V]$_{ki/zi}$ ## [-accent]
v.	Generalization of [V -accent] to all MCs		
v.	Shift to V2:	[S O V]$_{RC}$	[S V2 O V]$_{MC}$ ## [-accent]
		[S V2 O V]$_{MC}$ [-accent]	[S V2 O V]$_{ki/zi}$ ## [-accent]

If this scenario is on the right track, Kashmiri shares strong similarities with both Modern Persian and Vedic. However, given the chronological gap between Vedic on one hand and Modern Persian and Kashmiri on the other, it is not likely that there is any direct connection between the developments. Rather, it seems that we have to accept developments of this sort as possible independent responses to Verb Finality.

6 Conclusions and implications

As I hope to have demonstrated, Utterance Finality — especially in the form of Verb Finality — plays an important role, both synchronically and in linguistic change, both crosslinguistically and more specifically in South Asian languages. However, there are also questions as to how prosodic effects such as the sensitivity of 'be'-deletion, or the ordering of finite verb and negation in Hindi, can be accounted for in synchronic grammar.

In principle, Chomsky's (1995) Minimalist Program opens ample opportunities for exploring such effects in terms of the interfaces between the syntax and other components of the grammar, including the prosodic one.

With some notable exceptions, however, syntacticians have been reluctant to entertain accounts that pay serious attention to phrasal prosody and its interface with syntax. Consider for instance many of the contributions on P2 clitics in Halpern & Zwicky (1996), such as Hale (1996) on Vedic

and Progovac on Serbo-Croatian.[12] Prosodic approaches, such as Radanović-Kocić (1996) and Hock (1996) regarding P2 clitics, have tended to be side-lined or ignored by syntacticians such as Erschler (2009) regarding Radanović-Kocić, and Hale (1996), Lowe (2014) regarding Hock (1996), in spite of strong empirical evidence showing that P2 can only be accounted for prosodically, at least in Serbo-Croatian and Vedic Sanskrit. Consider Serbo-Croatian (19), which shows that when the first element of the clause is followed by an appositive elaboration, P2 clitics cannot occur after the first word (19b) or after that word together with its appositive extension (19c), but must appear after the first prosodically prominent element that follows the prosodic break after the appositive (19a). Similarly, Vedic Sanskrit (20) shows that P2 elements may occur after the first word that follows a poetically imposed prosodic break (in this case a caesura, indicated by a colon).

(19) Serbo-Croatian

a.	*Ja* |	*tvoja mama* |	*OBEČALA*	***sam***	***ti***	*igračku*
	I	your Mom	promised	AUX.1SG.CLT	you.SG.CLT	toy
	'I, your Mom, have promised you a toy.'					
b.	**JA*	***sam ti*** |	*tvoja mama* |	*obečala*	*igračku*	
c.	**JA* |	*TVOJA MAMA* |	***sam ti***	*obečala*	*igračku*	

(20) Vedic Sanskrit

apā́m	*tokásya*			
water.GEN.FPL	offspring.GEN.NSG			
tánayasya	*jeṣá* |			
offspring.GEN.NSG	winning.LOC.MSG			
índra	*sūrī́n*			
Indra.VOC.MSG	patron.ACC.MPL			
: *KṚṆUHÍ*	***smā***	***no***	*ardhám*	
make.IMPV.2SG	PCLE	our.CLT	part.ACC.MSG	

'In the winning of water for our offspring, Indra, make our patrons (take) part.' (RV 6.44.18cd)

In light of evidence of this sort it is possible to take the strong position that those continuing to argue for purely syntactic approaches to P2 clitic placement do so at their own risk, by ignoring clear empirical evidence or, in the case of Vedic, by shunting it aside under the assumption that poets can treat certain poetic prosodic contexts as if they are clause-initial (e.g. Hale 1996).

Only relatively recently have prosodic accounts found wider recognition, such as Bošković (2001), Agbayani & Goldston (2010), and Agbayani, Goldston, & Hederer (2011). This is an encouraging sign, and it is to be hoped that the work will continue and lead to an increasingly well-developed theory of phrasal prosody.

I hope that the present paper will be a useful contribution in this context by adding to the number of cases for which a prosodic account is required or, at least, a viable alternative to purely syntactic explanations.

References

Agbayani, Brian and Chris Golston. 2010. Phonological movement in Classical Greek. *Language* 86: 133–167.

Agbayani, Brian, Chris Golston, and Dasha Hederer. 2011. Prosodic movement. In Mary Byram Washburn et al., eds. *Proceedings of the 28th West Coast Conference on Formal Linguistics*, pages 231–239. Somerville, MA: Cascadilla Proceedings Project.

Becker, Lee A. 1977. Leftward movement of high tone. LSA Annual Meeting.

[12] The closest that many of these contributions come to recognizing the relevance of prosody to P2 clitic placement is by accepting Halpern's (1992) notion of Prosodic Flip from initial position to the position after the first word.

Becker, Lee A. 1979. A contribution to an explanation of the Neo-Stokavian accent retraction. *Zbornik za filologiju i lingvistiku* 22(1): 87–94.
Beckman, Mary and Janet Pierrehumbert. 1986. Intonational structure in English and Japanese. *Phonology* 3: 255–309.
Bhatt, Rakesh Mohan. 1999. *Verb movement and the syntax of Kashmiri*. Dordrecht: Kluwer.
Bloomfield, Leonard. 1917. *Tagalog texts with grammatical analysis*. (University of Illinois Studies in Language and Literature, vol. 3, no. 2–4.)
Bošković, Željko, 2001. *On the nature of the syntax-phonology interface*. Amsterdam: Elsevier.
Canger, Una. 1990. Philology in America: Nahuatl. In Jacek Fisiak, ed., *Historical linguistics and philology*, pages 107–118. Berlin/New York: Mouton de Gruyter
Cheng, Chin-Chuan, and Charles W. Kisseberth. 1979. Ikorovere Makua tonology (part 1). *Studies in the Linguistic Sciences* 9(1): 31–63.
Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
Crowley, Terry. 1992. *An introduction to historical linguistics*. 2nd ed. Auckland: Oxford University Press.
Erschler, David. 2009. Modularity and 2P clitics: Arguments from Digor Ossetic. *Proceedings of IATL 25* [Israel Association for Theoretical Linguistics]. Online publication: http://linguistics.huji.ac.il/IATL/25/Erschler.pdf
Grimes, J. E. 1959. Huichol tone and intonation. *International Journal of American Linguistics* 25: 221–232.
Hale, Mark R. 1996. Deriving Wackernagel's Law: Prosodic and syntactic factors determining clitic placement in the language of the Rigveda'. In Halpern & Zwicky 1996: 165–197.
Halpern, Aaron. 1992. *Topics in the syntax and placement of clitics*. PhD thesis, Stanford University.
Halpern, Aaron and Arnold Zwicky, eds. 1996. *Approaching second: Second position clitics and related phenomena*. Stanford: CSLI Publications.
Harms, Robert T. 1964. Review of L. Hakulinen, *The structure and development of the Finnish language*. *Word* 20:105–114.
Harms, Robert T. 1990. Synchronic rules and diachronic "laws": The Saussurean dichotomy reaffirmed. In Edgar C. Polomé, ed., *Research guide on language change*, pages 313–323. Berlin: Mouton de Gruyter.
Hayes, Bruce and Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9: 47–96.
Hock, Hans Henrich. 1982. AUX-cliticization as a motivation for word order change. *Studies in the Linguistic Sciences* 12(1): 91–101,
Hock, Hans Henrich. 1986/1991. *Principles of historical linguistics.* 1st ed. 1986, 2nd rev. ed. 1991. Berlin/New York: Mouton de Gruyter.
Hock, Hans Henrich. 1996. Who's on first? Toward a prosodic account of P2 clitics. In Halpern & Zwicky 1996: 199–270.
Hock, Hans Henrich. 1999. Finality, prosody, and change. In Osamu Fujimura, Brian D. Joseph, and Bohumil Palek, eds., *Proceedings of LP'98*, pages 15–30. Prague: The Karolinum Press
Hock, Hans Henrich. 2012. Phrasal prosody and the Indo-European verb. In H. Craig Melchert, ed., *The Indo-European verb: Proceedings of the Conference of the Society for Indo-European Studies, Los Angeles, 13–15 September 2010*, pages 115–126. Wiesbaden: Reichert.
Hock, Hans Henrich. 2014. Vedic verb accent revisited. In Jared Klein and Elizabeth Tucker, eds., *Vedic and Sanskrit historical linguistics: Papers from the 13th World Sanskrit Conference*, pages 153–178. Delhi: Motilal Banarsidass.
Hock, Hans Henrich. 2015. Yamuna Kachru's contributions to South Asian discourse and convergence studies. *World Englishes* 34(1): 68–77.
Hock, Hans Henrich and Indranil Dutta. 2006. Interaction of verb accentuation and utterance finality in Bangla. In Rüdiger Hoffmann and Hansjörg Mixdorff, eds., *CDROM Proceedings of Speech Prosody: 3rd International Conference, Dresden, May 2–5, 2006*. (Studientexte zur Sprachkommunikation, 40.) Dresden: TUDpress: 2006
Hock, Hans Henrich and Indranil Dutta. 2010. Prosody vs. syntax: Prosodic rebracketing of final vocatives in English. In: Mark Hasegawa-Johnson, ed. *Proceedings of Speech Prosody 2010, Chicago*. CD-publication (see also http://www.lulu.com/us/en/shop/mark-hasegawa-johnson/proceedings-of-speech-prosody-2010/paperback/product-11603457.html)

Hock, Hans Henrich and Indranil Dutta. 2013. Prosodic incorporation of English utterance final vocatives. *The EFL Journal* 4(2): 1–20.
Hooper, Joan B. 1976. *An introduction to natural generative phonology*. New York: Academic Press.
Hyman, Larry M. 1977. On the nature of linguistic stress. In Larry M. Hyman, ed., *Studies in stress and acc*ent, pages 37–82. (Southern California Occasional Papers in Linguistics, 4.) Los Angeles: University of Southern California.
Keating, Patricia A. 1988. The phonology-phonetics interface. In Frederick J. Newmeyer, ed., *Linguistics: The Cambridge survey, I: Linguistic theory: foundations*, pages 281–302. Cambridge: Cambridge University Press.
Klein, Jared S. 1992. *On verbal accentuation in the Rigveda*. (American Oriental Society Essay Number 11.) New Haven, CT: American Oriental Society.
Koul, Omkar N. 2003. Kashmiri. In George Cardona and Dhanesh Jain, eds., *The Indo-Aryan languages*, pages 895–952. London/New York: Routledge.
Kratzer, Angelika and Elisabeth Selkirk. 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24(2): 93–135
Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
Lazard, Gilbert. 1957. *Grammaire du persan contemporain*. Paris: Klincksieck.
Lazard, Gilbert. 1989. Le persan. In Rüdiger Schmitt, ed., *Compendium linguarum iranicarum*, pages 263–293. Wiesbaden: Reichert.
Liberman, Mark. 1975. *The intonational system of English*. Cambridge, MA: MIT PhD dissertation.
Lowe, John. 2014. Accented clitics in the Ṛgveda. *Transactions of the Philological Society* 112(1): 5–43.
Maling, Joan and Annie Zaenen. 1981. Germanic word order and the format of surface filters. In Frank Heny, ed., *Binding and filtering*, pages 255–278. London: Croom-Helm.
Mathesius, Vilém. 1911. O potenciálnosti jevů jazykovýh. *Věstník Král. české společnosti nauk* (Prague). Translation 'On the potentiality of the phenomena of language' by J. Vachek, 1964. In Josef Vachek, ed., *A Prague School Reader*, pages 1–32. Bloomington/London: Indiana University Press.
Nespor, Marina and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris. Repr. with a new foreword, 2012: Berlin/Boston: de Gruyter Mouton.
Progovac, Ljiljana. 1996. Clitics in Serbian/Croatian: Comp as the second position. In Halpern & Zwicky 1996: 411–428.
Ohala, John. 1993. The phonetics of sound change. *Historical linguistics: Problems and perspectives*, ed. by Charles Jones, 236–278. London/New York: Longman.
Pierrhumbert, Janet. 1980. *The phonetics and phonology of English intonation*. Cambridge, MA: MIT PhD dissertation.
Pike, Kenneth L. 1948. *Tone languages: a technique for determining the number and type of pitch contrasts in a language, with studies in tonemic substitution and fusion*. Ann Arbor: University of Michigan Press.
Radanović-Kocić, Vesna. 1996. The placement of Serbo-Croatian clitics: a prosodic approach. In Halpern & Zwicky 1996: 429–446.
Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure.* Cambridge, MA: MIT Press.
Trask, Robert Lawrence. 1996. *Historical linguistics*. London: Arnold.
Vennemann, Theo. 1974. Words and syllable in natural generative grammar. In Anthony Bruck et al., eds., *Papers from the Parasession on Natural Phonology*, pages 346–374. Chicago: Chicago Linguistic Society.
Vogel, Irene. 1986. External sandhi rules operating between sentences. In Henning Andersen, ed., *Sandhi phenomena in the languages of Europe*, pages 55–64. Berlin: Mouton de Gruyter.
Windfuhr, Gernot. 1987. Persian. In Bernard Comrie, ed., *The world's major languages*, pages 523–546. New York: Oxford University Press.

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

FRANK KÜGLER, *Goethe University Frankfurt*

Received APRIL 2019; Revised MARCH 2020

ABSTRACT

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

1 Introduction

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

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

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

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

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

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

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

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

1.1 Background on Hindi intonation

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

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

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

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

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

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

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

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

2 The production study

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

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

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

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

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

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

2.1 Methods

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

2.1.1 Design and Materials

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

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

2.1.2 Participants

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

2.1.3 Procedure

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

2.1.4 Data pre-processing and statistical analysis

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

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

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

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

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

2.2 Results

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

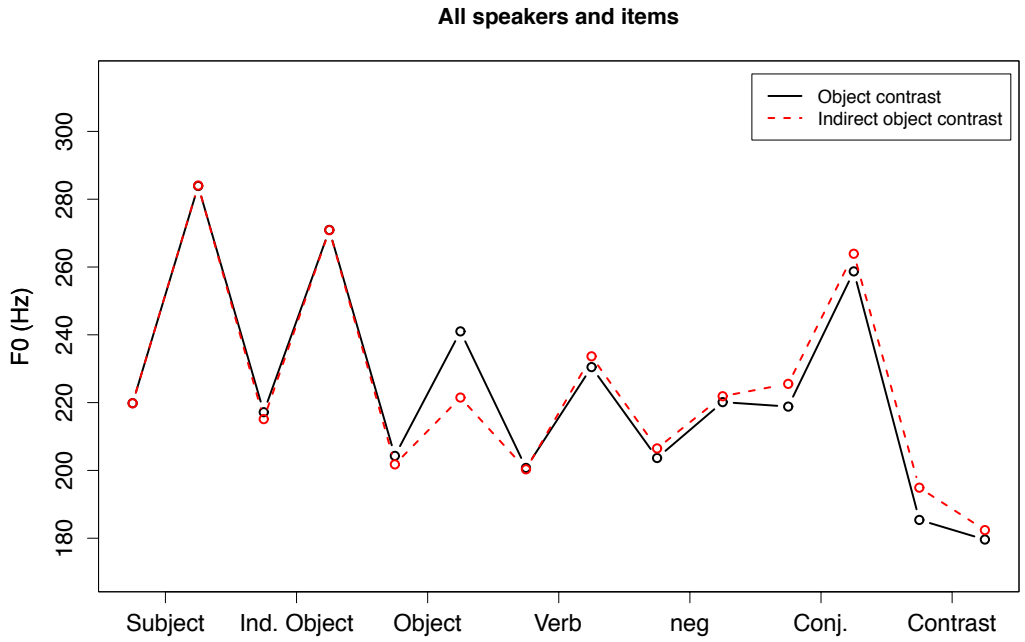

Figure 1. f0 plot of L and H tone per prosodic word averaged across all speakers (n = 30) and items (n = 6). The black solid line represents the object contrast condition, the dashed red line the indirect object contrast condition.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

