The Accentual Phrase in Urdu/Hindi: A prosodic unit at the interplay between rhythm and intonation

Farhat Jabeen[1], Elisabeth Delais-Roussarie[2]

[1]University of Konstanz, Germany
[2]University of Nantes, France

`farhat.jabeen@uni-konstanz.de, elisabeth.delais-roussarie@univ-nantes.fr`

Abstract

In Urdu/Hindi, the intonation of a sentence consists of a sequence of rising F0 contours (LH). These LH contours are the default intonation pattern associated with a basic prosodic unit which is analyzed either as a foot [1] or as a phonological phrase [2]. [1] , however, reported that a tonal pattern with two F0 rises (LHLH) was observed in morphologically complex words and syntactic phrases, but he did not provide an analysis to predict the occurrence of these double rises. The phonological status of these rises is not clear either. The current research addresses this issue and investigates the contexts in which the realization of double rises is observed. Using data from production experiments and speech corpus, we show that the LHLH pattern is not related to morphological complexity or morpheme type as these rises occur in monomorphemic words as well as in words with inflectional and derivational morphemes. We thus argue that the occurrence of these two rises is rhythmically constrained. Tonal alignment confirms such an analysis in which the first L and the last H tones are regular AP boundary tones, whereas the additional H and L tones are inserted for rhythmic reasons.

Index Terms: intonation, accentual phrase, rhythmic rise, phonology-syntax interface, rhythmic constraint, Urdu, Hindi

1. Introduction

An Accentual Phrase (AP) in Urdu/Hindi[1] minimally consists of a word although it frequently spans more than one word. For example, a noun with the following case marker and a verb followed by an auxiliary constitute an AP together by sharing one rising intonation contour (LH). The low and high tones occur respectively at the left and right edges of an AP [3, 4, 5] and are thus AP boundary tones. [5] reported that the alignment of low tone is fixed at the left edge of the AP but the alignment of high tone is flexible. It may align with either the last or the penultimate syllable of the AP. The alignment of these tones is not affected by the number of syllables and the position of the stressed syllable in the AP.

Both [1] and [6] observed that an AP in Hindi consisting of morphologically complex words may optionally carry two rising contours (LHLH). For instance, [1] stated that the word given in (1) may be produced with a single or a double rising contour. According to him, the selection between a single or

---

[1]Urdu is spoken mainly in Pakistan and Hindi is used in India. The main differences between these languages lie in their orthography and lexical borrowings. Urdu is written in Persio-Arabic script and borrows heavily from Persian and Arabic. Hindi uses Devanagari script and the bulk of its lexical borrowings comes from Sanskrit. In linguistic research, the observations made for one language are generally believed to be relevant for the other as well. We use the terms 'Hindi' or 'Urdu' while referring to the research based on that specific language. Otherwise, we use 'Urdu/Hindi.

a double rise is random and optional and does not change the meaning or the pragmatic import of an AP. Interestingly, morphologically complex words are never reported to have more than two rises even if they have more than two morphemes as shown in (1).

(1)

L H	+	L	+	H
hin.d̪ʊ	+	s.ʈɑː.n	+	i
Hindu	+	place	+	F.Sg
'Indian'				

Moreover, [1] also observed the appearance of double rises in nouns followed by case markers and verbs with the future tense morpheme *g(a,i,e)*. However, there is no existing analysis of the phonetic realization and phonological status of these double rises either in noun phrases or in polymorphemic words. We aim to fill this gap and address the following questions:

1. Is there a relationship between the occurrence of double rises and morphological complexity?

2. Do double rises form one long AP (LHLH)$_{AP}$, two iterative APs (LH)$_{AP}$(LH)$_{AP}$, or recursive APs?

In order to answer these questions, we used data from production experiments as well as from a corpus of Urdu speech. The first research question was addressed by investigating the intonation contour of monomorphemic and polymorphemic words. The phonological phrasing of APs with double rises in polymorphemic words and noun phrases is determined by referring to the alignment of tones with reference to the beginning of their target syllable. Further details are given in the next section.

