Immediately Postverbal Questions in Urdu

Farhat Jabeen,[1] *Tina Bögel,*[1] *Miriam Butt*[1]

[1]University of Konstanz, Germany

`firstname.lastname@uni-konstanz.de`

Abstract

This production study investigates the interaction of prosody, word order and information structure with respect to *wh*-constituents in Urdu. We contrasted immediately preverbal *wh*-constituents with immediately postverbal ones. The preverbal position is the default focus position in Urdu; the appearance of *wh*-constituents in the immediately postverbal position within the verbal complex is not well understood. In order to test various possible factors governing the appearance of immediately postverbal *wh*-constituents, target sentences with *wh*-constituents in both pre- and postverbal positions were presented in different pragmatic contexts and given to native speakers to pronounce. The results show a clear difference in prosodic realization between the pre- and the postverbal position. The preverbal position is consistent with focus prosody, the postverbal *wh*-phrases appear to occur when the verb is in focus.

Index Terms: prosody-syntax interface, Urdu, focus, intonation, word order, *wh*-questions

1. Introduction

In this paper, we contribute towards an understanding of the variable placement of *wh*-constituents in Urdu.[1] We are currently engaged in laying the foundations for a Text-to-Speech (TTS) system for Urdu. Concrete target applications are question-answering (Q&A) systems (particularly in the area of eGovernance), foreign language learning and eBooks that can be read aloud via a TTS system. With respect to these, we are particularly interested in understanding Urdu intonation. There is very little previous research on Urdu prosody, in particular, there is next to no work on questions and on questions in interaction with the different word order possibilities.

Prior research on word order and questions in Urdu/Hindi has been of a primarily syntactic nature, with some consideration of semantic, pragmatic and information structural factors. Prosody has not been considered an important factor in understanding word order variation in Urdu/Hindi. In this paper, however, we show that prosody plays a central role in determining the placement of *wh*-constituents. In particular, it is the determining factor for the appearance of *wh*-constituents in the immediately postverbal position within the verbal complex.

---

[1]Urdu and Hindi are structurally so similar that researchers tend to refer to them as one language (e.g., variously Urdu/Hindi, Hindi/Urdu, Hindi-Urdu). There are small differences in phonemic inventory and derivational morphology. It is not clear to date whether there are major differences in terms of prosody. We believe there are not, but as our research has been conducted exclusively with Urdu speakers, we refer to the language as Urdu.

2. Background and Motivation

Urdu is an Indo-Aryan language spoken mainly in Pakistan. Its default sentence structure is Subject-Object-Verb but major constituents may scramble. Word order variation is correlated with the expression of information structure [1, 2, 3, 4]. The clause initial position in a sentence is generally the topic. But this position can also be used for scene setting (e.g., *With regard to the party, ...*). The immediately preverbal position in a sentence generally expresses focus. [3] further shows that the postverbal position performs many functions such as de-emphasizing, presenting afterthoughts, providing background information or adding emphasis to the new information, and to create suspense. This position is also reserved for phrasally "heavy" items.

While the interaction between word order and information structure is generally acknowledged, information structural concerns are not generally articulated with respect to the positioning of *wh*-phrases. We became interested in data as in (1) and (2), first discussed in detail in [5], which proposes that the postverbal positioning of the *wh*-constituent in (2) is the result of syntactic movement of the verb to the left for verb topicalization purposes. However, the clause internal position occupied by the verb is not one that has been associated with topicality (and [5] offers no independent evidence for this claim).

(1) sita=ne d^hyan=se **kıs=ko** dek^h-a t^h-a?
S.F=Erg carefully who.Obl=Acc see-Perf.M.Sg be.Past-M.Sg
'Who had Sita looked at carefully?'

(2) sita=ne d^hyan=se dek^h-a **kıs=ko** t^h-a?
S.F=Erg carefully see-Perf.M.Sg who.Obl=Acc be.Past-M.Sg
'Who had Sita looked at carefully?'

