

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 from 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.'	<i>Mere hoš o havās javāb de rahe hāī.</i> /ˈme:re: ˈho:ʃ o: həˈva:s ʒəˈva:b ɖe: ˈrəhe: hæ:/
2	بچپن سے انہوں نے اب تک پرکشش زندگی گزاری۔ 'From childhood, he has lived a luxurious life.'	<i>Bacpan se unhō ne ab tak purkašiš zindagī guzārī.</i> /ˈbətʃpən se: ˈunhō: ne: əb tək ˈpurkəʃi:ʃ ˈzindəgi: guˈzɑ:ri:/
3	اس کے پاؤں تھک گئے۔ 'His feet got tired.'	<i>Us ke pāō thak gae.</i> /ˈus ke: ˈpa:ō: ˈtʰək gæ:/
4	درویش کمرے سے نکل گیا۔ 'The saint went out of the room.'	<i>Darveš kamre se nikal gēā.</i> /dərˈve:ʃ ˈkəmre: se: ˈnikəl ˈgæa:/
5	نوجوانوں کو بھی پہرے کی ذمہ داری سونپی گئی۔ 'The young people were also given the responsibility of guarding.'	<i>Naujavanō ko bhī pēhre kī zimmedārī saumpī gāī.</i> /no:ʒəˈva:nō: ko: bʰi: ˈpəhre: ki: zimme:ˈɖɑ:ri: ˈsō:mpi: ˈgəi:/
6	اوکاڑہ ہمارا آبائی شہر ہے۔ 'Okara is our hometown.'	<i>Okārā hamārā ābāī šēhēr hai.</i> /o:ˈka:ɾɑ: həˈma:ra: a:ˈba:i: ˈʃəhær hæ:/
7	بشارت مت سماجت کرتے رہے۔ 'Basharat kept on insisting.'	<i>Bašārat minnat samājat karte rahe.</i> /bəˈʃɑ:rət ˈminnət səˈma:ʒət ˈkəɾte: ˈrəhe:/
8	اخوت نے مسلمانوں کو ایک قوم بنا دیا ہے۔ 'Brotherhood has made Muslims a nation.'	<i>Axuvvat ne musulmānō ko ek qaum banā diyā hai.</i> /əˈxuvvət ne: musulˈma:nō: ko: ek ˈqə:m ˈbənɑ: ˈɖijɑ: hæ:/
9	نماز تو وہ ہے جو ہم پڑھ رہے ہیں۔ 'The prayer is what we are offering.'	<i>Namāz to vo hai jo han paṛ rahe hāī.</i> /nəˈma:z to: vo: hæ: ʒo: həm pəɾ ˈrəhe: hæ:/
10	دعا میں گڑگڑانے والے کو اللہ محبوب رکھتا ہے۔ 'God loves the one who laments in prayer.'	<i>Duā mē giṛgiṛāne vāle ko Allāh mēhbūb rakhtā hai.</i> /ˈɖuɑ: mē: giɾgiˈɾɑ:ne: ˈva:le: ko: əlˈla:h məhˈbu:b ˈrəkʰtɑ: hæ:/
11	محبت ان کے لیے کافی ہے۔ 'Love is enough for them.'	<i>Muhabbat un ke lie kāfī hai.</i> /mʊˈhəbbət ˈun ke: ˈlie: ˈka:fi: hæ:/
12	ہماری ذمہ داریوں میں غیر معمولی اضافہ ہو گیا ہے۔ 'Our responsibilities have grown significantly.'	<i>Hamārī zimmedārīō mē ġair māmūlī izāfā ho gēā hai.</i> /həˈma:ri: zimme:ˈɖɑ:riō: mē: ɣæ:r ma:ˈmu:li: ɾˈzɑ:fa: ho: ˈgæa: hæ:/
13	ان کے ذہن بدل گئے۔ 'Their minds changed.'	<i>Un ke zēhēn badal gae.</i> /ˈun ke: ˈzəhæn ˈbədəl ˈgæ:/
14	ان کے دل کی شریانیں خراب ہو گئی تھیں۔ 'The arteries of his heart were damaged.'	<i>Un ke dil kī šaryānē xarāb ho gāī thī.</i> /un ke: ɖil ki: ʃərˈja:nē: xəˈra:b ho: ˈgəi: ˈtʰi:/
15	عذرانے نعیم کے منہ پر پانی چھڑکا۔ 'Azra sprayed water on Naeem's face.'	<i>Azrā ne Naīm ke mū par pānī chirka.</i> /ˈəzra: ne: ˈnəi:m ke: mū: pər ˈpa:ni: ˈtʃʰɪɾka:/