2. Data set & analysis

2.1. Laboratory speech

The data for double rises in noun phrases (noun followed by case marker) was collected via production experiment. The target noun phrases were placed at initial position in declarative sentences produced in wide focus. Twenty speakers of Urdu were recorded for this experiment but not all speakers provided data for each noun phrase.

2.2. Speech corpus

All the examples of monomorphemic and morphologically complex words analyzed here were taken from One Hour Urdu Speech Corpus. This corpus was developed by the Center for Language Engineering, University of Engineering and Technology, Lahore, Pakistan. It consisted of written sentences read by a professionally trained female speaker of Urdu. The target words were extracted manually from this corpus and analyzed for their intonation contour and tone alignment.

2.3. Data analysis

We found in preliminary analysis that the tones did not consistently align with a particular syllable in the target word/phrase. Sometimes, the first rising contour aligned with the first morpheme and the second rise was produced on the following stem. However, it happened only in words with prefixes. The three examples presented below indicate no clear preference for aligning rising contours with a particular syllable or morpheme. This argues for an analysis of the alignment of tones within syllables in a word rather than with morpheme boundaries to determine the pattern of tone alignment.

(2)   L H   L H
       nə.ˈsi:.   hət̪   Monomorphemic
       'advice'

(3)   LH   L H
       ləʈ.   kʰə.ˈʈɑ. + t̪i   Derivational
       'staggering'

(4)   L   H   L H
       ˈmʊs.sə.rə.   t̪ + ẽ   Inflectional
       'pleasures'

Therefore, we measured the absolute alignment of a given tone from the beginning of the syllable that tone aligned with. In the next section, mean alignment (in seconds) of low and high tones in monomorphemic and polymorphemic words as well as in noun phrases is presented and discussed in detail.

3. Results

In order to confirm the relationship between double rises and morphological complexity, we analyzed the intonation contour of monomorphemic as well as polymorphemic words. The results of our analysis are as follows.

3.1. Polymorphemic words

Our analysis of words extracted from One Hour Urdu Speech Corpus confirmed the claims made by [1] regarding the production of double rises in morphologically complex words. These double rises appear frequently in words with derivational morphemes, both prefixes, and suffixes. An example is given in (5) and its F0 contour is presented in Figure 1.

(5)   LH   L H
       ˈbe   +   in.t̪ɑ.hɑ:
       without + limit
       'limitless'

Figure 1: *F0 of a polymorphemic word with a double rise.*

It is not clear from the example presented in (5) as well as from [1]'s data if words with inflectional morphemes are eligible to carry double rises as well. The lack of double rises in words

with inflectional morphemes could result from the fact that most of these morphemes in Urdu/Hindi are single vowels attached to a word to add information regarding number and gender. In this case, the lack of double rises in these words is not because inflectional morphemes are prosodically deficient in a way that derivational morphemes are not. This could simply happen because most of the words with inflectional morphemes are not long enough for the speakers to produce them with double rises. To resolve this, we investigated the intonation contour of long words with inflectional morphemes. (6) exemplifies this. The F0 contour of this word is presented in Figure 2. It shows that, just as with derivational morphemes, words with inflectional morphemes may also be produced with double rises. It also indicates that the production of double rises is not related to morpheme types.

(6)   L H   L H
       ʃər.ˈjɑ:.n + ẽ
       vein   + Pl.
       'veins'

Figure 2: *Double rise in a word with an inflectional morpheme.*

3.2. Monomorphemic words

As for the monomorphemic words, we found in the corpus examples of monomorphemic nominative nouns and adjectives located at non-final positions and realized with double rises. (7) is an example of a monomorphemic noun produced with a double F0 rise as shown in Figure 3. This shows that double rises are not specifically associated with morphological complexity.