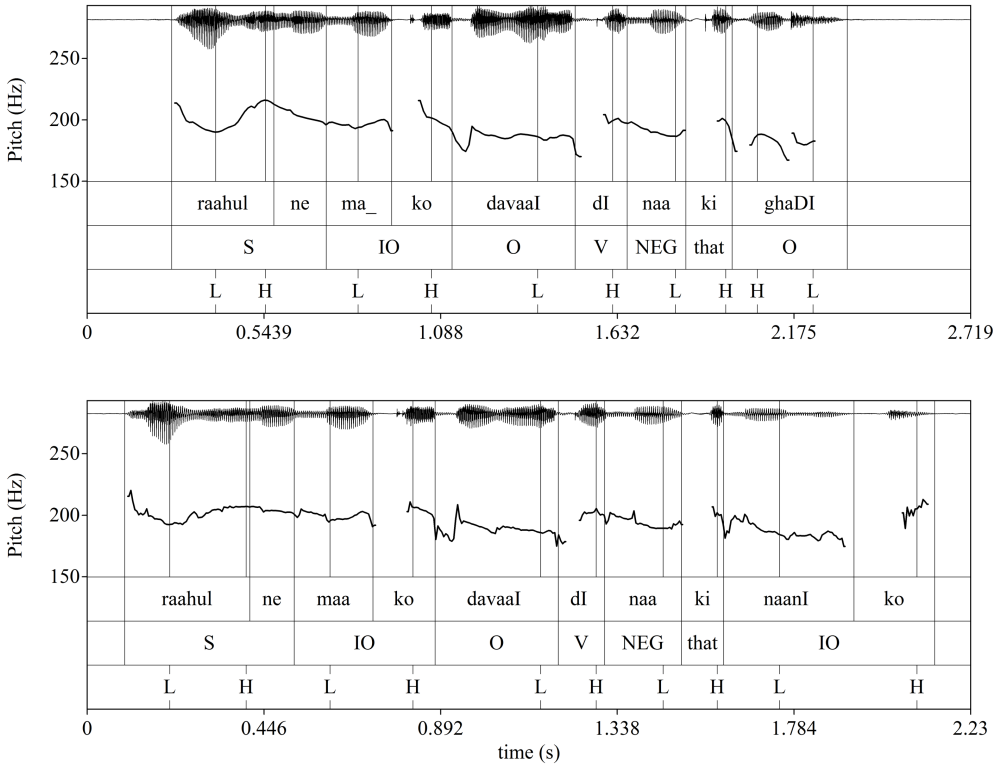

Figure 2. A comparison between conditions for a speaker who does not express PFC on the direct object: the object contrast condition is shown in the upper panel, and the indirect object contrast condition in the lower panel.

2.3 Discussion

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

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

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

and no prosodic cues were needed.

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

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

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

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

3 The completion task

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

3.1 Methods

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

3.1.1 Speech materials

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

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

b. ghaṛī
watch
'the watch.'

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

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

3.1.2 Participants

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

3.1.3 Procedure

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

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

3.1.4 Predictions

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

3.2 Results

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

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

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

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

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

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

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

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

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

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

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

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

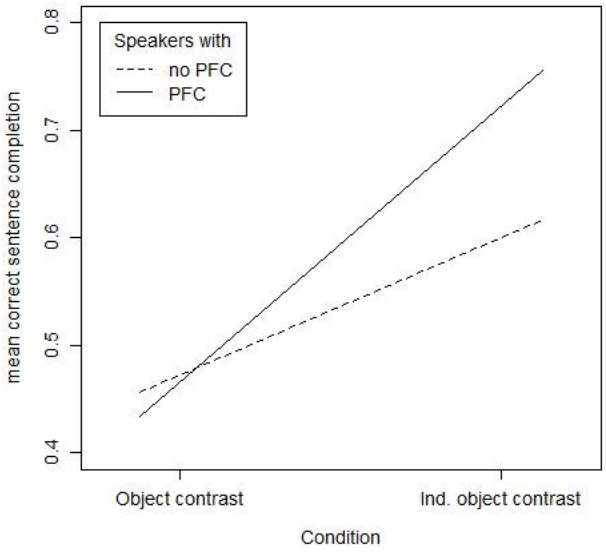

Figure 3. Interaction plot: responses to recordings of speakers without PFC (dotted line) compared to those of speakers with PFC (solid line), split by condition (contrast on DO vs. contrast on IO).

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

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

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

3.3 Discussion

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

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

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

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

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

4 General discussion

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

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

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

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

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

(4)

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

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

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

4.1 On the perception of focus

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

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

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

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

4.2 The role of speaker variation

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

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

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

4.3 The role of cues of intonation in sentence processing

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

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

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

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

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

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

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

5 Conclusion

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

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

Acknowledgements

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Appendix

Speech materials

E03-C01-01-30-079-

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

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

E03-C02-01-30-107-

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

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

E03-C01-02-30-153-

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

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

E03-C02-02-30-214-

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

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

E03-C01-03-30-072-

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

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

E03-C02-03-30-034-

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

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

E03-C01-04-30-197-

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

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

E03-C02-04-30-076-

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

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

E03-C01-05-30-132-

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

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

E03-C02-05-30-012-

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

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

E03-C01-06-30-135-

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

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

E03-C02-06-30-092-

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

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

Prosody of discontinuous nominal phrases in Indian languages

CAROLINE FÉRY, *University of Frankfurt*
GISBERT FANSELOW, *University of Potsdam*

Received DECEMBER 2018; Revised OCTOBER 2019

ABSTRACT

The purpose of this survey is to compare the tonal and prosodic structure of discontinuous nominal phrases in several Indian languages with those of the better studied intonation languages, such as English and German. From a syntactic perspective, the SOV base order and the free constituent order property in nearly all Indo-Aryan, Dravidian, and Tibeto-Burman languages lead to a fairly rich system of discontinuous NPs of the type involving two independently generated NPs. From a prosodic perspective, the Indian languages discussed in the article are prototypical examples of phrase languages, i.e., the intonation is not dependent on variable pitch accent placement, but is rather based on the existence of prosodic domains, each with a characteristic 'phrasal' intonation. We will see that the division between cohesive and non-cohesive patterns that has been made for intonation languages is blurred in these languages. In line with this observation on prosody, the hierarchy-preserving and hierarchy-inverting discontinuous nominal phrases are not always easy to tell apart syntactically.

1 Introduction

More often than not, the study of prosody is confined to base word order without any change in information structure, i.e. declarative sentences with broad focus. This strategy is used to identify and study the basic tonal structure of a language, the position of pitch accents, boundary tones, and other intonational peculiarities. However, the relationship between tones, prosody, and syntax also needs to be studied with non-canonical word orders, as only special word orders can reveal more subtle properties of the prosodic and tonal patterns of languages. This paper examines the prosody of discontinuous nominal phrases (NPs) in several Indian languages: three Indo-Aryan languages (Assamese, Bengali, and Hindi), two Dravidian languages (Tamil and Malayalam), and two Tibeto-Burman languages (Bodo and Meithei).[1] The purpose of this survey is to compare the tonal and prosodic structure of these languages with those of the better studied intonation languages, such as English and German. Indian languages are still understudied, even though this area of linguistic investigation has experienced a rapid development in the last decades. The present article is a modest contribution to this research area.

All data in this paper were elicited in the course of interviews with native speakers and linguists. The data come from a long questionnaire on discontinuous NPs that was elaborated for the elicitation of data on this topic. We also asked our colleagues to pronounce the sentences elicited, sometimes in different information structural contexts if available.

The article is structured as follows. In Section 2, discontinuous NPs are first given a definition and two principled distinctions in their grammar are introduced. The first one is syntactic. Discontinuous NPs can be 'hierarchy-inverting' or 'hierarchy-preserving'; see Fanselow & Féry (in preparation) for this distinction. The second division concerns their prosodic structure and distinguishes 'cohesive' from 'non-cohesive' prosodic structures. These two distinctions are useful in a large number of European languages, especially in those where pitch accents change along with pragmatics, for example Slavic and Germanic languages, but also Baltic, Caucasian, Greek, and most Romance languages, although it is not clear whether they can be considered universally valid.

[1] In the context of this research enterprise, we also collected data from four further Indo-Aryan languages (Gujarati, Maithili, Marathi, Oriya/Odia), and two further Dravidian languages (Kannada, Telugu) that we will not go into here because data collection could not reach a satisfactory level.

Section 3 addresses the role of information structure in the formation of discontinuous NPs: Focus, Givenness, and Topic are the categories used. It is shown that discontinuity is often related to differences in the information structural roles of the parts of the discontinuous construction.

Section 4 is a short survey of the main properties of the intonation system of some Indian languages. It is suggested that these Indian languages show a great deal of similarities across several families of languages.

In the following sections (Sections 5 to 7), the relationship between syntax and prosody of discontinuous NPs in the Indian languages mentioned above is examined from the following perspective: do we find a main division between cohesive and non-cohesive prosodic structure, especially in relation to a main syntactic division between hierarchy-preserving and hierarchy-inverting discontinuous NPs? The answer differs among the languages considered. From a syntactic perspective, the SOV base order and the free constituent order property present in nearly all Indo-Aryan languages lead us to expect a fairly rich system of discontinuous NPs of the type involving an independent generation of two NPs. From a prosodic perspective, Indian languages are prototypical examples of phrase languages, i.e. the intonation is not dependent on variable pitch accent placement, but is rather based on the existence of prosodic domains, each with a characteristic 'phrasal' intonation. We will see that the division between cohesive and non-cohesive patterns that has been made for intonation languages and that is summed up in Section 2.2 is blurred in the Indian languages. In line with this observation on prosody, we will also see that the hierarchy-preserving and hierarchy-inverting discontinuous NPs, see Section 2.1, are not always easy to tell apart syntactically in many Indian languages.

The last section contains a conclusion and a summary of the main results.

2 Two basic distinctions

This section introduces two basic distinctions. The first concerns the syntactic distinction between hierarchy-inverting and hierarchy-preserving discontinuous NPs, and the second concerns the prosodic distinctions between cohesive and non-cohesive prosodic structure.

2.1 Syntax: Hierarchy-inverting and hierarchy-preserving discontinuous NPs

Let us begin by introducing a major distinction between two types of discontinuous NPs. The core idea behind the notion of a discontinuous NP is, of course, that material that could fit into a single, standard, continuous NP can also appear scattered in two or more places in a single sentence in many languages, as illustrated by the German example in (1)a–b.

(1) a. Sie hat viele Bücher geschrieben (German)
she has many books written
'She has written many books.'
b. Bücher hat sie viele geschrieben
books has she many written
'As for books, she has written many.'

Not all constructions that pattern grammatically with simple instances of a discontinuous NP such as (1)b easily allow a reconstruction of their parts into a continuous form, a point to which we will briefly return below. For the standard case, reconstructability is the rule, however.

In the study of discontinuous NPs, it has proven useful to distinguish two different types of discontinuity. Normally, the left part of a discontinuous NP occupies a higher structural position than the right part, with the former c-commanding the latter, as illustrated in (1c).

(1) c. [CP **Bücher** [C' [C hat] [TP sie [**viele** [geschrieben]]]]] (German)
books has she many written
'As for books, she has written many.'

We can now classify discontinuous no NPs as to whether the structural hierarchies that hold among their parts are identical to the hierarchies we find in the corresponding continuous construction, or

whether that fails to hold. In particular, we focus on the lowest head of the continuous construction, which is normally the noun, and distinguish discontinuous NPs in which the lowest head[2] is structurally higher than the other part of the discontinuous constructions from discontinuous NPs in which this is not the case. For an illustration, see (2): in a continuous NP, the higher functional heads (or the phrases they form) such as quantifiers and determiners asymmetrically c-command the noun, and exactly this asymmetric c-command relation also holds in discontinuous (2). E.g., *wieviel* 'how many' c-commands *Bücher* 'books' in both (2)c and (2)d. We will label such discontinuous NPs 'hierarchy-preserving', a term that is not fully accurate but which captures the core cases. We called such constructions 'simple' discontinuous NPs in our earlier work (Fanselow & Ćavar 2002, Féry et al. 2007); a frequently used alternative label is 'left branch extraction'. The prosodic phrasing is indicated by means of subscripted Φ standing for Φ-phrase (prosodic phrase), roughly equivalent to a syntactic phrase (Selkirk 2009, 2011, Elfner 2015, Ito & Mester 2013).

(2) Hierarchy-preserving discontinuous NP (left branch extraction)

a. (**Hodně** má Marie **židlí.**)$_{\Phi}$ (Czech)
many has Mary chairs.GEN
'Mary has many chairs.'

b. (**U jake** vin pojide **misto**?)$_{\Phi}$ (Ukrainian)
in which he will.go town
'To which town will he go?'

c. (**Wieviel** hat Maria spannende **Büch-er** gelesen?)$_{\Phi}$ (German)
how.many has Mary fascinating book-PL read
'How many fascinating books did Mary read?'

d. **[Wieviel** [spannende **Büch-er]]** (German)
how.many fascinating book-PL
'how many fascinating books'

In the constructions in (3), the lowest head of the continuous NPs, viz. the noun, appears at the left periphery, the highest structural position. One can thus say that the hierarchy relative to the nominal head is inverted in this construction – the lowest head of the continuous construction (e.g., *Bücher* 'books' in (3)d c-commands the other heads in the discontinuous construction in (3)c. It makes sense to label these 'hierarchy-inverting' discontinuous NPs. The construction is often discussed under the label 'split topicalization'; in earlier work we spoke of 'inverted' discontinuous NPs.

(3) Hierarchy-inverting discontinuous NP (split topicalization)

a. (**Krastavic-i**)$_{\Phi}$ (vseki obica **presn-i i makl-i**.)$_{\Phi}$ (Bulgarian)
cucumber-PL everyone likes fresh-PL and small-PL
'Everyone likes fresh and small cucumbers.'

b. (**Knyžk-u**)$_{\Phi}$ (Marija pročytala **cikavui**.)$_{\Phi}$ (Ukrainian)
book.ACC.F Mary has.read interesting.ACC.F
'Mary has read an interesting book.'

c. (**Büch-er**)$_{\Phi}$ (hat sie **keine** spannenden gelesen.)$_{\Phi}$ (German)
book-PL has she none fascinating read
'She read no fascinating books.'

d. ([keine [spannenden **Büch-er**]])$_{\Phi}$ (German)
no fascinating book-PL
'no fascinating books'

There are several arguments for making a principled distinction between these two construction types. First, we observe that they need not co-occur. Georgian and the Slavic languages allow both kinds of discontinuous NPs (see, e.g., the Ukrainian examples above), but in German, the hierarchy-preserving version is quite restricted (unlike its inverting counterpart) – it is practically confined to wh-heads as in (2)c, and in this respect, the other modern Germanic languages pattern with German. In general, it is not uncommon that a language has hierarchy-inverting discontinuous NPs but lacks

[2] or, more precisely, the part that contains the lowest head.

the hierarchy-preserving counterpart altogether, as in Yucatec Mayan (Skopeteas et al., to appear), while the reverse constellation appears less frequently – it is, e.g., typical of North American languages (cf. Fanselow & Féry, in preparation).

Second, the grammatical details of the two constructions can differ even when they coexist in a single language. Often, the hierarchy-inverting construction comes with morphological changes that are absent in the hierarchy-preserving counterpart. In (4), a Bulgarian example, the continuous version (4)a and the hierarchy-preserving split (4)b have the same form of the head noun, namely accusative, while the hierarchy-inverting one (4)c needs a genitive plural form for the noun.