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	<p>نایاب نے دوکاندار سے لیموں مانگا۔ <i>Nāyāb ne dukāndār se līmū māṅgā.</i> /na:'ja:b ne: dʊ'ka:nɖɑ:r se: 'li:mū: 'ma:ŋga:/ 'Nayab asked for limes from the shopkeeper.'</p>	Declarative
2	<p>کس نے دوکاندار سے لیموں مانگا؟ <i>Kis ne dukāndār se līmū māṅgā?</i> /'kɪs ne: dʊ'ka:nɖɑ:r se: 'li:mū: 'ma:ŋga:/ 'Who asked for limes from the shopkeeper?'</p>	Wh-question
3	<p>کیا نایاب نے دوکاندار سے لیموں مانگا؟ <i>Kyā Nāyāb ne dukāndār se līmū māṅgā?</i> /kja: na:'ja:b ne: dʊ'ka:nɖɑ:r se: 'li:mū: 'ma:ŋga:/ 'Did Nayab ask for limes from the shop-keeper?'</p>	Yes/no question
4	<p>دوکاندار سے لیموں مانگو۔ <i>Dukāndār se līmū māṅgo.</i> /dʊ'ka:nɖɑ:r se: 'li:mū: 'ma:ŋgo:/ 'Ask for limes from the shopkeeper.'</p>	Semi-honorific imperative
5	<p>دوکاندار سے لیموں مانگیے۔ <i>Dukāndār se līmū māṅgie.</i> /dʊ'ka:nɖɑ:r se: 'li:mū: 'ma:ŋgie:/ 'Please ask for limes from the shopkeeper.'</p>	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.

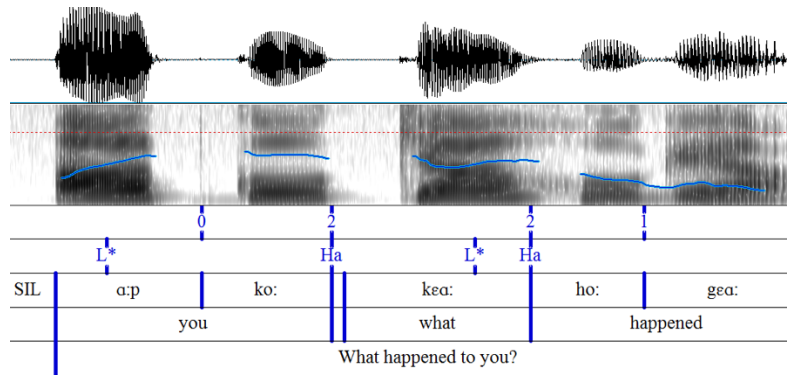


Figure 1. Example of an annotated utterance of Urdu

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).

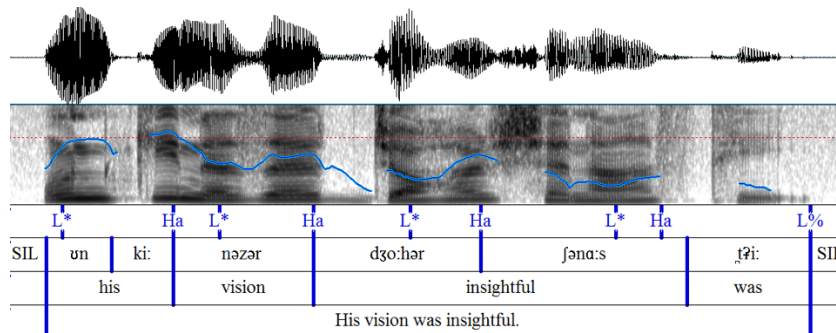


Figure 2. The words *un kī* /'ʊn k-i/ '3HON.OBL GEN-F', *nazar* /'nəzət/ 'vision', and each word in the compound *johar śanās* /'dʒo:hər ʃə'nəs/ 'insightful' bear rising APs composed of L* and Ha.

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* /nə'məz/ 'prayer' in Figure 3. AP-initial tones are also found in Korean (Jun, 2004).