(7)   L H   LH
       mʊ.ˈɑ:.ʃ.rɑ:
       'society'

Figure 3: *F0 of a monomorphemic word with a double rise.*

3.3. Size constraint in APs

Based on the discussion above, we propose that the extra rising contour in some APs is produced for rhythmic reasons. Our analysis of the appearance of double rises showed that the rhythmic rises appear when an AP consists of six or more moras. The

application of this size constraint explains why the words with the same derivational prefix *be* 'without' carry a single rise in Group A (Table 1) but have a double rise in Group B. The prefix *be* is produced with a short vowel in Group A but with a long vowel in Group B. This changes the mora count of words in each group which leads to the application of the size constraint in Group B but not in Group A.

Table 1: *Examples for the application of size constraint.*

	Word	**Meaning**	**Mora**
Group A	ˈbe + hәd	endless	$\mu + \mu\mu$
	be + ˈho.ʃ+ i	unconsciousness	$\mu + \mu\mu + \mu\mu$
Group B	be + tә.ˈhɑ:.ʃɑ	recklessly	$\mu\mu + \mu.\mu\mu.\mu\mu$
	be + ˈin.ta.hɑ:	limitless	$\mu\mu + \mu\mu.\mu.\mu\mu$
	dʊn.ja+ ˈdɑ:r	worldly	$\mu\mu.\mu\mu + \mu\mu\mu$

It could be argued that the insertion of the rhythmic rise is dependent on the foot or syllable structure. However, the second word in Group A *be + ho.ʃi* 'unconscious' and the last word *dʊn.ja+ da:r* 'worldly' in Table 1 both have three syllables. Yet the former has only one rising contour whereas the latter has two. As for the foot structure, both the words in (8) and (9) have three feet,[2] yet (8) has two rises but (9) has only one rising contour as Figure 4 illustrates. This indicates that neither foot structure nor syllable count can correctly predict if a word has a single or a double rise. Notice, however, that (8) has six moras ($\mu\mu + \mu\mu.\mu.\mu < \mu >$) whereas (9) has five moras ($\mu\mu.\mu.\mu\mu < \mu >$) as the last mora is extrametrical [8]. So the mora count based analysis correctly predicts the occurrence of double rises. This also shows that the tonal structure of an AP is weight sensitive and that the relevant feature to account for the realization of double rises is the mora, and not the syllable or foot.

(8) $\mu\mu + \mu\mu.\mu.\mu < \mu >$ Mora
(x .) + (x) (x .) $<$(x .)$>$ Foot
ˈbe + ˈin.ta.hɑ:
'limitless'

(9) $\mu\mu + \mu.\mu\mu < \mu >$ Mora
(x) + (x .) (x) $<$(x .)$>$ Foot
ˈnɑ: + gʊˈzi:r
'inevitable'

4. Phonetic alignment of tones

As mentioned above, it is not clear if the rhythmic tones are phonetically different from the default rising contour reported by [5]. In order to investigate this, we analyzed the alignment of tones in APs produced with single and double rises. In what follows, we present the results of the alignment of tones in single rises and double rises in APs consisting of polymorphemic words and noun phrases (nouns followed by case markers) separately. As it is corpus data and various variables were not controlled, we present the results of descriptive statistics only and report mean alignment of tones. The standard deviation is given in parentheses.

---

[2]According to [7], Hindi has bimoraic trochaic feet constructed from right to left. Degenerate feet, e.g. in (8) and (9) are allowed only at metrically strong positions. He analyzed the last foot as extrametrical.
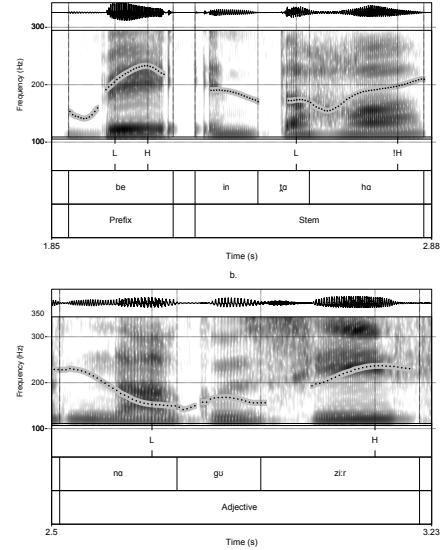