(4) Hierarchy-inverting discontinuous NP (split topicalization)

a.	Toj	ima	**tri**	**stol-a**.	(Bulgarian)
	he	has	three	chair-PL.ACC	
b.	**Tri**	ima	toj	**stol-a**.	
	three	has	he	chair-PL.ACC	
c.	**Stol-ove**	toj	ima	**tri**.	
	chair-PL.GEN	he	has	three	

'He has three chairs.'

This difference in grammatical behavior has been attributed to a fundamental difference in the status of the two parts in each case. In a hierarchy-inverting discontinuous NP, *both* parts appear to have the status of complete NPs (Fanselow 1988, van Riemsdijk 1989). This explains the need for morphological changes, and may also be responsible for the greater flexibility often observed with hierarchy-inverting discontinuity.[3] There are several proposals for how discontinuous constructions with two full NPs are generated, involving either movement (van Riemsdijk 1989), base generation (Fanselow 1988), or a mixture of both (Ott 2012) – a decision among these is not relevant for the purposes of the present paper. Hierarchy-preserving discontinuous NPs, on the other hand, do not have a uniform generation in the world's languages, but they all involve a very restrictive kind of movement process that does not create two complete and fully independent NPs, hence the absence of repair operations and the stricter locality we normally observe.

2.2 Prosody: Cohesive and non-cohesive prosodic structure

In addition to the syntactic subdivision just discussed, there is also a major prosodic distinction among discontinuous NPs. As will be shown below, it comes with the interesting potential of overwriting the syntactic distinction, for example for purposes of morphology or syntactic restrictions.

The major division in the prosodic structure of discontinuous NPs is referred to as 'non-cohesive' vs. 'cohesive'. The natural pairing with the two syntactic structures just introduced is non-cohesive with hierarchy-inverting, and cohesive with hierarchy-preserving. The first pattern is illustrated in Figure 1 for (3)a, reproduced with prosodic and tonal information in (5).

(5) Hierarchy-inverting discontinuous NP (split topicalization)

	H*L	H_Φ		H*	L*H		H*L	L_Φ L_ι
	(**Krastavic-i**)$_\Phi$	(vseki	obica	**presn-i**	**i**	**makl-i**)$_\Phi$		(Bulgarian)
	cucumber-PL	everyone	likes	fresh-PL	and	small-PL		

'Everyone likes fresh and small cucumbers.'

[3] For instance, what would be a syntactic island for movement can, nevertheless, be split up in a hierarchy-inverting fashion, and potential interveners such as negation do not affect the grammaticality of a hierarchy-inverting discontinuity (see Fanselow & Ćavar 2002 among others) in contrast to what holds for hierarchy-preserving discontinuity.

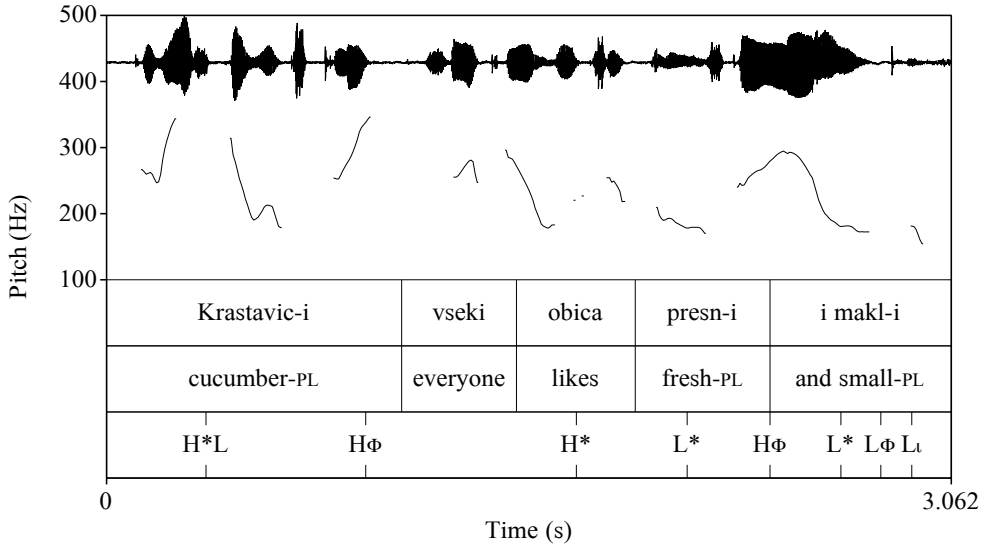

Figure 1. Non-cohesive intonation in a hierarchy-inverting Bulgarian sentence (3)a, recorded in June 2013 in Potsdam.

In the non-cohesive pattern, at least two (maximal) prosodic phrases (Φ-phrases) or two intonation phrases (ι-phrases) are present. Each of these prosodic domains must be well-formed. In particular, each one needs a pitch accent, at least in languages with pitch accents (and culminativity; see Hyman 2006, Féry 2017), and each one needs a boundary tone (Pierrehumbert 1980). Moreover, tone scaling is dependent on the relationship between the two parts of the discontinuous NP: there may be a downstep relationship between the two parts of the NP, but this is not obligatory. Downstep refers to the lowering of a high tone relative to a preceding high tone. Furthermore, the prosodic domains formed on each part of the discontinuous NPs do not need to be adjacent, which implies that more than two Φ-phrases may be involved in a non-cohesive prosodic pattern: there may be a prosodic phrase separating the discontinuous NP. The non-cohesive pattern is preferred in hierarchy-inverting splits. The prototypical case involves a topic on the fronted noun and a focus on the remnant element.

In the cohesive pattern, by contrast, illustrated in Figure 2 for (4)b, reproduced in (6) with prosodic and tonal patterns, only one (maximal) Φ-phrase (or ι-phrase) is typically present.

(6) Hierarchy-inverting discontinuous NP (split topicalization)

H*L LΦ Lι

(**Tri** ima toj **stol-a**.)Φ (Bulgarian)

three has he chair-PL.ACC

'He has three chairs.'

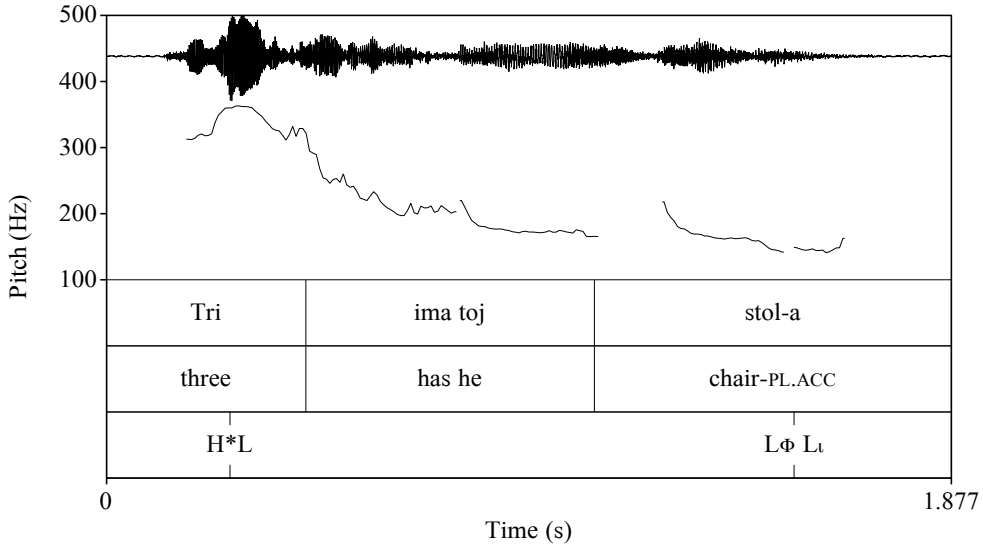

Figure 2. Cohesive intonation in hierarchy-preserving Bulgarian sentence (4)a, recorded in June 2004 in Potsdam.

The unique prosodic phrase implies that only one pitch accent is needed, as well as only one boundary tone of a prosodic phrase. The single pitch accent is often located on the fronted part of the discontinuous NP. The two parts of the discontinuous NP are minimally separated. Adding a prosodic phrase between the two parts may lead to ungrammaticality (but see Fanselow & Féry 2013 for examples involving non-cohesive prosodic structure in hierarchy-preserving discontinuous NPs in Slavic languages).

The natural pairing between syntactic and prosodic patterns illustrated in Figures 1 and 2 is also disrupted when a cohesive pattern is realized on a hierarchy-inverting split, as in (7), which requires a focus on the fronted noun and givenness on the remainder of the sentence.

(7) { A: Many of what did Mary read? }
B: (**Bücher** hat sie **viele** gelesen.)$_{\Phi}$ (German)
books.FOC has she many read
'She read many books.'

As Féry et al. (2007) show, it is the syntactic and not the prosodic type that determines the morpho-syntactic properties of the discontinuous NP, at least in Ukrainian.

3 The role of information structure

That information structure plays a major role in the formation of discontinuous NPs has been noticed by several authors for various languages; see Fanselow & Ćavar (2002), van Hoof (2007), Ott (2012), and others. It has been assumed by these authors that specific information structural features are responsible for such NPs, as well as for movement (and deletion) of parts of them. For instance, in the case of 'split topicalization', the feature [Topic] defining the makeup of the left periphery of a clause determines which part of a discontinuous NP can be placed in that position. According to Ott, the left-peripheral part of a hierarchy-inverting discontinuous NP is necessarily a frame-setting topic: it is the sole reason why the NP part of the discontinuous NP is placed peripherally, while the DP part of the NP remains behind.

But there are problems with this view related to the fact that the information structural roles of the parts of a discontinuous NP are not invariable. In fact, even though there are preferred roles for some positions in the sentence, it is typically the case that any position of a part of a discontinuous

NP can have several discourse functions.[4] This is illustrated in (8) through (10). A possible context for each variant is set in curly brackets.

(8) (Contrastive) topic on first part and focused second part
{How many Italian books and French newspapers did she buy?}
[Italienische Bücher]$_{Top}$ hat sie [drei]$_{F}$ gekauft. (German)
Italian books has she three bought
'She bought three Italian books.'

(9) Focused first part and given second part
{She bought three (Italian) watches, didn't she?}
Italienische [Bücher]$_{F}$ hat sie drei gekauft. (German)
Italian books has she three bought
'She bought three Italian books.'

(10) (Aboutness) topic for first part and a (contrastive) topic for second part
{She did something with three Italian books
(and something else with another four), what was that?}
[Italienische Bücher]$_{Top}$ hat sie drei [gekauft]$_{F}$. (German)
Italian books has she three bought
'She bought three Italian books.'

Moreover, under special prosodic conditions, the entire discontinuous NP can be part of a wide focus; see Fanselow & Lenertová (2011). However, in this case, the second part is necessarily unaccented. As a result, only one cohesive prosodic phrase is formed on the entire sentence.

An alternative explanation taking the role of information structure into account is that in an intonation language like German, a discontinuous NP is preferred when the two parts of the NP have different information structural roles, see Fanselow & Féry (in preparation) for detail.

4 Intonation of Indian languages

When investigating the prosody of Indian languages, it is important to be aware of the differences between the intonation of these languages and that of the better studied Germanic languages, such as English. At the phonetic level, all languages have melodies that can be decomposed into a series of low and high tones, but the function of these tones in the grammar can differ a great deal from one language to another (see Gussenhoven 2004 and Féry 2017 for explicit accounts). Most Indian languages, especially Indo-Aryan and Dravidian ones, are prototypical examples of so-called 'phrase languages'. The intonation system of these languages is based on phrasal tones, assigned at the prosodic level of the Φ-phrase, rather than on pitch accents, which are typical for intonation languages, or on lexical tones, assigned to words or to lexically specified stressed syllables. In several Indian phrase languages, each non-final Φ-phrase has an initial prominent low tone and a final high boundary tone. The final Φ-phrase of a declarative sentence has an initial high tone and a final low tone. According to Hayes & Lahiri (1991), Bengali weakly stresses the initial syllable of each word. However, phrasal tones assigned at the Φ-phrase level sometimes overwrite any tone that can be associated with lexical stress: in these cases, intonation at the phrase level is all that is left.

In the remainder of this paper, we do not try to formulate rules for the formation of Φ-phrases based on the morpho-syntax. Instead, we assume that in the default case, a grammatical word forms a Φ-phrase. In some cases, based on the tonal scaling and the tonal structure, it is assumed that a Φ-phrase is embedded in a larger one.

Sentence (11), illustrated in Figure 3, is from Bengali (Bangla) and it illustrates important prosodic properties of this language. The sentence has a complex syntactic structure, but the prosodic structure is quite simple. It forms a single ι-phrase that consists of a sequence of Φ-phrases, differing

[4] This has not escaped the attention of the authors just mentioned, who propose various solutions.

in length and in tonal scaling. The division of the ι-phrase into Φ-phrases correlates with the syntactic structure. In the pitch track, the typical rising contour of each Φ-phrase is clearly visible on all non-final Φ-phrases, and the final Φ-phrase has a falling contour. The rising contour of non-final Φ-phrases is analyzed as an initial prominent L* and a final H_{Φ}, following Hayes & Lahiri (1991) and Khan (2008, 2014). These authors introduce rules for the formation of Φ-phrases, and a survey of different tunes used for the expression of pragmatic meanings. A syntactic head forms a Φ-phrase together with a constituent that precedes it within its maximal projection. Additional evidence for Φ-phrases come from segmental processes like /r/-assimilation and voicing assimilation, which only take place inside Φ-phrases. Some variations in phrasing occur as a consequence of rhythm, style, and information structural roles, but these variations are still subject to special syntactic constraints.

In the phonological notation, the contour of the final Φ-phrase is simplified to H^*L_{ι} (not $H^*L_{\Phi}L_{\iota}$), since there is just one falling contour.

(11) L* H_{Φ} L* H_{Φ} L*H_{Φ} L* H_{Φ}
(((chele-bêla-te)$_{\Phi}$ (piṭar)$_{\Phi}$ ((skul)$_{\Phi}$ šeše)$_{\Phi}$)$_{\Phi}$ (Bengali)
child-time-LOC Peter school after
'When Peter was a child, after school…'

L* H_{Φ} L* H_{Φ}
(ta-r bondhu-der)$_{\Phi}$ (šathe dêkha kor-t-o)$_{\Phi}$
3-GEN friend-PL.GEN with seeing do-HAB-3
'…he used to meet with his friends…'

L* H_{Φ} L* H_{Φ} H* L_{Φ} L_{ι}
(eboṅ ta-der)$_{\Phi}$ (šathe)$_{\Phi}$ (ônek-khon) (khel-t-o.)$_{\Phi}$)$_{\iota}$
and 3-PL.GEN with much-time play-HAB-3
'…and play with them for a long time.'

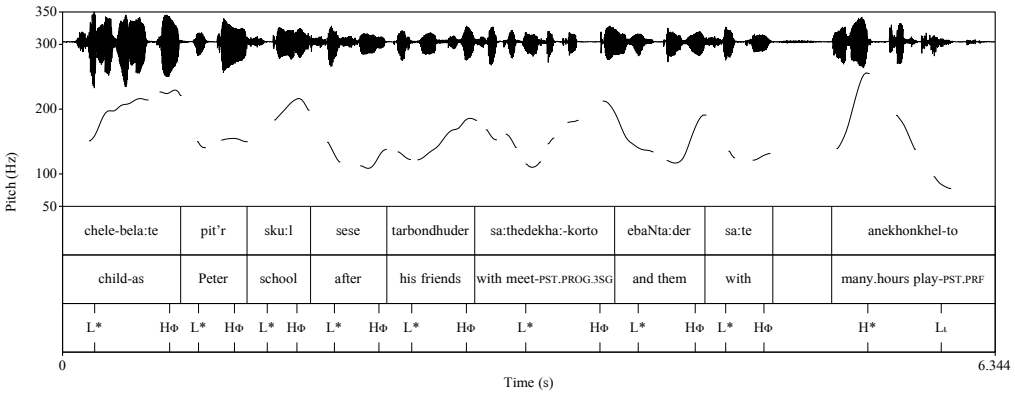

Figure 3. Tonal structure of a long Bengali sentence (11) in canonical word order.