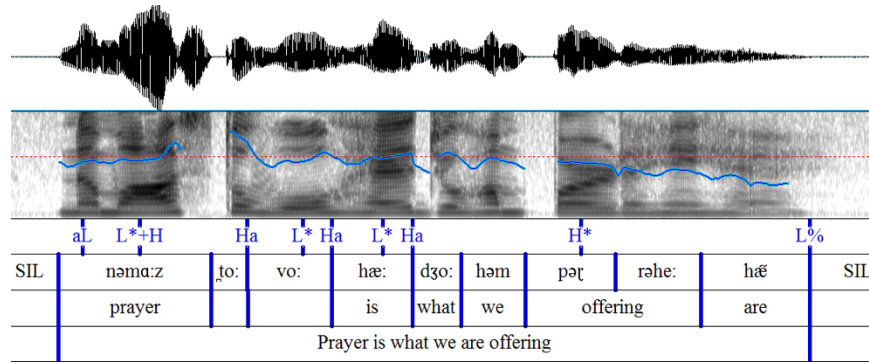


Figure 3. AP-initial aL tone on *namāz* /nə'mɑ:z/ 'prayer'

This L tone appears on the initial unstressed syllable of an AP, as shown on the words *namāz* /nə'mɑ:z/ 'prayer' in Figure 3 and *gīrgīrāne* /gɪrɡɪ'ɾɑ:ne/ 'lament' in Figure 4.

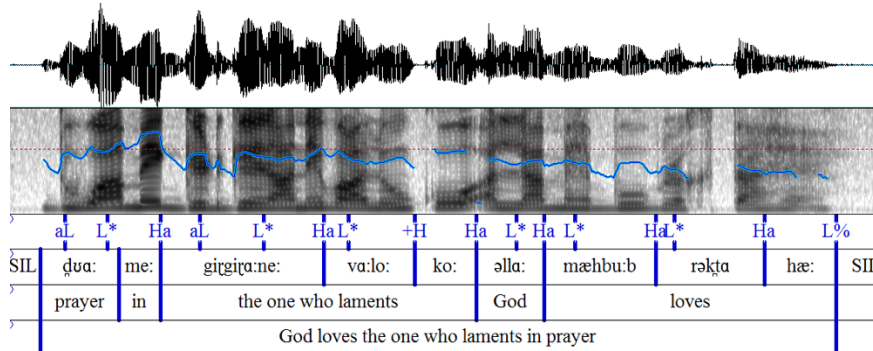


Figure 4. Two examples of the AP-initial aL tone

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* /'əzrɑ: ne:/ '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.

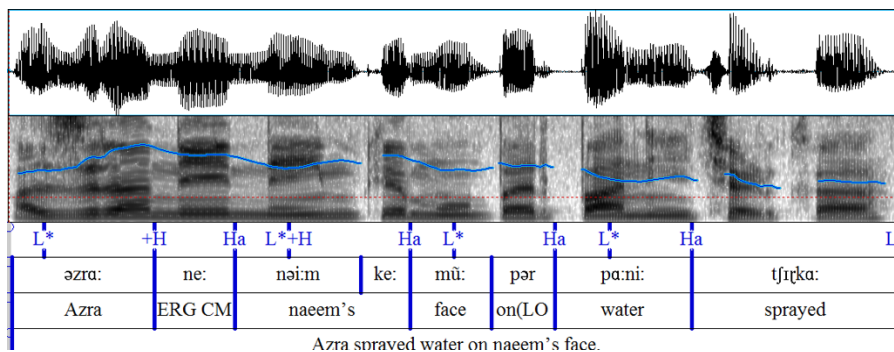


Figure 5. Two examples of a bitonal pitch accent L*+H followed by Ha

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.

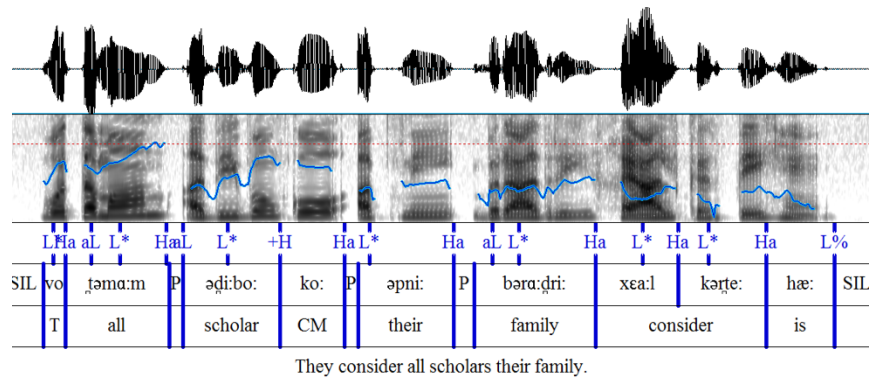


Figure 6. Fully realized AP tonal pattern in *adībo ko /ə'di:bo: ko:/* ‘scholars ACC’

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 /tʰək/* ‘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.