Figure 4: *F0 contour of polymorphemic words with single and double rises.*

4.1. Tone alignment: Polymorphemic words

4.1.1. Alignment of L tone

One cue to decide if rhythmic rises are different from default rises is the temporal alignment of low tone in polymorphemic APs with single and double rises. To do that, we measured the alignment of low tones with reference to the beginning of their target syllable (Beg_syll). There is variation in the syllable structure of prefixes and suffixes. Therefore, it is hypothesized that the alignment patterns do not result from a particular syllable structure.

The results presented in Table 2 show that in polymorphemic APs with double rises, the timing of both the low tones is different. The first low tone (L1) in the LHLH contour occurs later in the syllable as compared with the second low tone (L2). Furthermore, there is a lot of variance, shown by standard deviation, in the alignment of the first low tone indicating that it can occur quite late in the syllable. On the other hand, the variance in the realization of the second low tone is very small. The temporal realization of low tone in polymorphemic APs with a single rise is similar to that of L1 in the APs with double rises. This shows that the first low tone in both these instances is a phrase boundary tone marking the left edge of an AP.

Table 2: *Mean temporal alignment (sec.) of low tones in polymorphemic APs.*

	Tone	**Beg_syll. (sd)**
Double rises	L1	0.156 (0.33)
	L2	0.08 (0.03)
Single rise	L	0.131 (0.07)

4.1.2. Alignment of H tone

When a polymorphemic AP carries a single rising contour, the high tone has variable surface alignment [5]. Furthermore, there is variation in the absolute alignment of the first and second high tones with reference to the beginning of the syllable as shown in Table 3. The first high tone in the double rises is realized later

in the syllable as compared with the F0 peak in the following rise. Moreover, the first high tone has higher standard deviation as compared with the second peak indicating that the first high tone can be pushed even farther into the syllable.

Table 3: *Mean temporal alignment (sec.) of high tones in poly-morphemic APs.*

	Tone	Beg_syll. (sd)
Double rises	H1	0.234 (0.34)
	H2	0.151 (0.07)
Single rises	H	0.165 (0.07)

The last row in Table 3 shows the alignment of high tone in polymorphemic APs with a single rise. The alignment of this high tone is similar to that of H2 in double rises. This indicates that the last high tone in an AP with either single or double rises marks its right edge. It also shows that the rhythm induced high tone (H1) is realized later in the syllable and has higher variation in its alignment than the phrase edge marking high tone.

4.2. Tone alignment: Noun phrases

Apart from rhythmic rises produced in morphologically complex words, double rises are also observed in noun phrases produced in careful, slow speech in Urdu/Hindi [1]. For example, a phrase consisting of a noun followed by a case marker may optionally carry a double rise. A pertinent question is if the alignment of tones in the double rises in these noun phrases is similar to that of the double rises produced in morphologically complex words. In what follows, we compare the temporal alignment of tones in noun phrases with double rises with the tone alignment in morphologically complex words.

Table 4 presents the alignment of low and high tones in noun phrases when both the noun and the following case marker had a rising contour of their own. Compare the temporal alignment of tones in Table 4 with the alignment given in Tables 2 and 3. It is clear that the overall alignment of tones in noun phrases and polymorphemic words is not very different. The main difference appears in the alignment of high tones. The H2 is realized earlier in polymorphemic words in comparison with the noun phrases. However, as [5] has shown that the alignment of high tones in Urdu is variable, it was to be expected. The data presented here shows that the temporal alignment of LHLH tones in polymorphemic words is not different from the double rises in noun phrases.