2.1. Syntax

Urdu/Hindi is generally characterized as a *wh-in-situ* language, meaning that the *wh*-element can remain in place, rather than being moved to the front of the clause as in English, for example. However, the default position for *wh*-constituents is actually the immediately preverbal position. Given that questions are generally associated with focus and that the immediately preverbal position is the default focus position in Urdu, the default appearance of *wh*-constituents in the immediately preverbal position is unsurprising.

The appearance of a *wh*-constituent in the immediately postverbal position and *within* the verbal complex, on the other hand, is surprising, particularly since non-*wh*-phrases cannot appear there, as illustrated in (3).[2] Otherwise, *wh*-questions have the same syntactic distribution as standard, non-*wh*-constituents [6]. The only other elements that may appear in

---

[2]A reviewer wonders what '*' stands for. As per standard linguistic notation, the '*' denotes that (3) is ungrammatical.

this position are the negative markers and the discourse clitics (*hi* 'only', *bʰi* 'also').

(3) *sita=ne dʰyan=se dekʰ-a **ram=ko** tʰ-a?
 S.F=Erg carefully see-Perf.M.Sg R.M=Acc be.Past-M.Sg
 'Sita had looked at Ram carefully?'

In [5] (2) is reported as a standard information-seeking question, but in [7] a rhetorical question reading is invoked. Under this reading, Sita had not really looked at anybody carefully and the speaker is exclaiming about this fact and no answer is expected to the question. Data as in (4) and (5) is further supportive of this hypothesis. Phrases containing a *which* generally demand an answer. As (5) shows, the *which* phrase is dispreferred (indicated by #) in the immediately postverbal position.

(4) sita [konsi ɣazal] ga-ti
 Sita.F.Nom which.M.Sg song.M.Sg.Nom sing-Impf.F.Sg
 rah-ti tʰ-i?
 stay-Impf.F.Sg be.Past-F.Sg
 'Which song (ghazal) did Sita keep singing?'

(5) #sita ga-ti [konsi ɣazal]
 Sita.F.Nom sing-Impf.F.Sg which.M.Sg song.M.Sg.Nom
 rah-ti tʰ-i?
 stay-Impf.F.Sg be.Past-F.Sg
 'Which song (ghazal) did Sita keep singing?'

In order to explain the postverbal positioning of the *wh*-phrase, a secondary focus position in the immediately postverbal position is postulated in [7] and comparative evidence from Romance [8] is adduced. [6] put forwards a different, yet related proposal in which a purely syntactic account posits rightward scrambling that is motivated by feature checking (the features pertain to echo questions and question formation in general).

All the approaches dealing with data as in (2) have concentrated on understanding the positioning of *wh*-constituents from a purely syntactic perspective, with some reference to information structure and semantics/pragmatics. However, when one attends to the intonation of utterances with the word order in (2), the verb appears to be stressed. This could be indicative of verb focus, a hypothesis that has so far not been pursued in the literature (though the verb topicalization analysis in [5] could be seen as working essentially along those lines).

2.2. Prosody

Of the very little work on intonation that exists for Urdu/Hindi, some has already investigated focus. [9] is an investigation of the relationship between intonation, word order and focus in Hindi. It concludes that in contrast to languages like English, focus in Hindi is realised by reduction in the post focal pitch contour rather than by the actual raising of the pitch on the focused element. In contrast, according to [10], focus in Hindi is realized by an increase in the pitch span and duration of the focused element. In [9] the focused element is defined as being the one that was produced in answer to a question, regardless of its position. Given that the immediately preverbal position is the default focus position, their experimental data may have included both default focus as well contrastive focus. [10] also looked at question-answer pairs in Hindi, however, it investigated the expression of contrast with respect to adjectives that were part of a constituent answering the question.

In [9] it is further stated that each content word (corresponding to a constituent in their data) forms a P-phrase in Hindi. Each P-phrase has a low tone and high phrase boundary at the right edge. This is also what has been found for in the genetically related South Asian language Bengali [11, 12]. Additionally a high tone is a marker of focus in Bengali [11, 12].