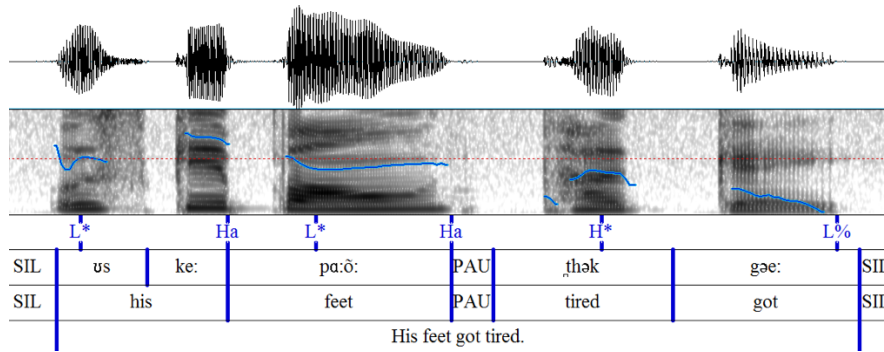


Figure 7. The less common AP tonal pattern H* L%

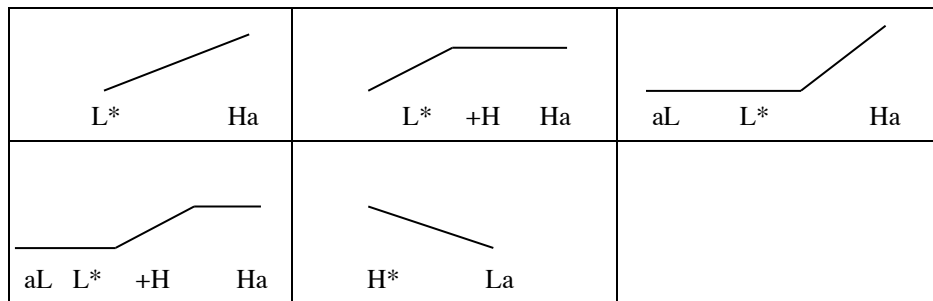


Figure 8. Schematic f0 contours of five type of APs

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 *sosāīfī* /soːˈsɑːiːfī/ ‘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* /ʔzrɑː/ ‘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 *mū* /mūː/ ‘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 mū par pānī chirkā*. /ʔzrɑː neː ˈnɑiːm keː mūː pər ˈpɑniː ˈtʃʰiːrkɑː/ ‘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	<i>Azrā</i> ʔzrɑː	<i>ne</i> neː	<i>Naīm</i> ˈnɑiːm	<i>ke</i> keː	<i>mū</i> mūː	<i>par</i> pər	<i>pānī</i> ˈpɑniː	<i>chirkā</i> ˈtʃʰiːrkɑː	
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 *ezāfe*), 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* /tʃʰɪʃm eː nɑm/ ‘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 /e:/* 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.

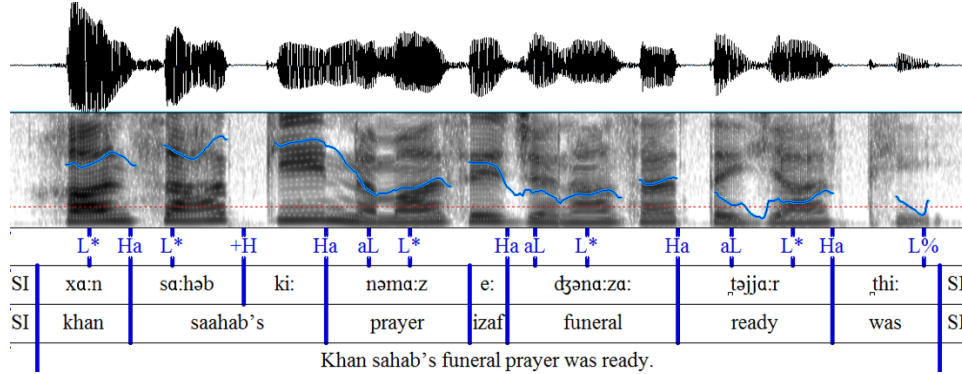


Figure 9. *Izāfat e /e:/* joining with the preceding word *namāz /nə'mɑ:z/* 'prayer'

4.1.2 Conjunctive *vāo*

In Urdu, the conjunctive *vāo* 'and' is pronounced as *o /o:/* is used to conjoin two nouns (Mangrio, 2016). Prosodically, the noun preceding 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 /ho:ʃ/* 'consciousness' to form one AP. The conjoined noun *havās /hə'vɑ:s/* 'senses' then forms a separate AP as shown in Figure 10.