Table 4: *Mean temporal alignment (sec.) of tones in noun phrases with double rises along with standard deviation (sd).*

L1 (sd)	H1 (sd)	L2 (sd)	H2 (sd)
0.12 (0.05)	0.186 (0.08)	0.09 (0.06)	0.197 (0.08)

5. Phonology of double rises

We have shown that the alignment of tones in single and double rises is different. According to this analysis, the first and the last tones in double rises are the same AP boundary tones (Lp Hp) as the ones found in APs with single rising contours. Despite the double rises, there are only two phrase boundary tones in these words and the extra tones are rhythmically induced. The appearance of the extra high (H1) and low (L2) tones can be explained in two ways.

(10) **a.** H1 and L2 are inserted between left and right edge marking AP boundaries.
b. This involves the application of an OCP constraint. The proposal then is that only H1 is inserted in APs with six or more moras due to the application of rhythm constraint. However, OCP forbids the appearance of two consecutive high tones within an AP and thereby leads to the insertion of L2 between two consecutive F0 peaks. According to this, only H1 is inserted in long words and L2 is merely sagging F0 between two high tones.

At this stage, none of these analyses is superior to the other, so we refrain from selecting one over the other.

6. Conclusion

The data presented here shows that the F0 contour of APs in Urdu/Hindi is a combination of phrasal and rhythm-induced tones. Moreover, the tonal contour of these APs is not affected by morphological complexity and is instead sensitive to mora count. These double rises tend to occur in slow speech and thus have stylistic import. Moreover, the analysis reported here shows that double rises in morphologically complex words and noun phrases do not differ in terms of tonal alignment. One reviewer pointed out that the rises in these two contexts may differ according to other acoustic/prosodic parameters. To follow up, we analyzed the scaling of F0 peaks as well as the steepness of F0 after the high tone in single and double rises in noun phrases and morphologically complex words and found no difference in terms of F0 peak scaling and degree of steepness in F0 on the basis of the intonation contour of APs.

The finding that noun phrases and morphologically complex words in Urdu/Hindi may carry double rises is in consonance with the occurrence of double rises in Tamil [9]. However, Keane had abstained from offering a definitive analysis of these rises. Our analysis of double rises here not only explains the data for Urdu/Hindi, it also provides a measure against which the double rises in Tamil can be analyzed.

7. Acknowledgements

8. References

[1] R. R. Moore, "A study of Hindi intonation," Ph.D. dissertation, University of Michigan, 1965.

[2] U. Patil, G. Kentner, A. Gollrad, F. Kügler, C. Féry, and S. Vasishth, "Focus, word order and intonation in Hindi," *Journal of South Asian Linguistics*, vol. 1, no. 1, pp. 55–72, 2008.

[3] C. Féry, "The intonation of Indian languages: an areal phenomenon," in *Festschrift for Ramakant Agnihotri*, I. Hasnain and S. Chaudhury, Eds. Akar publishers, 2010, pp. 288–312.

[4] C. Féry, *Intonation and prosodic structure*. Cambridge: Cambridge University Press, 2017.

[5] F. Jabeen, "Interpretation of LH intonation contour in Urdu/Hindi," in *Proceedings of International Congress of Phonetic Science*, Melbourne, 2019. [Online]. Available: https://www.icphs2019.org/

[6] J. D. Harnsberger, "Towards an intonational phonology of Hindi," Master's thesis, University of Florida, 1994.

[7] B. Hayes, *Metrical structure theory: Principles and case studies*. Chicago: The University of Chicago Press, 1995.

[8] S. Hussain, "Phonetic correlates of stress in Urdu," Ph.D. dissertation, Northwestern University, 1997.

[9] E. Keane, "The intonational phonology of Tamil," in *Prosodic Typology II: The Phonology of Intonation and Phrasing*, S.-A. Jun, Ed. Oxford: Oxford University Press, 2014, pp. 118–153.