Two comments are in order that hold for most of the languages addressed in the following subsections. The first one concerns the alignment of the two tones in their respective Φ-phrase. The final H_{Φ} is typically aligned as far to the right in its Φ-phrase as possible. But the initial low tone is not always linked to the first syllable of the Φ-phrase, though it seems to be systematically linked to the first syllable of a word, as in the Φ-phrase (*šathe dêkha korto*)$_{\Phi}$ 'used to meet with', where the first word *šathe* 'with' just smoothly interpolates from the high tone of the preceding Φ-phrase to the low tone on the first syllable of *dêkha korto* 'used to meet'. The same holds for *eboṅ tader* 'and them', where the low tone is associated with the first syllable of *tader*.

Second, tonal scaling of the high tones is crucial as it reflects the syntactic structure. Some H_{Φ} are higher than others. Compare for instance the high tone at the end of *skul* 'school' with the high tone at the end of *šeše* 'at the end', a postposition. The former one is much higher, and we assume that the scaling between these two tones reflects the syntactic relation they have with each other.

We express this by a recursive prosodic structure, homomorphic with the syntactic structure of the sentence (Féry 2017). Other high tones also display meaningful scaling in relation to each other. For instance, the first Φ-phrase *chelebêlate* 'as a child' ends much higher than the second one *piṭar* 'Peter'. The third Φ-phrase *skul* 'school' returns to nearly the same height as the first one. There is an upstep relationship between *tar bondhuder* 'his friends' and *šathe dêkha korto* 'used to meet with' that is probably motivated by the syntactic structure. The higher rising tone on the verb expresses a continuation rise. The same is true for the final three Φ-phrases. The high tone of the prefinal Φ-phrase is higher than the preceding ones. Remarkably, it is even higher than the first high tone in the sentence. This is related to the status of the preverbal position in Bengali as the focus position.

Pitch scaling relationships have been studied in syntactically complex sentences in Bengali by Khan (2008, 2014) and in Hindi by Kügler (2020), but we do not try to address this topic here. It seems to us that the relative paucity of tonal contours in these phrase languages may be compensated for by the richness of pitch scaling, and the richness of the devices for expressing phrasing.

Before turning to intonation in discontinuous NPs in the next sections, let us briefly examine post-focal compression in Assamese. This language has been given a detailed and careful prosodic analysis by Mahanta (2010) and Twaha (2017). Assamese resembles other Indian languages as far as the intonational structure is concerned. The 'building blocks of an intonational contour' (Keane 2014 for Tamil) are provided by the prosodic phrases. As in Bengali, the building blocks of non-final Φ-phrases are characterized by a low tone at the beginning of the prosodic phrase and a high tone at the end; see Figure 4 with the same sentence (12) in different information structural contexts. The end of the intonation phrase is delimited with a low $L_ι$ and an optional $H_ι$. Like Hayes & Lahiri (1991) and Khan (2008, 2014) do for Bengali, both Mahanta (2010) and Twaha (2017) analyze the initial low tones of Assamese as pitch accents.

(12)

L*	$H_Φ$	L*	$H_Φ$	L*	$H_Φ$	$L_ι$	
((nôgên-ê)$_Φ$		(nôyôn-ôk)$_Φ$		(mala)$_Φ$		(khuz-il-ê.)$_Φ$)$_ι$	(Assamese)
Nagen-ERG		Nayan-DAT		garland		ask-PST-3SG	
'Nagen asked Nayan for a garland.'							

As can be seen in Figure 4(c), the phrasal tones following the focused phrase are reduced or even suppressed. However, narrow focus on the pre-verbal argument does not change the prosodic pattern, as can be seen from (b).

Similar patterns have been shown for Indo-Aryan languages Bengali (Hayes & Lahiri 1991; Khan 2008, 2014) and Hindi (Patil et al. 2008), but also for the Dravidian language Tamil (Keane 2014), among other Indian languages.

Given the prosodic and intonational properties of the Indian languages discussed in this section, the question arises of how discontinuous NPs are realized. Do they present any special contour? Is there a difference between cohesive and non-cohesive prosodic contours?

The short answer to be developed in the remainder of this article is that there seems to be no clear prosodic difference between hierarchy-preserving and hierarchy-inverting discontinuous NPs, and that this parallels the absence of a clear difference between hierarchy-inverting and hierarchy-preserving syntax. The left parts of inverting discontinuous NPs are neither prosodically more integrated into the clause than their inverted counterparts nor have there been observations of differences in 'accentuation'.

Discontinuity of NPs may trigger the emergence of a new Φ-phrase on the displaced constituent. However, this Φ-phrase has no particular properties that would distinguish it from Φ-phrases triggered by other non-canonical syntactic structures.

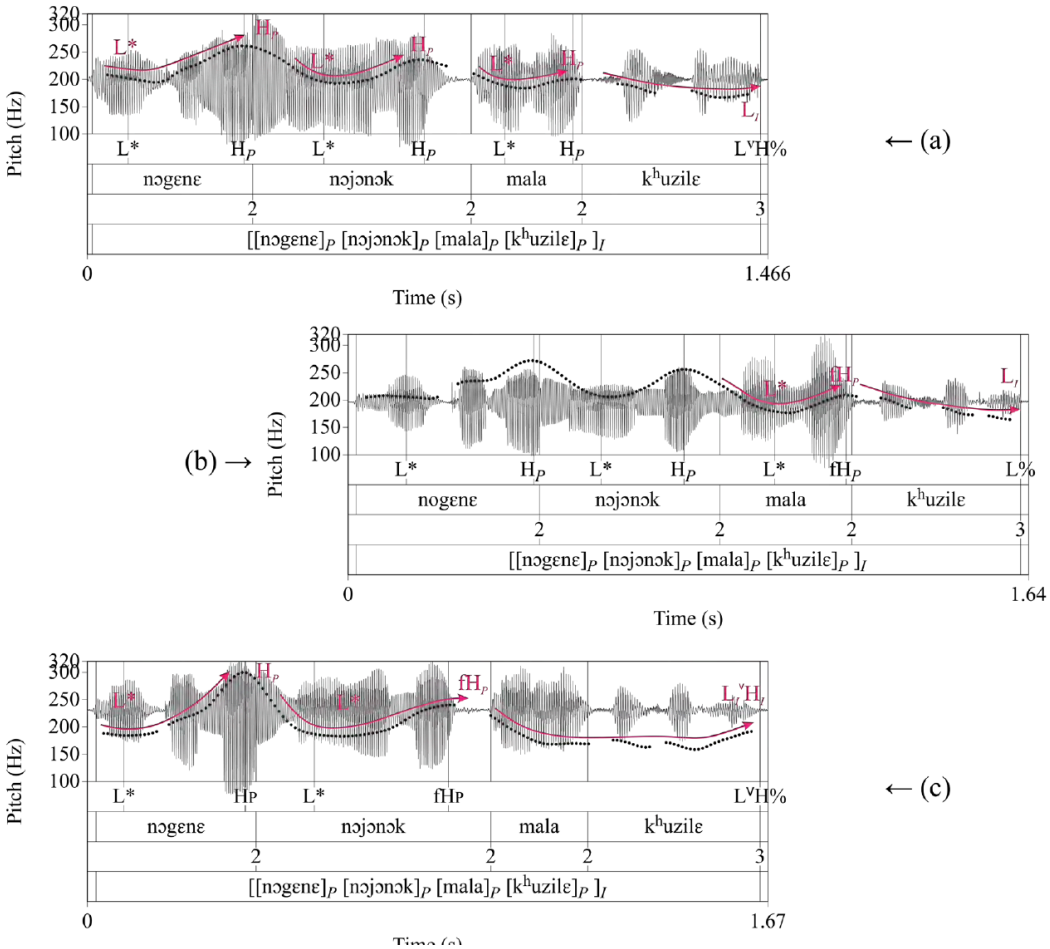

Figure 4. Tonal structure of Assamese sentence (12) in different information structural contexts: (a) in a wide focus, (b) with narrow focus on *mala* 'garland', and (c) with narrow focus on *nôyôn-ôk* 'Nayan-DAT' (from Twaha 2017).

5 Indo-Aryan: Hindi

By and large, the Indo-Aryan languages share a number of properties that make them tend to have discontinuous NPs among their repertoire of syntactic constructions. They come with an underlying SOV structure, and possess relatively free constituent order. At least since Déprez (1989), Mahajan (1990), and Dayal (1994), we know that both A-bar- and A-scrambling occur in Hindi, the latter being a close-to-perfect indicator of the possibility of discontinuous NPs (Fanselow & Féry, in preparation). Consequently, it is no surprise that all Indo-Aryan languages for which we have collected data possess discontinuous NPs, with the possible exception of Kashmiri.[5] For the other Indo-Aryan languages we have investigated (Assamese, Bengali, Gujarati, Hindi, Maithili, Marathi, Nepali, Odia, Panjabi, Sindhi, Sinhala), the existence of hierarchy-inverting discontinuous NPs is beyond doubt. It is quite remarkable that the distinction between inverting and preserving discontinuous NPs is blurred in these languages – quite in contrast to what we observe in other languages. This appears to be correlated with the observation that we also see no two classes of discontinuous NPs

[5] Claims in the literature that Kashmiri lacks discontinuous NPs are difficult to reconcile with sentences such as (i) which sound fine to at least some speakers (Darakshan Mir, p.c.).

(i) aer khyav yimav wāriy. (Kashmiri)
plums ate they.ERG many
'They ate many plums.'

of cohesive and non-cohesive type.

Two of these languages (Bengali, Odia) lack hierarchy-preserving discontinuous NPs, as illustrated by (13) for Odia (Kalyanamalini Sahoo, p.c.). They differ from other Indo-Aryan languages such as Hindi or Nepali in that scrambling is more restrictive, in particular in the post-verbal domain (see Simpson & Choudhury 2015 among others).[6]

(13) a. **bôhi**, piṭôr **bôhutô** guḍae pôḍh-i-ch-i. (Odia)
book Peter many CLA read-PRF-AUX-3
'Peter has read many books.'
b. **côuki**, piṭôr **ketu-ṭa** kiṇ-i-ch-i?
chair Peter how.many-CLA buy-PRF-AUX-3
'How many chairs has Peter bought?'
c. ***ketu-ṭa** piṭôr **côuki** kiṇ-i-ch-i?
how.many-CLA Peter chair buy-PRF-AUX-3
'How many chairs has Peter bought?'

Regarding discontinuous NPs in Hindi, we observe a high degree of flexibility (14). NPs may be split up even when both parts appear in the post-verbal domain of the SOV language; see (14)b.

(14) a. **kursi-yɛ̃** xarīd-ī th-ī̃ **tīn** rām-ne. (Hindi)
chair-PL buy-PRF.F be.PST-FPL three Ram-ERG
b. xarīd-ī th-ī̃ **kursi-yɛ̃** rām-ne **tīn**.
buy-PRF.F be.PST-FPL chair-PL Ram-ERG three
c. xarīd-ī **kursi-yɛ̃** th-ī̃ rām-ne **tīn**.
buy-PRF.F chair-PL be.PST-FPL Ram-ERG three
'Ram bought three chairs.'

Even the specifiers of an NP may be discontinuous, as shown in (15).

(15) a. kitne tum-ne *athletes*-k-ī ek foṭo dekh-ī? (Hindi)
how.many you-ERG athletes-GEN-F a photo see-PRF.F
'A picture of how many athletes did you see?'
b. kis-k-ī tum bahan-ke patī-se mil-e?
who-GEN-F you sister-GEN husband-ABL meet-PRF.M
'Whose sister's husband did you meet?'

As shown in (16), we can also observe constructions that come with the appearance of a hierarchy-preserving discontinuous NP: the quantifier precedes the noun. However, the possibility of splitting across a negation (*nahĩ*), as illustrated in (16)a–b, shows the absence of negative intervention effects in Hindi hierarchy-preserving NPs – while such intervention effects often occur with Left Branch Extraction in other languages. In this respect, and also with respect to locality, the two NP types are quite similar in Hindi.

---

[6] A reviewer notes that in some varieties of Bengali, sentences such as (i) and (ii) are fine:

(i) tin-ṭe kin-e-ch-il-o ram cear. (Bengali)
three-CLA buy-PRF-AUX-PST-3 Ram chair
(ii) ram tin-ṭe kin-e-ch-il-o cear.
Ram three-CLA buy-PRF-AUX-PST-3 chair
'Ram had bought three chairs.'

Note that these structures place the noun in the postverbal domain, i.e. they are more liberal with postverbal scrambling than what is reported by Simpson & Choudhury (2015). Furthermore, it is not entirely clear that (i) involves hierarchy-preserving discontinuous NPs. If the construction comes about by a rightward scrambling of the noun, it is the noun that ends up in the highest position, c-commanding the numeral, so that the discontinuity would indeed be of the inverting type.

(16) a. rām-ne **kitn-ī** nahĩ̃ pasand k-ĩ̃ **gāṛī-yā̃**? (Hindi)
Ram-ERG how.many-F NEG like do-PRF.FPL car.F-PL
b. **kitn-ī** rām-ne nahĩ̃ pasand k-ĩ̃ **gāṛī-yā̃**?
how.many-F Ram-ERG NEG like do-PRF.FPL car.F-PL
'How many cars does Ram not like?'

There is a further observation that casts some doubt on the expectation that Hindi discontinuous NPs follow the distinction between preserving and inverting subtypes neatly. Quite in contrast to what is observed in other languages, in which the formation of hierarchy-inverting discontinuity is always at least as flexible with respect to grammatical functions as for the hierarchy-preserving discontinuous NPs, the latter ones can be constructed with more grammatical functions than hierarchy-inverting ones. The sentences (17) through (20) show hierarchy-inverting NPs are impossible for indirect objects and the ergative subjects of transitive verbs.

(17) bahut sāre tum-ne ḍrāivar-õ-ko šahar-k-ā rāstā dikhā-y-ā? (Hindi)
many.M you-ERG driver.M-PL-DAT city-GEN-M way show-PRF-M
'Did you show the way to the city to many drivers?'

(18) *ḍrāivar-õ-ko tum-ne bahut sāre is šahar-k-ā rāstā batā-y-ā? (Hindi)
driver.M-PL-DAT you-ERG many.M this city-GEN-M way show-PRF-M
'Did you show the way to the city to many drivers?'

(19) a. kitne kal laṛk-õ-ne tumhẽ bulā-y-ā? (Hindi)
how.many yesterday boy-PL-ERG you.DAT call-PRF-M
'How many boys invited you yesterday?'
b. bahut sāre laṛk-õ-ne mujhe bulā-y-ā.
many.M boy-PL-ERG I.DAT call-PRF-M
'Many boys invited me.'
c. bahut sāre mujhe bulā-y-ā laṛk-õ-ne.
many.M I.DAT call-PRF-M boy-PL-ERG
'Many boys invited me.'

(20) *laṛk-õ-ne bulā-y-ā bahut sāre tumhẽ. (Hindi)
boy-PL-ERG call-PRF-M many.M you.DAT
'Many boys invited you.'

Some speakers of Hindi (Alok 2016) do not accept the formation of hierarchy-preserving NPs in the above constellation either, in line with our data for Gujarati. On the other hand, subjects and indirect objects can be split up even in the restrictive languages Bengali and Odia.