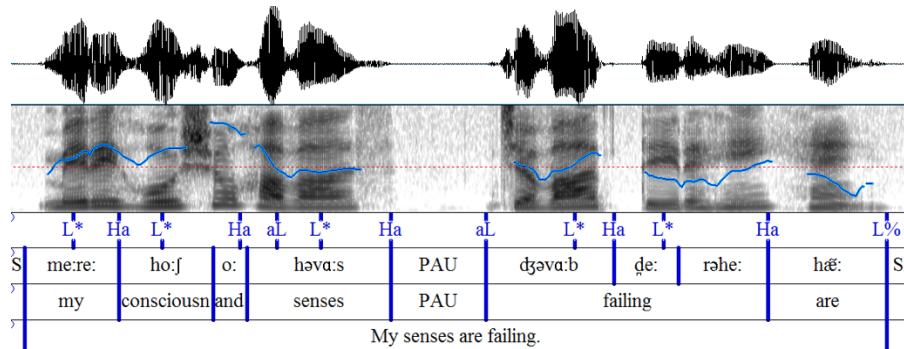


Figure 10. Conjunctive *vāo o /o:/* phrased with the preceding noun *hoš /ho:ʃ/* 'consciousness'

4.1.3 Case markers

Prosodically, case markers (e.g., *ne /ne:/* 'ERG', *ko /ko:/* 'ACC', *ke /ke:/* 'GEN', *mē /mē:/* '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 /ke:/* 'GEN' instead of the pronoun *us /us/* '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 mē* /'ɑ:grɑ: mē:/ '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.

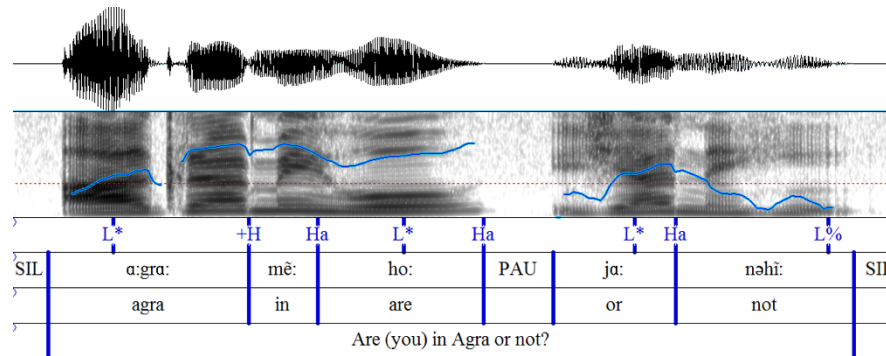


Figure 11. Noun and case marker phrased as one AP

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* /ke:/ 'GEN' takes low initial AP boundary aL, and a L* can be observed on first syllable of the postposition *liye* /'lije:/ 'for' as shown in Figure 12. The case marker and postposition are sometimes also joined orthographically.

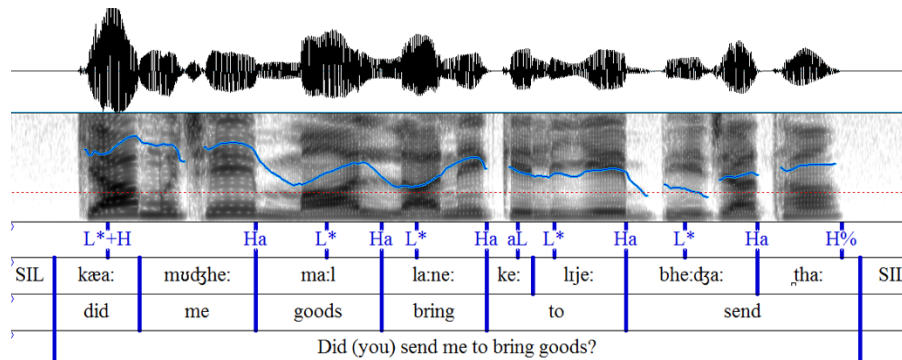


Figure 12. Complex postposition *ke liye* /ke: 'lije:/ 'GEN for' forming one AP

4.1.5 Verb auxiliaries

A complex verb forms a single AP. For example, the main verb *de* /ɖe:/ 'give' takes L* and the following auxiliary *rahe* /'rəhe:/ '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* /tʰi:/ ‘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.

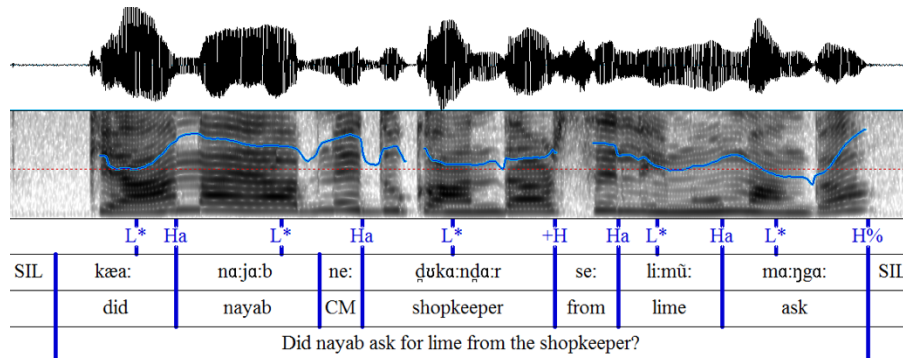


Figure 13. High IP boundary tone H%

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.