2.3. Outlook

Given the inconclusive nature of the findings of the previous studies on Urdu/Hindi, we decided to perform a production experiment. The interaction of word order, information structure and intonation in Urdu/Hindi is complex and the studies conducted so far have made important first steps in understanding the factors combining to produce a speech signal in a particular manner. However, many more experiments teasing apart various factors will most likely be necessary.

3. Experiment

The experiment sought to determine the conditions governing the immediately preverbal placement of *wh*-constituents vs. the immediately postverbal one. We decided to investigate the hypothesis put forward in [7], which implicates a pragmatic difference in that the postverbal position signals a non-canonical question — no answer is expected when the *wh*-constituent is immediately postverbal within the verbal complex.

We thus set up contexts which led either towards straight information-seeking questions or a rhetorical question interpretation. We presented both the pre- and the postverbal orders in each of these contexts, asked speakers to pronounce these sentences and analyzed their prosodic realization.

3.1. Participants

Seven male speakers of Urdu were recorded for this experiment. They were required to fill out a questionnaire about their linguistic background and other demographic details. They were aged between 23–30 years. All of them were natives of Pakistan and were enrolled in a Masters program at the University of Konstanz, Germany. They were paid a fee based on the standard rate for participating in recordings. All the participants were multilingual, but had learned Urdu as their mother tongue.

A note on this is in order. Urdu is mainly spoken in Pakistan and India, both of which are multilingual countries. Pakistan has 72 individual languages including the non-South Asian languages English, Arabic and Chinese [13]. Most of Pakistan's population can speak three languages including English, Urdu and at least one indigenous language. According to the Ethnologue, the L2 users of Urdu are far more in number (94,000,000) than the total native speakers (10,000,000). It is thus scarcely possible to find monolingual Urdu speakers. In fact, due to the high prestige of Urdu in Pakistan, there is a whole generation of Urdu speakers in Pakistan whose mother tongue is an indigenous language, but who use Urdu with their children as their first language. Consequently, the variety of Urdu used by younger speakers carries strong traces of those indigenous languages and the speakers of this indigenized Urdu outnumber the "pure" native speakers. Moreover, regardless of the linguistic background of their parents, these "First Generation Native Speakers" identify themselves as native speakers of Urdu. Our participants also belong to this category. We are aware that this multilingual background influences their language output but in studying South Asian languages it is near to impossible to avoid issues of multilinguality.

3.2. Methodology

A total of 56 sentences containing *wh*-constituents in either the immediately pre- or postverbal sentences were constructed. They follow contexts which are designed to prime the subject to interpret them as either information seeking or rhetorical ques-

tions. The subjects were presented with the contexts in English. With every context two questions, written in Urdu script, were presented. In the two questions, the *wh*-phrases were placed alternatively in pre- and postverbal position. The subjects were asked to select the sentence which seemed to best fit the context and to pronounce that sentence. Fillers were not added to the data set. The contexts were provided in English so as to avoid bias towards a particular linear order.

3.3. Material

We decided to use four different *wh*-phrases, *kıs=ne* 'who', *kıs=ko* 'whom', *kya* 'what', *kıs=se* 'with what/whom' for the experiment. These *wh*-phrases represent different case markings and the three most typical ones for verbal arguments. The bare *kya* 'what' is used extensively in Urdu and we included it to see whether bare *wh*-phrases showed different effects from the case marked ones. This turned out to be not the case.

The two sentence orders presented to the subjects are schematized in (6). An auxiliary was always included to ensure that the postverbal *wh*-constituent had to be interpreted as being within the verbal complex. Postverbal *wh*-constituents without the verbal complex receive a backgrounded interpretation (or one of the other interpretations detailed in section 2).

(6)   XP   WhP   V   AUX
        XP   V   WhP   AUX

We selected 7 verbs for the data set. All the verbs are high frequency disyllabic words. All the other words in the sentence were monosyllabic. All the sentences started with pronouns except one which began with a noun *kʰana* 'food'.