Closer inspection reveals that the constraint in question is due to a ban on the appearance of an overt case marker in hierarchy-inverting discontinuous NPs. Thus, when the subject appears in absolutive rather than ergative case, i.e. when it bears no case particle, it can be discontinuous, as shown by the contrast in (21) (Anoop Mahajan, p.c.).

(21) a. *bacc-õ-ne kal bahut sāre yah gānā gā-y-ā thā. (Hindi)
child-PL-ERG yesterday many this song sing-PRF-M be.PST
'Many children sang this song yesterday.'
b. bacc-e kal bahut sāre yah gānā gā-ẽ-g-e.
child-PL tomorrow many this song sing-PL-FUT-PL
'Many children will sing this song tomorrow.'

Data such as (21) show that the difference between the two discontinuous NP types with respect to grammatical functions is epiphenomenal – the topical noun in the left-peripheral position simply seems to be unable to bear an overt case marker. The relevant constraint not only affects 'standard' discontinuous NPs, but also constructions with two overt nouns, as in (22).

(22) a. pakšī to use sirf koyal pasand hãĩ. (Hindi)
bird TOP he.DAT only cuckoos like be.PRS
'As for birds, he likes only cuckoos.'
b. ??ciṛiyɑ̃̃ to bhārat me sirf koyal-ko saṅgīt pasand hai.
birds TOP India in only cuckoos-DAT music like be.PRS
'As for birds, in India only cuckoos like music.'
c. ciṛiyɑ̃̃ to bhārat me kal sirf cīlõ-ne hāthī-ko *attack* ki-y-ā.
birds TOP India in y'day only eagles-ERG elephants-DAT attack do-PRF-M
'As for birds, in India only eagles attacked elephants yesterday.'

We therefore assume that the two types of discontinuous NPs have roughly the same grammatical analysis in Hindi. This fundamental parallelism notwithstanding, constructions with a left-peripheral nominal topic are subject to a further case restriction (that we will not characterize in more detail here) that blocks certain discontinuous NPs, but is orthogonal to their formation. From a syntactic perspective, there is thus little reason to distinguish the two discontinuous NP types in Hindi.

This leads us to expect that the prosodic distinctions introduced above are also absent, or blurred, in Hindi. First, examine a sentence of the kind discussed above in its normal word order, uttered in an all-new context in (23).[7]

(23) Normal word order in Hindi
L* HΦ L*HΦ L* HΦ H* Lι
((rām-ne)Φ (**tīn**)Φ (**kursi-yɑ̃̃**)Φ (xarīd-ī th-ĩ̃.)Φ)ι (Hindi)
Ram-ERG three chair-PL buy-PRF.F be.PST-FPL
'Ram bought three chairs.'

Each word is realized in a separate Φ-phrase, and each high tone ending a Φ-phrase is (slightly) downstepped relative to the preceding one; the H* in the last Φ-phrase has the largest downstep. The low tones starting the Φ-phrases are clearly visible in the pitch track in Figure 5, and are perceptively prominent, even though the ones on *tīn* 'three' and *kursiyɑ̃̃* 'chairs' are located rather high in the register of this speaker.

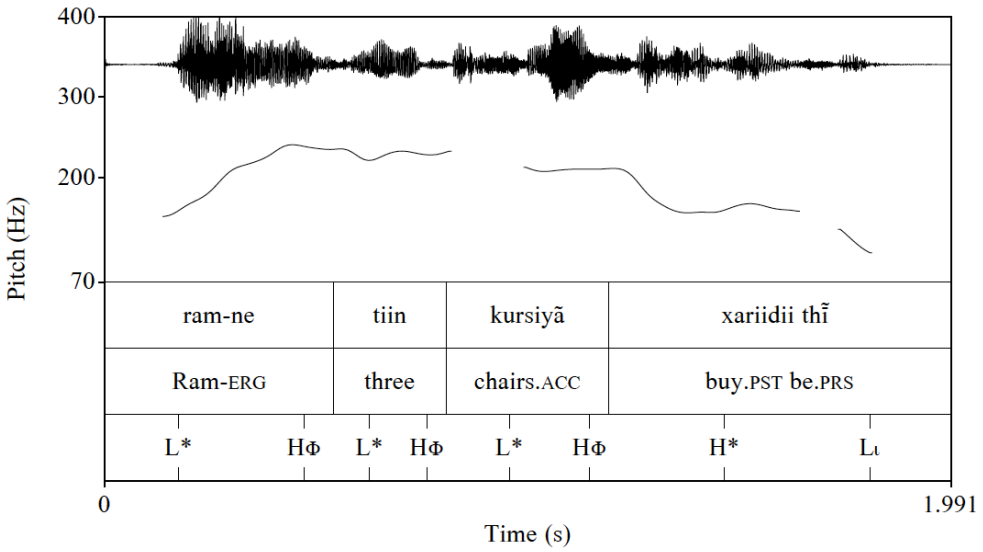

Figure 5. Canonical word order of Hindi, from sentence (23).[8]

[7] Realizations and comments on possible contexts for Hindi were provided by Rajesh Bhatt (p.c.).
[8] All Hindi sentences were recorded during the CreteLing Summer School in July 2018.

As a relatively unmarked example of a hierarchy-preserving discontinuous NP, consider (24) with verb finality. *Kursiyã̃* 'chairs' is not more focused or contrasted than it is in (23). It is in the pre-verbal position in both cases. Sentence-initial *tīn* 'three' may be felt to be slightly more prominent, but the reason is related to the fact that it is separated from its head noun, rather than because of its hypothetical topic role.

(24) Hierarchy-preserving discontinuous NP with verb finality in Hindi

L*HΦ	L* HΦ	L* HΦ	H*	Lι	
((**tīn**)Φ	(rām-ne)Φ	(**kursi-yã̃**)Φ	(xarīd-ī	th-ī̃.)Φ)ι	(Hindi)
three	Ram-ERG	chair-PL	buy-PRF.F	be.PST-FPL	

'Ram bought three chairs.'

Low and high tones defining the Φ-phrases can easily be spotted in the pitch track of this sentence in Figure 6. As shown for the sentence in its base word order, every word forms a separate Φ-phrase and the high tones of the sentence are in a downstep relation to each other. What changes is the tonal scaling among the tones rather than the phrasing itself or the distribution of the tones. The downstep is larger in Figure 6 than in Figure 5. However, at least in the present case, this difference does not seem to have an impact on the interpretation of the sentence.

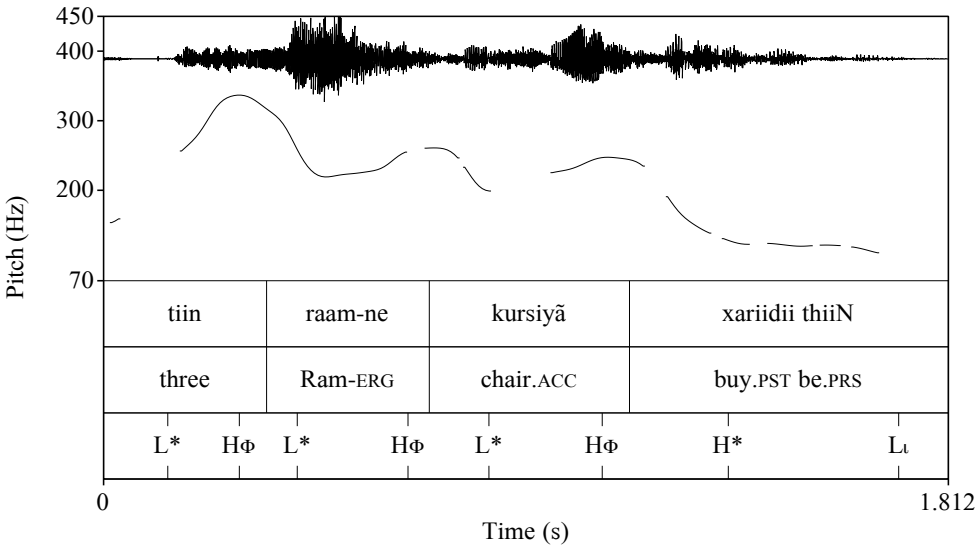

Figure 6. Hierarchy-preserving discontinuous NP with verb finality in (24).

In the next version of the sentence in (25), the head noun is post-verbal and focused. The context in which such a sentence may be uttered can be paraphrased as 'What has Ram bought three of?'[9] An important property of the pitch track in Figure 7 is that the word *kursiyã̃* 'chairs' is not particularly prominent from the point of view of prosody. It is realized with the typical final default tonal structure that has been illustrated in Figures 5 and 6, and analyzed phonologically with H* and Lι. A similar result, pointing to the absence of reliable prosodic differences between focused and given material, has been found for Hindi (Jyothi et al. 2014). However, a crucial prosodic cue appears to be the high boundary separating the auxiliary from the following narrow focus, which is very prominent due to cancellation of downstep. In other words, it is the tonal scaling of the boundary tone preceding the focused element that attracts attention to the focus; see also Féry et al. (2016) for a

[9] If the subject and the verb are inverted (***tīn*** *xarīdī thī̃ rāmne* ***kursiyã̃***), an even stronger contrast on *kursiyã̃* is called for. This latter version is felicitous when the sentence is continued by *aur tīn kitābẽ̃*, 'and three books'.

similar observation. A last interesting feature in this sentence is the recursive phrasing of the participle plus auxiliary. The participle alone has the tonal pattern of a Φ-phrase, and together with the auxiliary, it forms a larger Φ-phrase.

(25) Hierarchy-preserving discontinuous NP with verb finality in Hindi
L*HΦ L* HΦ L* HΦ L*HΦ H* Lι
((**tīn**)Φ (rām-ne)Φ ((xarīd-ī)Φ th-ɪ̃)Φ)Φ (**kursi-yɑ̃**.)Φ)ι (Hindi)
three Ram-ERG buy-PRF.F be.PST-FPL chair-PL
'Ram bought three chairs.'

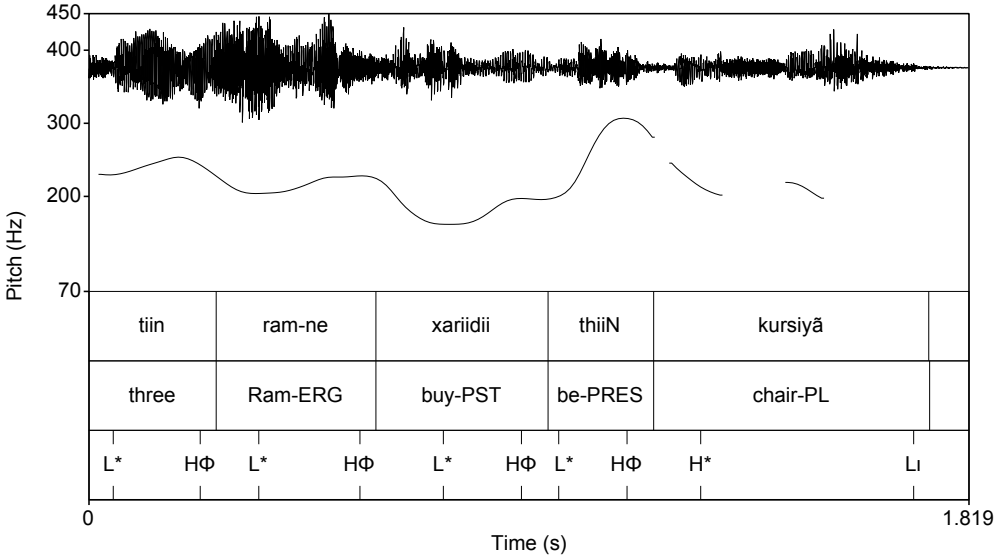

Figure 7. Hierarchy-preserving discontinuous NP in (25), in which the head noun is post-verbal and focused.

That the post-verbal position may be a preferred place of focus is also illustrated by the word order in (26), where the subject is post-verbal and focused: *it was Ram who bought three chairs*.

(26) Hierarchy-preserving discontinuous NP with post-verbal head noun and subject in Hindi
((**tīn**)Φ (xarīd-ī th-ɪ̃)Φ (**kursi-yɑ̃**)Φ (rām-ne.)Φ)ι (Hindi)
three buy-PRF.F be.PST-FPL chair-PL Ram-ERG
'Ram bought three chairs.'

However, in some cases, the post-verbal element is not focused, but given. Compare the next version in (27) and its pitch track in Figure 8. In this case, the focused element is pre-verbal *tīn* 'three' rather than post-verbal *kursiyɑ̃*, which is preferably interpreted as given. A prosodic difference between (25) and (26) on the one hand and (27) on the other hand lies in the prosodic attachment of the auxiliary. While it is this element that carries the boundary tone in Figure 7, it is part of the last Φ-phrase in Figure 8. In this case, it is the verb *xarīdī* that carries the high boundary tone. *Kursiyɑ̃* is uttered entirely at a low level, and the final fall takes place on *thɪ̃*.

(27) Hierarchy-preserving discontinuous NP with a pre-verbal numeral in Hindi
((rām-ne)Φ (**tīn**F)Φ (xarīd-ī)Φ (th-ɪ̃ **kursi-yɑ̃**.)Φ)ι (Hindi)
Ram-ERG three buy-PRF.F be.PST-FPL chair-PL
'Ram bought three chairs.'

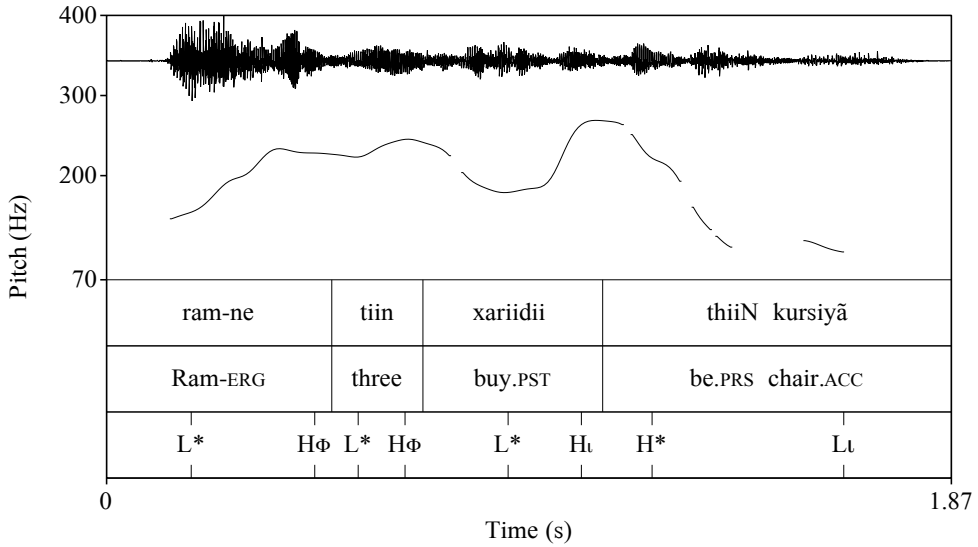

Figure 8. Hierarchy-preserving discontinuous NP in (27), in which the head noun is post-verbal but given.

Hierarchy-preserving discontinuous NPs can be formed with adjectives as well; see (28) and (29). Hindi allows intermediate discontinuous NPs, as in (29). In this particular case only the adjective is fronted, while the quantifier remains adjacent to the noun even though the adjective is located between Q and N in the base word order.

(28) **lāl** xarīd-ī th-ī us-ne **gāṛī**. (Hindi)
red buy-PRF.F be.PST-F he-ERG car.F
'He bought a red car.'

(29) **kāl-ī** mãĩ-ne dekh-ī th-ĩ̃ **tīn** **billi-yã̃**. (Hindi)
black-F I-ERG see-PRF.F be.PST-FPL three cat.F-PL
'I had seen three black cats.'