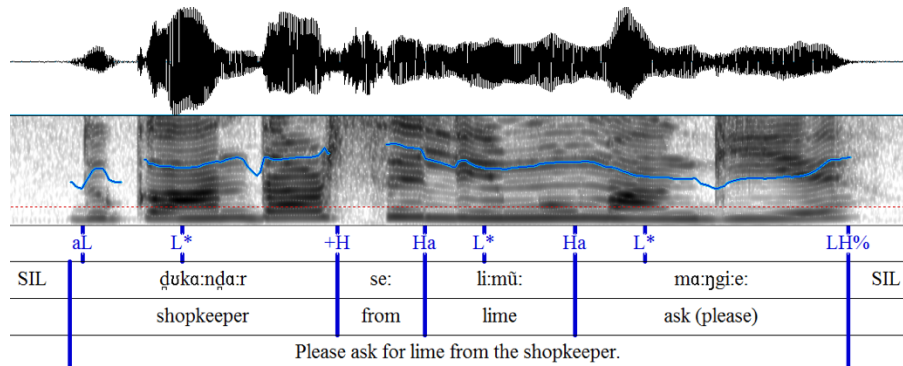


Figure 14. Rising IP boundary tone LH%

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.

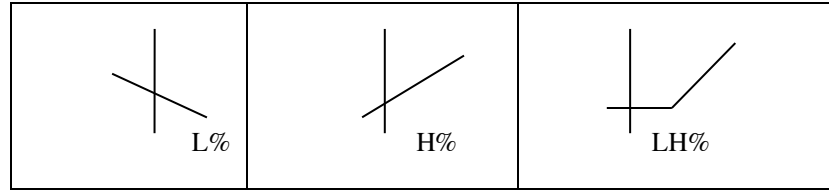


Figure 15: Schematic f0 contours of three types of IP boundary tones

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* /'va:qif/ ‘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.

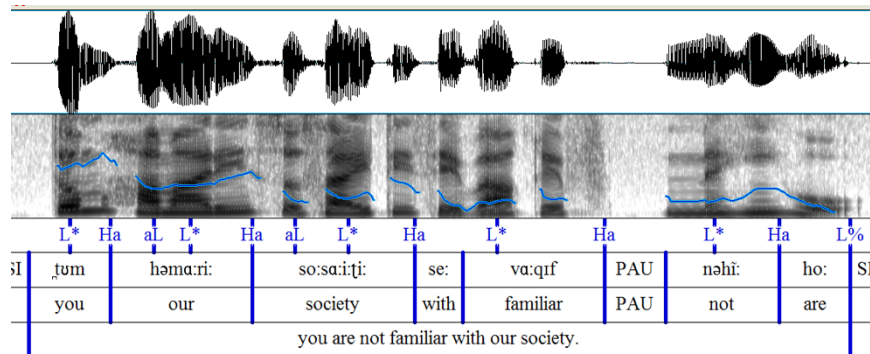


Figure 16. Absence of pitch reset

5 Sentence-specific prosody

To determine the differences in prosody for the various types of sentences, the sentence *Nāyāb ne dukāndār se līmū māngā* /nɑ:'jɑ:b ne: d̪u'ka:n̪ɑ:r se: 'li:m̪u: 'mɑ:ŋgɑ:/ ‘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.