In order to avoid boundary effects, the target sentences included an interrogative and a follow-up, explanatory statement. An example is provided in (7) and (8). The follow-up, explanatory sentence followed both of the target versions immediately.

(7)   a.   **Context:** You tell a friend that you have been hunting. So your friend asks what game you killed.

      b.   ap=ne   kya mar-a   tʰ-a?
          you.Pol=Erg what kill-Perf.M.Sg be.Past-M.Sg
          'What did you kill?'   (**Target1**)

      c.   ap=ne   mar-a   kya tʰ-a?
          you.Pol=Erg kill-Perf.M.Sg what be.Past-M.Sg
          'What did you kill? '   (**Target 2**)

      d.   aj-kɑl   ʃikar=ka   mosam to   nɑhĩ hɛ
          nowadays hunting=Gen.M.Sg season Top not   be.Pres.3.Sg
          'It's not hunting season yet.'   (**Follow-up**)

(8)   a.   **Context:** You have been bragging about your hunting. So your brother taunts that you killed only a partridge.

      b.   ap=ne   kya mar-a   tʰ-a?
          you.Pol=Erg what kill-Perf.M.Sg be.Past-M.Sg
          'What did you kill?'   (**Target1**)

      c.   ap=ne   mar-a   kya tʰ-a?
          you.Pol=Erg kill-Perf.M.Sg what be.Past-M.Sg
          'What did you kill? '   (**Target 2**)

      d.   sırf ek titɑr
          only one partridge.M.Sg   (**Follow-up**)
          'Only one partridge.'

The context in (7) sets up an information-seeking question for which the preferred answer should be Target 1, according to [7]. The context in (8) sets up a non-canonical question, for which the preferred answer should be Target 2.

3.4. Task

The task was presented on MS PowerPoint and the sentences were randomly ordered. Participants were given verbal instructions and were also guided to repeat the sentences in case of stuttering, coughing or laughing. They were asked to be as natural as possible in their rendering of the sentences. While the exact nature of the experiment was not shared with the participants, it was made clear that the research was concerned with interrogatives. The experiment was divided into the following steps: reading the context silently, determining the more appropriate linear order in accordance with the context and then reading the corresponding sentence aloud. Every speaker pronounced 56 sentences containing seven different verbs and four *wh*-phrases. Each of them was pronounced in information seeking and rhetorical contexts by seven speakers. Thus a total of 392 items were recorded and analyzed. The duration of the whole familiarization and recording process was between 15–25 minutes. Some respondents took more time to read the contexts. There was no time constraint for them to read the context and pronounce the sentences. The sentences were recorded using a headset in a sound proof cabin.

3.5. Data analysis

The data was analyzed by various means. The selection of syntactic structures for information seeking and rhetorical contexts was documented simply by hearing the recordings. They were also visually evaluated using the spectrogram in PRAAT (http://www.fon.hum.uva.nl/praat/). The data was studied to measure the F_0 in the middle of the vowel in every syllable of the target sentences. The overall duration of all syllables in the sentences was also measured.

The statistical analysis calculated the fundamental frequency and the duration with a linear mixed effects regression model (LMER) with the type (pre-/postverbal) as fixed factors and subjects and items as random factors on the basis of 380 sentences. 12 sentences were discarded due to errors in the data points.[3]

3.6. Results

The results show that the prosodic realization of the verb and the *wh*-phrases differs significantly on the basis of their position in the sentence. However, the results also indicate that the pre- and postverbal positions of the *wh*-phrases are not associated with either information seeking or non-canonical question context. Neither pitch nor duration show any significant correlation with the change in context.

The contours in Figures 1 and 2 represent the mean pitch track for *wh*-phrases at pre- and postverbal positions respectively. Each condition is averaged over seven speakers and normalized over time. F_0 has been analyzed syllable-wise with the single syllables encoded on the y-axis. The crucial parts are represented by V_S1/V_S2 which refers to the two syllables of the verb. Additionally, WH represents the position of *wh*-phrase.