Turning now to the prosodic structure of hierarchy-inverting versions of discontinuous NP, the same word order freedom as before is observed. In (30) the focus-given relation among the two parts of the discontinuous NP is inverted relative to (27). It is again the pre-verbal word *kursiyã̃* that is focused, and the post-verbal numeral *tīn* 'three' is 'out of the way'. A possible context for this word order is: 'What was it that Ram bought three of?' Figure 9 shows that the prosodic structure remains unexceptional: each word, except the auxiliary, forms its own Φ-phrase.

(30) L* HΦ L* HΦ L* HΦ H* Lι
((rām-ne)Φ (**kursi-yã̃**)Φ (xarīd-ī th-ĩ̃)Φ (**tīn**.)Φ)ι (Hindi)
Ram-ERG chair-PL buy-PRF.F be.PST-FPL three
'Ram bought three chairs.'

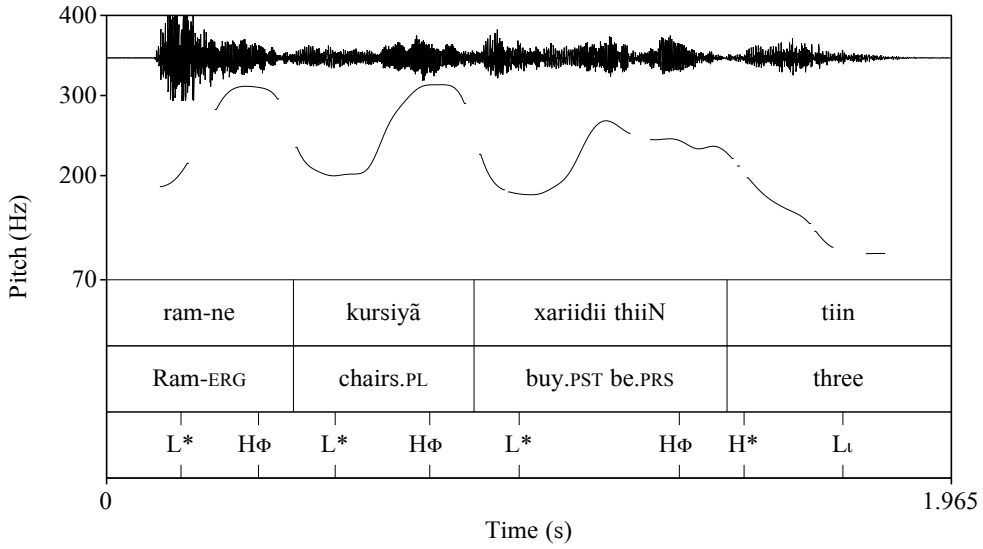

Figure 9. Hierarchy-inverting discontinuous NP in (30), in which the head noun is pre-verbal and focused.

In the next examples, base word order (31)a is compared to two hierarchy-inverting word orders in which the quantifier is separated from the fronted head noun. In (31)b, the quantifier is post-verbal and in (31)c, it is pre-verbal. All three versions of this sentence are perfectly natural. Quantifiers are intrinsically focused, and both non-canonical positions favor a focused reading of the quantifier.

(31) Base word order and hierarchy-inverting discontinuous NPs in Hindi

a. ((pareš-ne)$_{Φ}$ (**bahut sār-ī**)$_{Φ}$ (**kitāb-ẽ**)$_{Φ}$ (paṛh-ī th-ī̃.)$_{Φ}$)$_{ι}$ (Hindi)
Paresh-ERG many-F book.F-PL read-PRF.F be.PST-FPL

b. ((**kitāb-ẽ**)$_{Φ}$ (pareš-ne)$_{Φ}$ (paṛh-ī th-ī̃)$_{Φ}$ (**bahut sār-ī**.)$_{Φ}$)$_{ι}$
book.f-PL Paresh-ERG read-PRF.F be.PST-FPL many-F

c. L* HΦ L* HΦ L* H L HΦ H* Lι
((**kitāb-ẽ**)$_{Φ}$ (pareš-ne)$_{Φ}$ (**bahut sār-ī**)$_{Φ}$ (paṛh-ī th-ī̃.)$_{Φ}$)$_{ι}$
book.F-PL Paresh-ERG many-F read-PRF.F be.PST-FPL
'Paresh read many books.'

The sentence (31)c is illustrated in Figure 10. It can be seen that prosodic and tonal structure are unchanged, except for the fact that the complex expression *bahut sārī* 'many' also has a complex tonal structure, analyzed as L*HLH$_{Φ}$ in (31)c. High tones are downstepped relative to each other (except in *bahut sārī*).

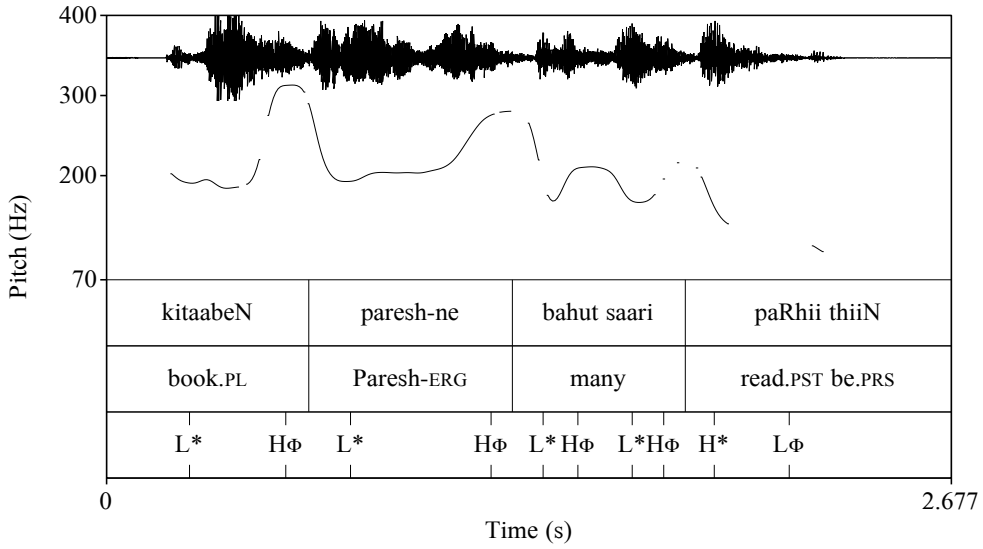

Figure 10. Hierarchy-inverting discontinuous NP in (31)c, with a pre-verbal quantifier.

By and large, the prosodic properties of hierarchy-inverting and hierarchy-preserving discontinuous NPs are quite similar in Hindi – in line with the profound syntactic non-distinctness of the two constructions.

6 Dravidian languages: Tamil and Malayalam

Dravidian languages do not allow discontinuous NP constructions as freely as Hindi or other Indo-Aryan languages, if at all, even though they all have a fairly free word order. Kannada seems to lack discontinuous NPs of the usual sort, while Malayalam (32), Tamil, and Telugu have hierarchy-inverting discontinuous NPs, but the construction is confined to underlying direct objects.

(32) **mantri-mār-e** jōṇ **pala-r-e** kaṇḍu. (Malayalam)
minister-PL-ACC John many-PL-ACC saw
'As for ministers, John saw many.'

6.1 Tamil

In this section, the prosodic structure of Tamil discontinuous NPs is investigated, which has been described by Keane (2007, 2014). Keane finds that the building blocks of intonation in Tamil consist of an initial low tone and a final high tone in the pre-final Φ-phrase, thus the same intonational pattern that was described for Hindi, Bengali, and Assamese. She observes that 'intonational differences between broad and narrow focus readings may be minimal. […] Intonational resources […] are limited: besides enforcing the presence of a rising contour on constituents that might otherwise lack one, manipulation of the relative scaling of f0 peaks appears to be the primary means of signaling semantic salience intonationally' (Keane 2014: 150). This description also corresponds to what has been found for Indo-Aryan languages.

Let us start the survey of the prosodic structure of discontinuous NPs in Dravidian languages with an example of a topic construction. Such sentences begin with an XP functioning as a free topic and marked as such by a postposition or similar devices. The topicalized NP is a hypernym of an NP that appears in the clause proper. We refer to such instances of a topicalized NP co-occurring with its clause-internal NP referent as 'double-noun constructions'.

As can be seen from Figure 11, a pitch track of this sentence, we find the pattern that is described

by Keane and that is also typical for Hindi. Each Φ-phrase has an initial low tone and a final phrasal high tone. The first Φ-phrase in this sentence is delimited by a very high boundary tone that separates the topicalized constituent from the remainder of the sentence. In Tamil, words are long and often complex and nearly each one of them consists of two series of L and H tones. Only the first of each Φ-phrase is marked by a star. In this sentence, except for the first high tone, we do not see much variation in tonal scaling, still the alternation of L and H tones is pervasive.

(33) Double-noun construction in Tamil

L* H L H_Φ	L* H L H_Φ	L*H L H_Φ	L*H L H_Φ	H* L_ι
(**paṟavai-gaḷ-uḷ**)$_\Phi$	(avan-ukku)$_\Phi$	(**nīla vaṇṇa.p**)$_\Phi$	(**paṟavai-gaḷ**)$_\Phi$	(piḍ-ikk-um.)$_\Phi$)$_\iota$
bird-PL-among	he-DAT	blue color	bird-PL	seize-FUT-3NSG

'As for birds, he likes blue ones.'

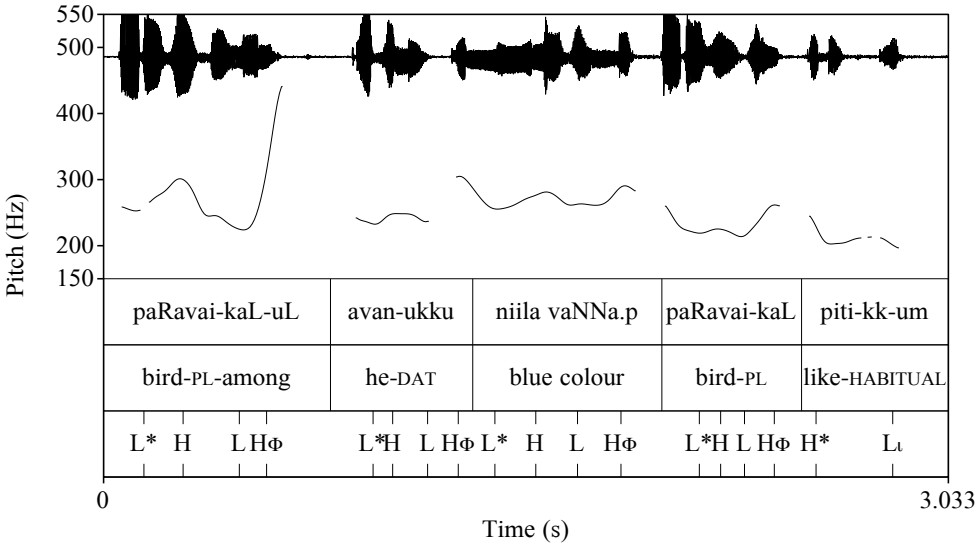

Figure 11. Tamil sentence (33) with a double-noun NP.[10]

Turning now to Φ-phrase formation and tonal structure in sentences containing a discontinuous NP, it can be seen once again that different word orders do not necessarily come along with different prosodic phrasings. In the two sentences in (34), which show base word order and an inverting discontinuous NP respectively, each word forms a separate Φ-phrase. Words are shorter than in (33), and alternation between L and H inside words is rarer. Furthermore, the quantifier *ettanai* 'how many' is prominent and bounded by a high boundary tone in both versions. In the base order in (34)a, illustrated in Figure 12, the Φ-phrases following the quantifier are compressed: the F0 register is smaller than at the start of the sentence. The tonal structure is present but not clearly perceptible.

(34) Base word order and hierarchy-inverting discontinuous NP in Tamil

a. ((pīṭṭar)$_\Phi$ (**ettanai**)$_\Phi$ (**cēr**)$_\Phi$ (nēttru)$_\Phi$ (pār-tt-ān?)$_\Phi$)$_\iota$ (Tamil)
Peter how.many chair yesterday see-PST-3MSG

b. ((pīṭṭar)$_\Phi$ (**cēr**)$_\Phi$ (nēttru)$_\Phi$ (**ettanai**)$_\Phi$ (pār-tt-ān?)$_\Phi$)$_\iota$
Peter chair yesterday how.many see-PST-3MSG
'How many chairs did Peter see yesterday?'

[10] All sentences from this and the next sections were recorded during the ICOLSI 39 in Patna in December 2017.

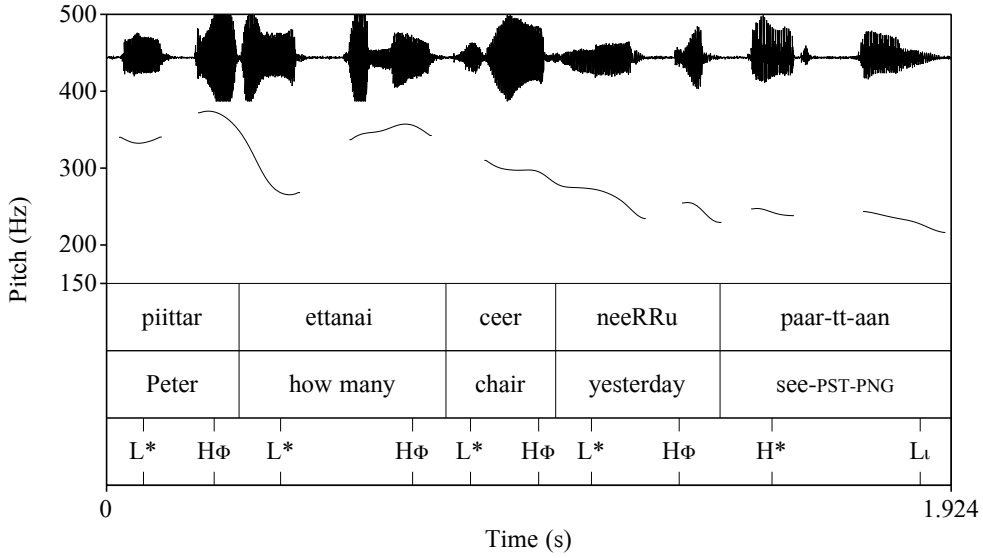

Figure 12. Base word order in Tamil sentence (34)a.

In the hierarchy-inverting NP (34)b illustrated in Figure 13, the first Φ-phrase has an initial low tone and a final phrasal H tone, but the following Φ-phrases, except for the one formed on the quantifier *ettanai* 'how many', are tonally inconspicuous with a tonal interpolation between the first H_Φ and the L* of the quantifier. This realization may correspond to integration of *cēr* 'chair' and *ettanai* 'how many' in one Φ-phrase, but since there is no strong reason to assume this, the solution adopted here is to assume that the prosodic phrasing is unchanged (each word forms its own Φ-phrase), but the tonal structure is eliminated.

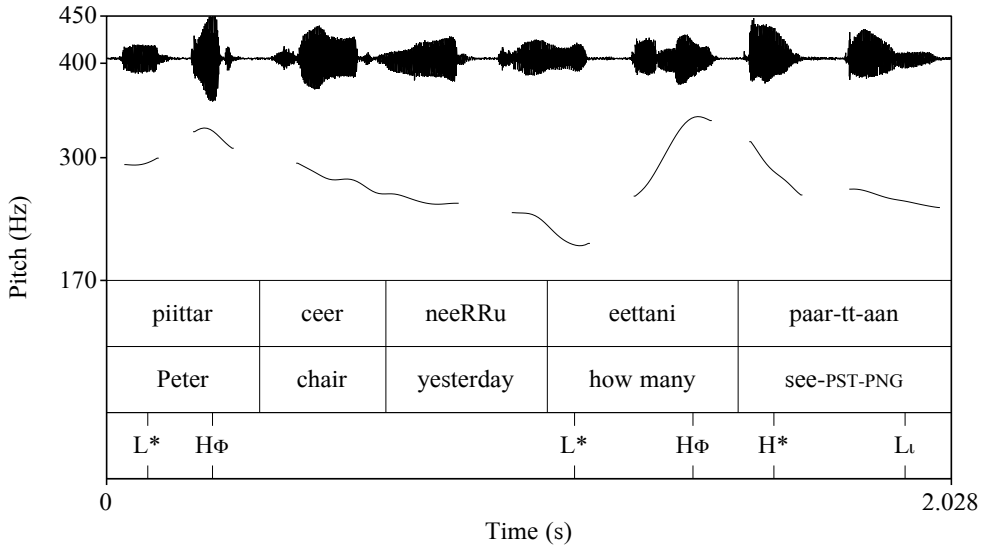

Figure 13. Hierarchy-inverting discontinuous NP in Tamil sentence (34)b.