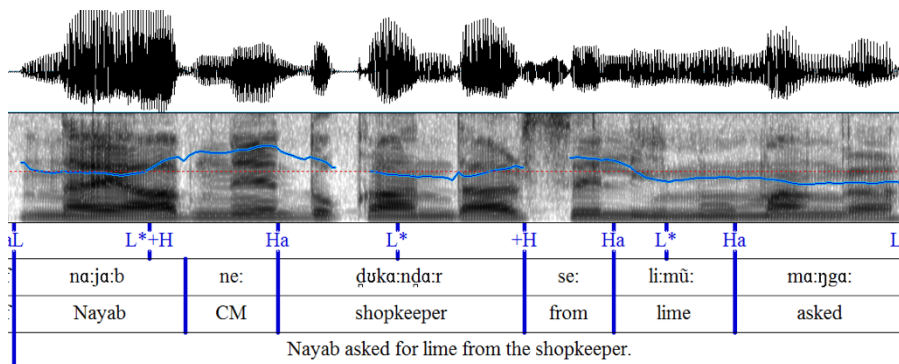


Figure 17. Pitch track of most frequent contour for declarative sentence

Table 6. Interspeaker tonal variation in declarative *Nāyāb ne dukāndār se līmū māngā* /na:'ja:b ne: ɖu'ka:ɳɖa:r se: 'li:mū: 'ma:ŋga:/ '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	<i>Nāyāb</i> na:'ja:b	<i>ne</i> ne:	<i>dukāndār</i> ɖu'ka:ɳɖa:r	<i>se</i> se:	<i>līmū</i> 'li:mū:	<i>māngā</i> 'ma:ŋga:		
Most frequent pattern¹	aL L*+H 4	Ha	L*+H 4	Ha	L* Ha 7	— 5	L% 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* /'kɪs ne:/ 'who ERG'

Speaker	f ₀ max:min (Hz)	f ₀ 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* /na:'ja:b ne:/ '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).

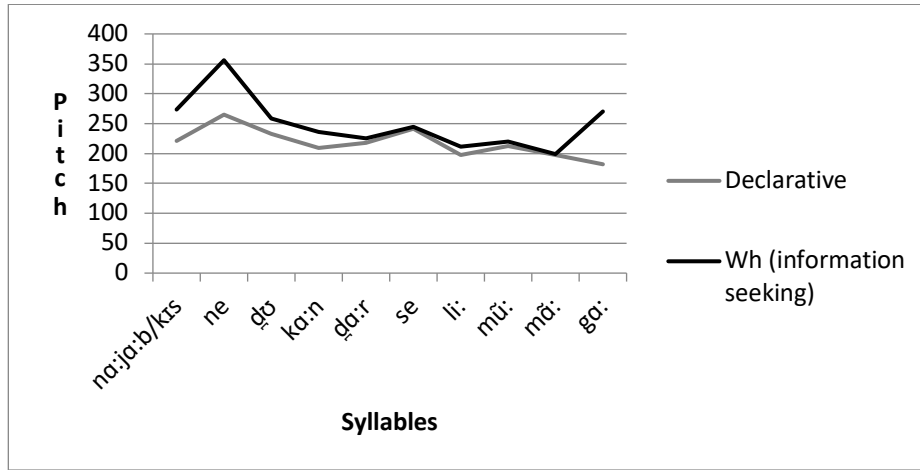


Figure 18. Declarative vs. wh-question uttered by the same speaker

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 līmū māngā* /'kis ne: du'ka:nda:r se: 'li:mū: 'ma:ŋga:/ '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	<i>kis</i> 'kis	<i>ne</i> ne:	<i>dukāndār</i> du'ka:nda:r	<i>se</i> se:	<i>līmū</i> 'li:mū:	<i>māngā</i> 'ma:ŋga:		
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

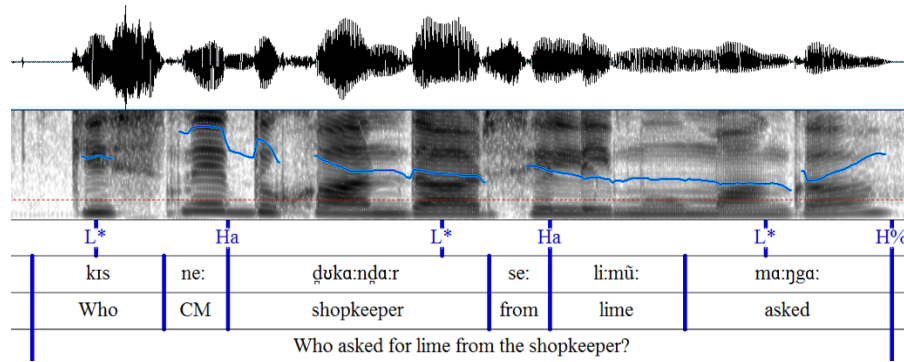


Figure 19. Pitch track of most frequent contour for wh-questions

5.3 Yes/no questions

Yes/no questions in Urdu are formed by adding an initial *kyā* /kja:/ (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 dukāṅḍār se līmū māṅgā?* /kja: na:’ja:b ne: ḍu’ka:ṅḍa:r se: ’li:mū: ’ma:ṅga:/ ‘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	<i>kyā</i> kja:	<i>Nāyāb</i> na:’ja:b	<i>ne</i> ne:	<i>dukāṅḍār</i> ḍu’ka:ṅḍa:r	<i>se</i> se:	<i>līmū</i> ’li:mū:	<i>māṅgā</i> ’ma:ṅga:		
Most frequent pattern	L* Ha 10	L* 5	Ha	L*+H 6	Ha	L* Ha 7	L* 8	H% 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