Figure 1 shows a clear rise in pitch for the *wh*-phrase in the preverbal condition and a following deaccentuation of the verb and the auxiliary. In contrast, as Figure 2 exemplifies, the *wh*-phrase is deaccented in the postverbal position with the verb receiving main stress.

---

[3]Reviewers wondered whether we normalized the data (duration of syllables, F_0, speaking rate). Our statistical model treated speakers and items as random factors, thus anticipating variation across speakers.
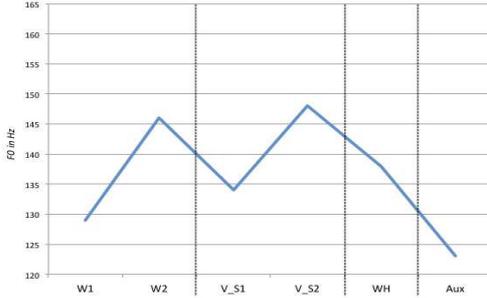
Figure 1: *The pitch contour of preverbal constructions.*

Figure 2: *The pitch contour of postverbal constructions.*

For the *wh*-phrase in the preverbal position, F_0 is significantly higher than its deaccented counterpart (β=12.68, SE=3.07, t=4.13). In contrast, when the *wh*-phrase is placed in the postverbal position, F_0 is significantly low. Instead, the second syllable of the verb receives higher pitch (V_S2, β= 22.45, SE=4.76, t=4.72). This difference is marked if compared to the pitch of the verb following a *wh*-phrase. Furthermore, the duration of this syllable is significantly longer when compared to the verb with the *wh*-phrase in the preverbal position (β=11.49, SE 2.7, t=4.26).

4. Discussion

The results show no significant correlation of immediately pre- vs. postverbal position of the *wh*-phrases with information seeking vs. non-canonical question interpretation. However, we do find a significant difference in the prosodic realization of the *wh*-questions depending on their position in the clause. We here suggest that our data can be understood if one hypothesizes that the verb is actually in focus when the *wh*-phrase is immediately postverbal within the verbal complex.

When the *wh*-phrase is preverbal, as in Figure 1, it is in focus. This focus is determined by a higher F_0 indicating a phonological phrase (PhP) boundary triggered by the focus.

(9) XP XP WH)$_{PhP}$ V AUX

This analysis is in accordance with [11, 12], which shows that there is an obligatory P-phrase boundary following a narrowly focused constituent. This would be the case in (9), whereby focus is realized via an H tone marked by longer duration and higher F_0. As in Bengali, the H tone is aligned with the right edge so that it appears on the second syllable of the verb.

The pitch contour when the *wh*-phrase is in the immediately postverbal position suggests that there is a prosodic boundary

following the verb. This would then lead to a different prosodic phrasing as compared to the one in (9).

(10) XP XP V)$_{PhP}$ WH AUX.

The drop in F_0 between *wh*-phrase and verb is considerably greater in the preverbal condition. However, while this difference in prosodic phrasing on the basis of position is also supported by durational cues, there was no significant difference in duration found for different *wh*-phrases.

The phrasing in (10) and the pitch contour is consistent with an analysis in which the verb is in focus. As per [11, 12] for Bengali and [14] for Chicheŵa, focus is assumed to trigger a P-phrase boundary at the right edge (also see [15]).

We further propose that the *wh*-phrase cannot be preverbal when the verb is in focus. This is because the verb cannot be focused via an H tone when the *wh*-phrase is preverbal. In the immediately preverbal position, the *wh*-phrase is correlated with question focus and demands the H tone (cf. [16] for Japanese) and a prosodic boundary. This makes it difficult to associate the verb with an H tone expressing focus. The solution in the language appears to be an undertaking of syntactic repositioning in order to accommodate prosodic constraints.

When the verb is in focus, the *wh*-phrase appears in the postverbal position. As a consequence, the immediately postverbal position appears to be defocused. This could account for the infelicity of the *which*-phrases in (5).