6.2 Malayalam

As we mentioned earlier, Malayalam has hierarchy-inverting discontinuous NPs as in (35) when the XP moves to the topic position, but again, only for direct objects and subjects of unaccusative verbs.

Malayalam seems to allow more discontinuous NPs than the other languages, but this may be due to the fact that this language has *-āṇŭ*, a copula playing the role of a focus particle (FOC in the glosses), as illustrated in (36).

(35) **mantri-mār-e** jōṇ **pala-r-e** kaṇḍu. (Malayalam)
minister-PL-ACC John many-PL-ACC saw
'As for ministers, John saw many.'

(36) Focus particles (copula) in Malayalam
a. **etra-āṇŭ** mēri kaṇḍa **kasēra-gaḷ**? (Malayalam)
how.many-FOC Mary saw chair-PL
'How many chairs did Mary see?'
b. mūnnŭ-āṇŭ mēri vāṅṅicca kasēra-gaḷ.
three-FOC Mary bought chair-PL
'Mary bought three chairs.'

Examples comparing a sentence with base word order and the same sentence with a discontinuous NP appear in (37). The focus particle is attached to *etra kasēragaḷ* 'how many chairs' in the continuous order, but to *etra* 'how many' in the discontinuous order. In both cases, it has its own rising contour, which is analyzed here as forming an embedded Φ-phrase in the Φ-phrase formed by its host, a recursive structure also seen in Basque (Elordieta 2015), Irish (Elfner 2015), and Japanese (Kubozono 2007, Ishihara 2014). As for as the remaining tonal pattern, the same pattern as before is found, with initial low tones and final high tones in most Φ-phrases and a fine tonal scaling. The last word of the sentence, the verb in (37)a and the head noun in (37)b, has a low and falling tonal pattern.

(37) Base word order and hierarchy-preserving discontinuous NP in Malayalam
L* H_Φ L* H L H_Φ L* H_Φ H* L_ι
a. ((**etra**)$_\Phi$ (**kasēra-gaḷ**-(**āṇŭ**)$_\Phi$)$_\Phi$ ((pīttar)$_\Phi$ vāṅṅicca-adŭ.)$_\Phi$)$_\iota$ (Malayalam)
how.many chair-PL-FOC Peter bought-NMZ
LH L H_Φ L* H_Φ L* H_Φ H* L_ι
b. ((**etra**-(**āṇŭ**)$_\Phi$)$_\Phi$ (pīttar)$_\Phi$ (vāṅṅicca)$_\Phi$ (**kasēra-gaḷ**.)$_\Phi$)$_\iota$
how.many-FOC Peter bought chair-PL
'How many chairs did Peter buy?'

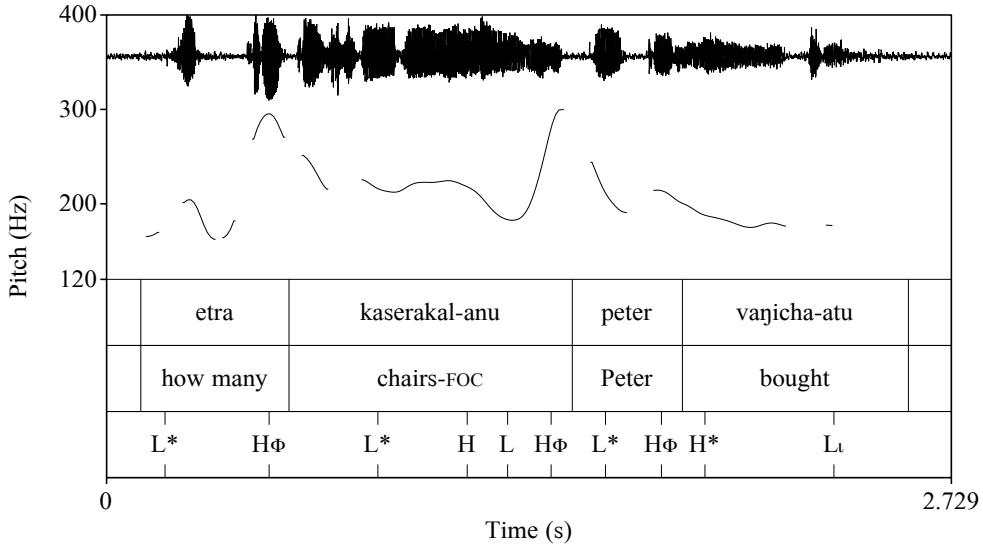

Figure 14. Malayalam sentence (37)a in base word order.

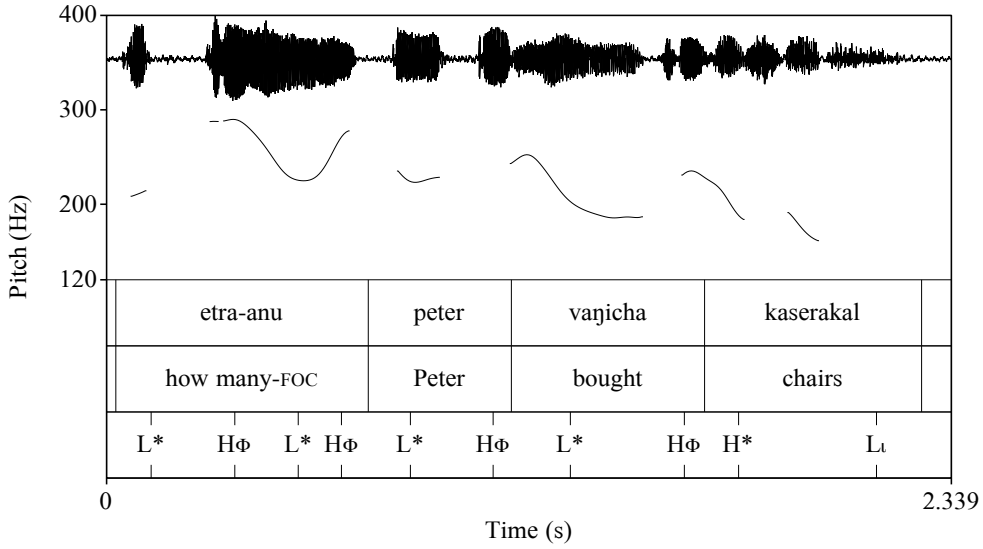

Figure 15. Malayalam sentence (37)b with hierarchy-preserving discontinuous NP.

To conclude this section, the tonal structure of both sentences with base word order and sentences with a discontinuous NP is very similar to that found in the Indo-Aryan languages examined.

7 Tibeto-Burman: Bodo and Meithei

In this section, two Tibeto-Burman languages are compared in regard to their intonational properties in base word order and in discontinuous NPs. Since there are only very few studies investigating the tonal and prosodic patterns of these languages, the results of this section are largely explorative and need more experimental investigation. On the basis of the descriptions of Indo-Aryan and Dravidian languages, it is tentatively proposed that the tonal patterns of Bodo and Meithei are superficially similar to those of these languages, but see some caveats below.

7.1 Bodo

Bodo (or Boro; Bodo-Garo, Brahmaputran) allows both hierarchy-inverting and hierarchy-preserving discontinuous NPs. Let us start the prosodic investigation of this language with a hierarchy-inverting discontinuous construction as in (38), drawing on intonational descriptions of the language in Das & Mahanta (2019). The left-peripheral noun has inflectional suffixes that participate in the tonal pattern: we find two rising contours, one on the nominal stem and one on the suffixes, and the same holds for the second nominal head followed by the exclusive particle. The pronominal subject has a simple rising contour and the final words, an adverb and the verb, carry the final falling contour.[11] In other words, we again find the same phrasal intonation that was described for the Indo-Aryan and Dravidian languages. In the Bodo sentence (38), downstep is present.

(38) Discontinuous construction in Bodo

L*HΦ L*	HΦ	L* HΦ	L* H L	HΦ	H*	Lι
((**dau-(phwr-khwu**)Φ)Φ		(bi-yw)Φ	(**gwthang-phwr-khwu**-lo)Φ		(mwzang	mwn-w.)Φ)ι
bird-PL-ACC		he-NOM	blue/green-PL-ACC-only		good	find-PRS

'As for birds, he only likes blue/green ones.'

[11] Das & Mahanta (2019) analyze the H tones of *bi-yw* 'he-NOM' and *dau-phwr-khwu* 'birds-ACC' and the L tones of *bai* 'buy' and *mwzang* 'good' as lexical tones.

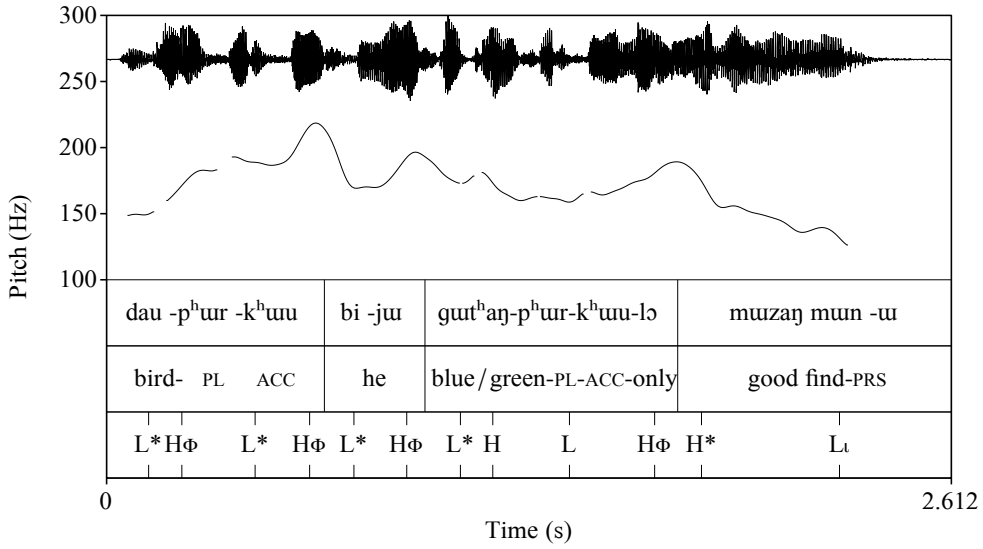

Figure 16. Hierarchy-inverting discontinuous construction in Bodo sentence (38).

In the discontinuous NP examples that we collected, the hierarchy-preserving NPs are grammatical only when the second part is placed post-verbally. In fact, the focal quantifier or numeral is the only preverbal element in these examples, giving Bodo the appearance of a verb-second language (but only in these contexts). As was the case for the other languages examined here, Bodo largely confines discontinuous NPs to direct objects.[12]

Turning to the prosodic and tonal structure, phrasing does not change much, only register changes are pervasive. In the pitch track of (39)a in Figure 17, the wh-word *gongbwise* 'how many' has a high boundary tone. The same is true for Figure 18, but there the wh-word is high anyway by virtue of being initial in the sentence.

(39) Bodo sentence in base word order and as discontinuous NP

L* HΦ L* HΦ L*HΦ H* Lι

a. ((pitar-a)Φ (**gongbwise**)Φ (**masi**)Φ (bai-khw?)Φ)ι (Bodo)
Peter-NOM how.many chair buy-PRF.Q

L* HΦ L* HΦ L* HΦ H* Lι

b. ((**gongbwise**)Φ (bai-khw)Φ (pitar-a)Φ (**masi**?)Φ)ι
how.many buy-PRF.Q Peter-NOM chair
'How many chairs did Peter buy?'

[12] Chelliah (1997: 120) observes for Meithei that post-verbal elements are given information. Predicate focus is a pragmatic condition that favors the presence of post-verbal material. In spite of the 'afterthought' nature of the second split part in pragmatic terms, it must be integrated quite firmly into the clause proper, because the construction is confined to direct objects in both Bodo and Meithei – neither subjects nor indirect objects can be split up.

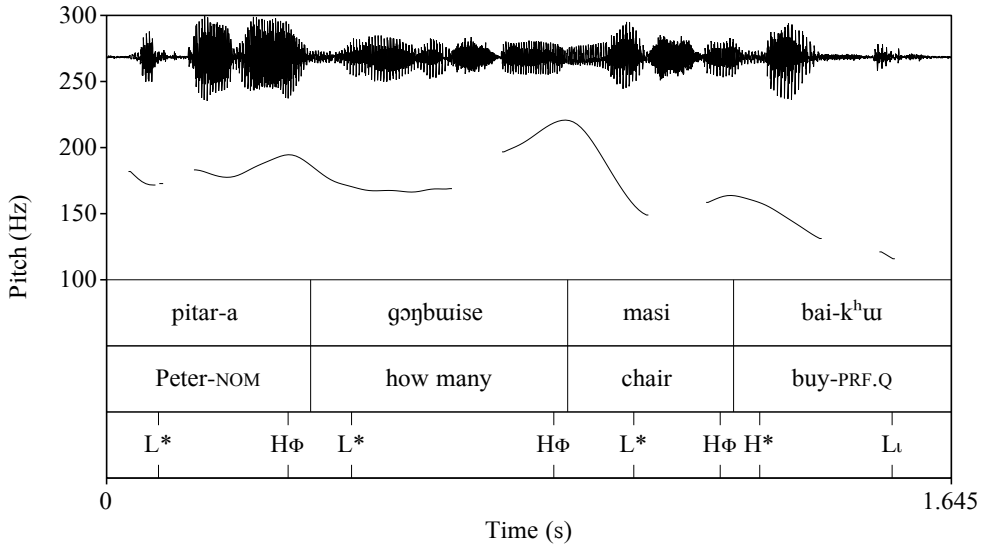

Figure 17. Sentence (39)a in base word order in Bodo.

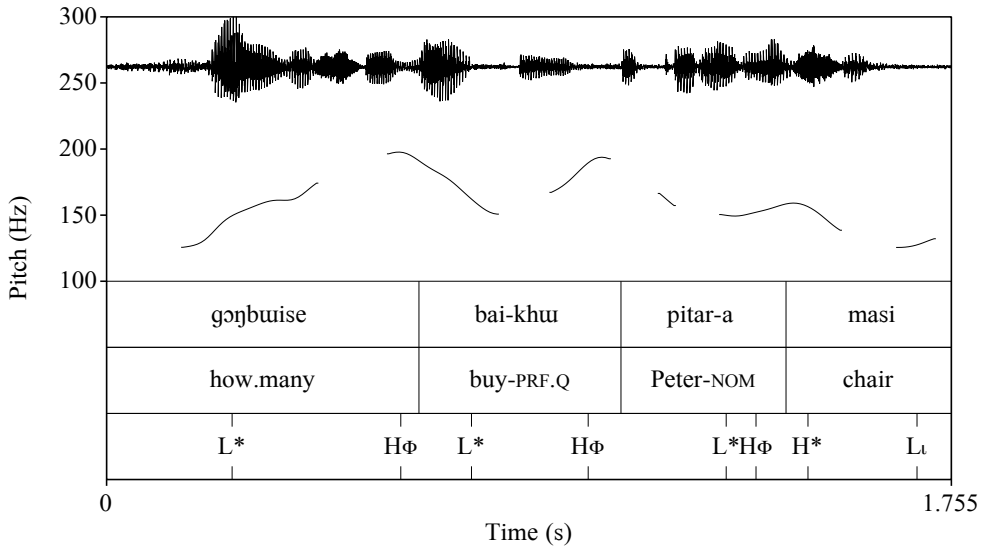

Figure 18. Hierarchy-inverting discontinuous NP with a post-verbal head in Bodo sentence (39)b.