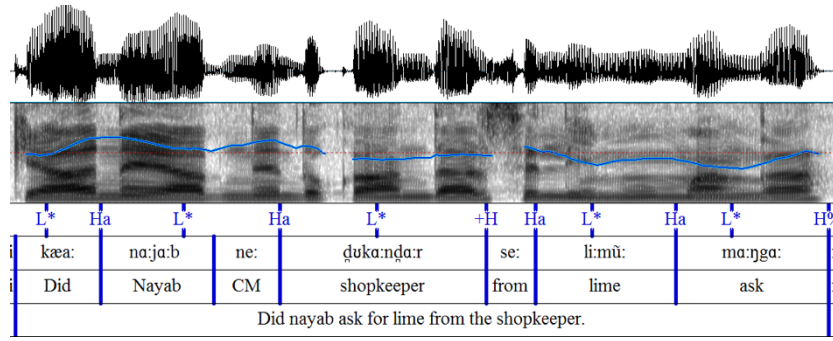


Figure 20. Pitch track of most frequent contour for Yes/no questions

5.4 Imperative sentences

In Urdu, imperatives have four forms: (i) non-honorific (e.g. *de* /d̪e:/ ‘give’), used with the singular non-honorific pronoun *tū* /tu:/, (ii) semi-honorific (e.g. *do* /d̪o:/), used with the plural and/or semi-honorific pronoun *tum* /t̪um/, (iii) subjunctive honorific (e.g. *dē* /d̪ē:/), used with the honorific pronoun *āp* /a:p/, and (iv) polite honorific (e.g. *d̪ij̪iye* /d̪i:d̪ʒje:/), also used with the honorific pronoun *āp* /a:p/, 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.

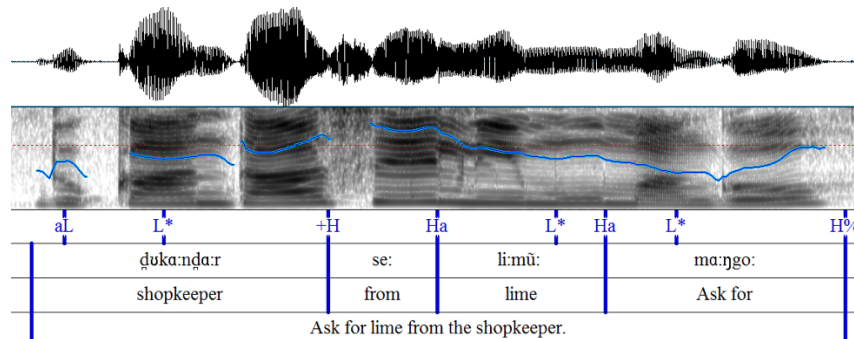


Figure 21. Pitch track of most frequent contour for semi-honorific imperatives

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 līmū māṅgo* /ḍu'ka:ṅḍa:r se: 'li:mū: 'ma:ṅgo:/ '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	<i>dukāndār</i> ḍu'ka:ṅḍa:r	<i>se</i> se:	<i>līmū</i> 'li:mū:	<i>māṅgo</i> 'ma:ṅgo:		
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 līmū māṅgie* /ḍu'ka:ṅḍa:r se: 'li:mū: 'ma:ṅgie:/ '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	<i>dukāndār</i> ḍu'ka:ṅḍa:r	<i>se</i> se:	<i>līmū</i> 'li:mū:	<i>māṅgie</i> 'ma:ṅgie:		
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

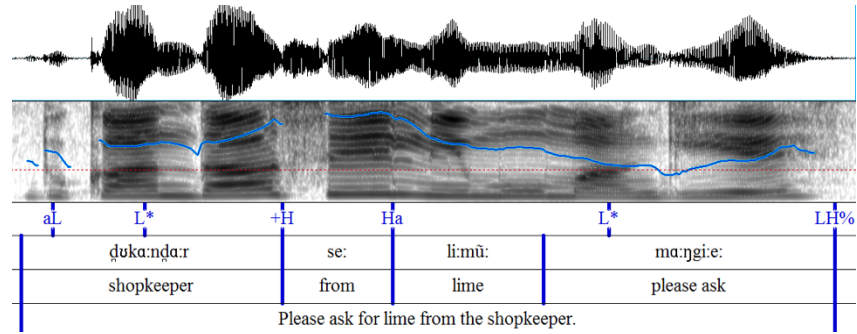


Figure 22. Pitch track of most frequent contour for polite honorific imperatives

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., & Vasisht, 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.