We further hypothesize that the *wh*-phrase appears in exactly the immediately postverbal position within the verbal complex because all the other positions in the clause are already associated with other information-structural notions (e.g, topic, backgrounding). The immediately postverbal position is also the first place that a *wh*-phrase can occur in terms of a prosodic landing site after the P-phrase that has been closed off after the verb (reflecting a type of prosodic inversion).

5. Conclusions

Analyses seeking to explain the word order variability in questions have so far been primarily syntactic in nature. We suggest that prosodic concerns are the crucial factor in determining the placement of *wh*-phrases in the immediately postverbal position within the verbal complex. This occurs when the verb is in focus and needs to attract the H tone that would otherwise be associated with the immediately preverbal focus position.

Our data indicate that focus is associated with increased pitch and duration in Urdu. This is consonant with [10], but not with the findings of [17]. However, each of the three studies (ours, [10, 17]) have concentrated on different aspects of the expression of focus in Hindi/Urdu. For one, the data differs considerably. For another, it is becoming clearer with each study that the interacting set of factors is complex and must be teased apart carefully. We are encouraged by the fact that our data is consonant with crosslinguistic findings. As far as we are aware, however, the type of prosodically triggered inversion with *wh*-phrases reported here has not been discussed elsewhere.

6. Acknowledgements

7. References

[1] M. Butt and T. H. King, "Structural topic and focus without movement," in *Proceedings of the First LFG Conference*, M. Butt and T. H. King, Eds. Stanford: CSLI Publications, 1996.

[2] ——, "Null elements in discourse structure," 1997, written to be part of a volume that never materialized, http://ling.uni-konstanz.de/pages/home/butt/main/papers/nulls97.pdf. [Online]. Available: http://ling.uni-konstanz.de/pages/home/butt/

[3] V. Gambhir, "Syntactic restrictions and discourse functions of word order in standard Hindi," Ph.D. dissertation, University of Pennsylvania, Philadelphia, 1981.

[4] A. Kidwai, *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. Oxford: Oxford University Press, 2000.

[5] R. Bhatt and V. Dayal, "Rightward scrambling as rightward movement," *Linguistics Inquiry*, vol. 38, no. 2, pp. 287–301, 2007.

[6] E. Manetta, "Reconsidering rightward scrambling: Postverbal constituents in Hindi-Urdu," *Linguistic Inquiry*, vol. 43, no. 1, pp. 43–74, 2012.

[7] M. Butt, "Questions and information structure in Urdu/Hindi," in *Proceedings of the LFG14 Conference*, M. Butt and T. H. King, Eds. Stanford: CSLI Publications, 2014.

[8] M. L. Zubizarreta, *Prosody, Focus and Word Order*. Cambridge, MA: The MIT Press, 1998.

[9] 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.

[10] S. Genzel and F. Kügler, "The prosodic expression of contrast in Hindi," in *The Proceedings of Speech Prosody 2010*, Chicago, 2010.

[11] B. Hayes and A. Lahiri, "Bengali intonational phonology," *Natural Language and Linguistic Theory*, vol. 9, pp. 47–96, 1991.

[12] A. Lahiri and J. Fitzpatrick-Cole, "Emphatic clitics and focus intonation in Bengali," in *Phrasal Phonology*, R. Kager and W. Zonneveld, Eds. Dordrecht: Foris Publications, 1999, pp. 119–144.

[13] G. F. S. Lewis, M. Paul and C. D. Fennig, Eds., *Ethnologue: Languages of the World*. Dallas: SIL International, 2015, eighteenth edition, Online version: http://www.ethnologue.com.

[14] H. Truckenbrodt, "The syntax-phonology interface," in *The Cambridge Handbook of Phonology*, P. de Lacy, Ed. Cambridge: Cambridge University Press, 2007, pp. 435–456.

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

[16] H. Truckenbrodt, "An analysis of prosodic f-effects in interrogatives: Prosody, syntax and semantics," *Lingua*, vol. 124, pp. 131–175, 2012.

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