7.2 Meithei

Meithei (or Meetei, Meitei, Manipuri; Kuki-Chin-Naga) is also very permissive as far as discontinuous NPs are concerned: it has both hierarchy-inverting and hierarchy-preserving examples, as well as double-noun constructions. The pair of examples in (40) shows a sentence in base word order and the same sentence with a hierarchy-inverting NP in which the adjective is post-verbal.

(40) Meithei sentence in base word order and as hierarchy-inverting discontinuous NP

L* HΦ L* HΦ L* HΦ L* Lι

a. ((pitar-nā)Φ (**nungaiba**)Φ (**lāirik amā**)Φ (pā-re.)Φ)ι (Meithei)
Peter-ERG interesting book one read-PST

L* HΦ L* HΦ L* HΦ L* Lι

b. ((pitar-nā)Φ (**lāirik**)Φ (pā-re)Φ (**nungaiba amā**.)Φ)ι
Peter-ERG book read-PST interesting one
'Peter read an interesting book.'

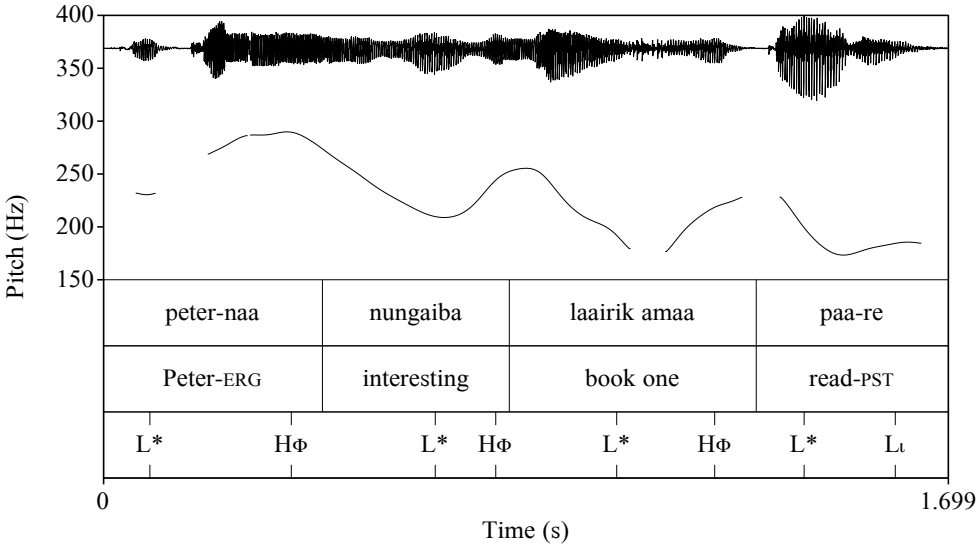

Figure 19. Meithei sentence (40)a in base word order.

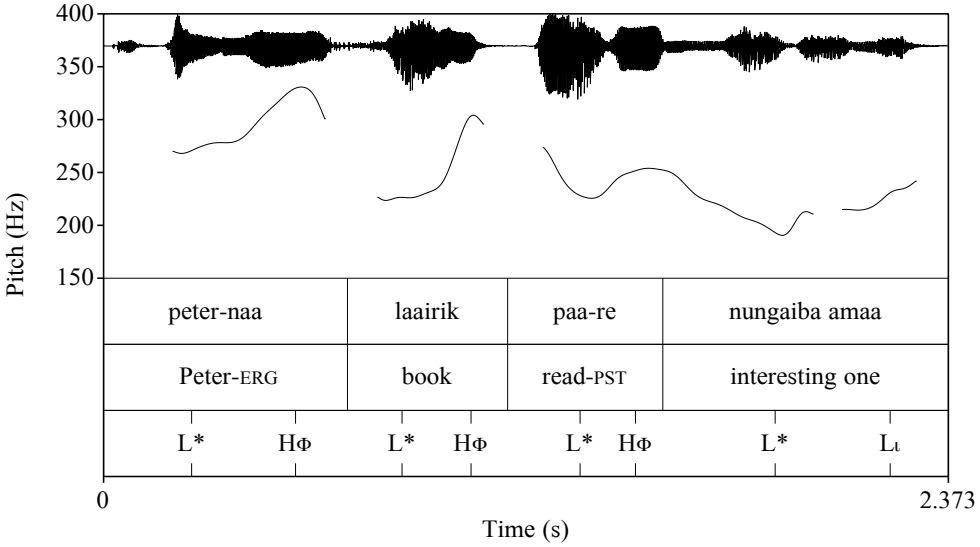

Figure 20. Meithei sentence (40)b, with a hierarchy-inverting discontinuous NP with a post-verbal quantifier.

When comparing the position of the low tones in the Tibeto-Burman languages with that in the Indo-

Aryan and Dravidian languages, it is conspicuous that they are placed later in their Φ-phrase, closer to the final H tone. It could be the case that the similarity in the tonal structure of all the Indian languages examined in this paper will turn out to be illusory and that the rising contour found in all the languages cannot be analyzed as resulting from the same underlying tones. It is not possible to give an informed answer to this question here.

8 Conclusion

In this paper, the prosodic structure of sentences containing a discontinuous NP has been examined in several Indian languages. The main research question was to test the division that Fanselow & Féry (in preparation) propose for a large number of languages, and that we call non-cohesive vs. cohesive intonation. In these languages, a continuous NP is typically included in a single prosodic phrase (Φ-phrase) that bears a unique information structural role. By rendering an NP discontinuous, the two parts of the NP may carry different roles and different tonal structures. Specifically, a special intonation can then be produced on the preposed phrase playing the role of the topic. There is also F0 raising on the focus and F0 compression on the post-focus material. We could also show that a non-cohesive intonational contour is typically associated with a hierarchy-inverting type of discontinuous NP and a cohesive intonational contour is preferred on a hierarchy-preserving one, although the pairing is not obligatory.

The specific question addressed above was whether the division between non-cohesive and cohesive intonation – and secondarily the division between hierarchy-inverting and hierarchy-preserving discontinuous NPs – is universal or whether it depends on specific intonational and prosodic properties. Indian languages are a good testing area because the intonation of these languages is different from that of intonation languages. They have a so-called phrase intonation because each content word typically forms a Φ-phrase of its own, and the tonal structure of the resulting Φ-phrases does not differ much, except for the sentence-final one in declarative sentences, which has a falling contour. The non-final phrases nearly always consist of an initial prominent low tone (written L*) and a final phrasal high tone (written H_Φ). It is important to realize that some grammatical features resulting from information structure are common to both intonation languages and Indian languages, namely word order changes and tonal scaling. In other words, NP discontinuity obviously elicits word order changes, and F0 can be raised or lowered depending on the focused or given status of the parts of the NP.

What does change in intonation languages is both the number of Φ-phrases and their tonal form, as a consequence of their pragmatic role in the sentence. And these are the features that do not change in Indian languages. The number of Φ-phrases is left unchanged because the noun and its modifier form different Φ-phrases to begin with, and the tonal pattern of the phrase does not change either: it does not depend on the information structural roles assumed here: focus, givenness, and topic.

To conclude, because of these properties, there is no clear difference between sentences in base word order, sentences containing hierarchy-inverting discontinuous NPs, and sentences containing hierarchy-preserving NPs as far as the prosodic and intonational patterns of the Indian languages examined in this paper are concerned. In other words, there is no clear prosodic division between non-cohesive and cohesive patterns.

The effect of information structure has not been tested systematically on the data presented in the paper, but for the cases that were tested, word order is crucial and tonal scaling is dependent on it. Tonal scaling is an important component of intonation in all the Indian languages discussed above, although its role is not completely clear. A constituent in focus is not necessarily raised in its F0, and sometimes, the preceding boundary seems to be at least as important.

What was not addressed in the paper is how Φ-phrases are mapped from morpho-syntax. Even though it is often the case that every word builds its own Φ-phrase, we can only suspect that the prosodic embedding that we could identify in some cases is much more common than we could demonstrate here. Tonal scaling is again the cue to prosodic embedding, but this must be the subject of separate research.

And the last point that needs further investigation concerns the alignment of the initial low tone, which seems to be later in the Tibeto-Burman languages than in the Indo-Aryan and Dravidian languages, although this point also needs more careful analysis.

Acknowledgements

This paper couldn't exist without the help and competence of many colleagues. Our thanks go to Rajesh Bhatt, Ayesha Kidwai, Anoop Mahajan, and Shravan Vasishth for Hindi; Ganesan Ambedkar, Arathi Kannan, and Keerthana Balasubramani for Tamil; Kiran Kishore, Rosmin Mathew, and Reshma Jacob for Malayalam; Aleendra Brahma and Ratul Mahela for Bodo; Tikaram Poudel for Meithei; Darakshan Mir for Kashmiri; Kalyanamalini Sahoo for Odia; and Svetlana Petrova, Milena Kühnast, and Bistra Andreeva for Bulgarian. A previous version of this paper was presented at the 39th International Conference of the Linguistic Society of India (ICOLSI-39) in Patna in December 2017 and at the Workshop on South Asian Prosody in Konstanz on 18 June 2018. Thanks are due to the organizers and the participants of these events. Last but not least, many thanks to the editor of this special issue, Sameer ud Dowla Khan, for his thoughtful reading of the draft of this paper. He provided many helpful comments.

References

Alok, Deepak. 2016. The syntax of split: The case of Hindi and Magahi. In M. Menon and S. Syed, eds., *Proceedings of Formal Approaches to South Asian Languages 6*, pages 53–61. https://ojs.ub.uni-konstanz.de/jsal/index.php/fasal/article/view/100

Chelliah, Shobhana. 1997. *A grammar of Meithei*. Berlin: De Gruyter.

Das, Kalyan and Shakuntala Mahanta. 2019. Intonational phonology of Boro. *Glossa: A Journal of General Linguistics* 4.

Dayal, Veneeta Srivastav. 1994. Binding facts in Hindi and the scrambling phenomenon. In M. Butt, T. H. King, and G. Ramchand, eds., *Theoretical Perspectives on Word Order Issues in South Asian Languages*, pages 237–61. CSLI Publications.

Déprez, Viviane. 1989. On the typology of syntactic positions and the nature of chains: move α to the specifier of functional projections. PhD. Thesis, MIT, Cambridge, MA.

Elfner, Emily. 2015. Recursion in prosodic phrasing: evidence from Connemara Irish. *Natural Language and Linguistic Theory* 33:1169–208.

Elordieta, Gorka. 2015. Recursive phonological phrasing in Basque. *Phonology* 32:49–78.

Fanselow, Gisbert. 1988. Aufspaltung von NPn und das Problem der "freien" Wortstellung. *Linguistische Berichte* 114: 91–113.

Fanselow, Gisbert and Damir Ćavar. 2002. Distributed deletion. In A. Alexiadou, ed., *Theoretical approaches to universals*, pages 65–107. Amsterdam: Benjamins.

Fanselow, Gisbert and Caroline Féry. 2013. A comparative perspective on intervention effects on left branch extractions in Slavic. In W. Sulym, ed., *Schriften für Alla Paslawka*. Lwiw University.

Fanselow, Gisbert and Caroline Féry (in preparation). *Discontinuous nominal phrases*. Cambridge: Cambridge University Press.

Fanselow, Gisbert and Denisa Lenertová. 2011. Left peripheral focus: Mismatches between syntax and information structure. *Natural Language and Linguistic Theory* 29:169–209.

Féry, Caroline. 2017. *Intonation and prosodic structure*. Cambridge: Cambridge University Press.

Féry, Caroline, Pramod Pandey, and Gerrit Kentner. 2016. The prosody of Focus and Givenness in Hindi and Indian English. *Studies in Language* 40:302–39.

Féry, Caroline, Alla Paslawska, and Gisbert Fanselow. 2007. Discontinuous constructions in Ukrainian. *Journal of Slavic Linguistics* 15:3–48.

Gussenhoven, Carlos. 2004. *The Phonology of tone and intonation*. Cambridge: Cambridge University Press.

Hayes, Bruce and Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9. 47–96. doi: 10.1007/BF00133326

van Hoof, Hanneke. 2007. Split topicalization. In M. Everaert and H. van Riemsdijk, eds., *The Blackwell companion to syntax*, pages 410–65. London: Blackwell.

Hyman, Larry. 2006. Word-prosodic typology. *Phonology* 23:225–57.

Ishihara, Shinichiro. 2014. Match theory and the recursivity problem. In S. Kawahara and M. Igarashi, eds., *Proceedings of Formal Approaches to Japanese Linguistics 7*, pages 69–88.

Ito, Junko and Armin Mester. 2013. Prosodic subcategories in Japanese. *Lingua* 124:20–40.

Jyothi, Preethi, Jennifer Cole, Mark Hasegawa-Johnson, and Vandana Puri. 2014. An investigation of prosody in Hindi narrative speech, *Proceedings of Speech Prosody*. http://www.ifp.illinois.edu/~pjyothi/files/SP2014.pdf.

Keane, Elinor. 2007. Distribution and alignment of F0 contours in Tamil. In *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany.

Keane, Elinor. 2014. The intonational phonology of Tamil. In S.-A. Jun, ed., *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Khan, Sameer ud Dowla. 2008. Intonational phonology and focus prosody of Bengali. PhD. Thesis, UCLA, Los Angeles, CA.

Khan, Sameer ud Dowla. 2014. In S.-A. Jun, ed., *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Kubozono, Haruo. 2007. Focus and Intonation in Japanese. In S. Ishihara, ed., *Interdisciplinary Studies on Information Structure 9: Proceedings of the 2nd Workshop on Prosody, Syntax and Information Structure*, pages 1–27. Potsdam: University of Potsdam Press.

Mahajan, Anoop. 1990. The A/A-Bar distinction and movement theory. PhD. Thesis, MIT, Cambridge, MA.

Mahanta, Shakuntala. 2010. Prominence in Oriya, Bangla and Assamese: a phonetic and phonological investigation. *International Journal of Dravidian Linguistics* 33:101–28.

Ott, Dennis. 2012. *Local Instability: The Syntax of Split Topics*. Cambridge, MA: MIT Press.

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

Pierrehumbert, Janet. 1980. The phonology and phonetics of English intonation. PhD. Thesis, MIT, Cambridge, MA.

van Riemsdijk, Henk C. 1989. Movement and regeneration. In P. Benincà, ed., *Dialectal Variation and the Theory of Grammar*, pages 105–36. Dordrecht.

Selkirk, Elisabeth O. 2009. On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu* 136:35–73.

Selkirk, Elisabeth O. 2011. The syntax–phonology interface. In J. Goldsmith, J. Riggle, and A. Yu, eds., *The Handbook of Phonological Theory*, 2nd edition.

Simpson, Andrew and Arunima Choudhury. 2015. The non-uniform syntax of post-verbal elements in SOV languages: Hindi, Bangla and the rightward scrambling debate. *Linguistic Inquiry* 46:533–51.

Skopeteas, Stavros, Elisabeth Verhoeven, and Gisbert Fanselow. to appear. Discontinuous noun phrases in Yucatec Maya. *Journal of Linguistics*.

Twaha, Asimul Islam. 2017. Intonational phonology and focus in two varieties of Assamese. PhD. Thesis, IIT Guwahati, Guwahati, Assam.