Lexical Stress in Urdu

Benazir Mumtaz, Tina Bögel, Miriam Butt

University of Konstanz

firstname.lastname@uni-konstanz.de

Abstract

This study looks at the role of lexical stress in Urdu/Hindi prosody. The literature on lexical stress is divided, with some authors developing algorithms for stress assignment, while others deny its relevance for prosody. We performed three experiments to investigate this issue and find evidence that a strong increase in the duration of a syllable indicates stress and that lexical stress and phrasal intonation interact in a non-trivial manner. We also find that stress perception varies according to syllable weight with weight clash being a determining factor.

Index Terms: Urdu/Hindi, lexical stress, duration, pitch increase, intonation, phrase languages, weight clash

1. Introduction

As part of a cooperative project, we are interested in developing a Text-to-Speech (TTS) system for Urdu.[1] Natural sounding TTS needs information about speech prosody [1], including lexical stress [2]. However, there is still comparatively little work on Urdu/Hindi[2] prosody and the literature is contradictory with reference to lexical stress, with some papers claiming that lexical stress does not play a role in Urdu/Hindi [3, 4]. We present three experiments designed to shed light on the issue of lexical stress in Urdu: two production experiments that probe the acoustic correlates of lexical stress and one perception experiment. Our results support the existence of lexical stress, but find that issues of phrasal intonation and syllable weight are complicating factors in the production and perception of stress.

2. Urdu Lexical Stress and Intonation

2.1. Intonation

Going back to at least [5], the consensus is that the basic prosodic structure of an Urdu/Hindi clause consists of a series of LH contours with a fall on the last constituent [6, 7, 8]. However, the precise nature and distribution of the LH contours remains the subject of debate. [5] lists three possibilities for the realization of the basic LH: 1) a bitonal pitch accent; 2) an L* pitch accent followed by an H boundary tone; 3) an LH accentual phrase. [3] surveys two Indo-Aryan (Hindi and Bangla) and two Dravidian languages (Tamil and Malayalam) and offers a new class in the typological space of intonational systems: "phrase languages". These phrase languages are characterized by a phrasal accent which determines the prosodic phrasing, rather than pitch accents or lexical stress. The association of L and H with syllables is therefore predicted to be variable, but continues to be tied to phrasing (L_p and H_p). There is consequently no predetermined syllable to receive the pitch accent. [3] posits this analysis in part because of an inability to find evidence for lexical stress in these languages. However, [9] show that definitive phonological cues can be adduced for the identification of lexical stress in Bangla: stress can only be word initial, nasal vowels have to be stressed and geminates can only exist in stressed syllables, see also [10].

2.2. Lexical Stress in Urdu

The literature on lexical stress in Urdu/Hindi is controversial. [3] argues that native speakers' inconsistent intuitions about lexical stress speak against the existence of lexical stress. On the other hand, while [11], [12], and [13] show variation in the placement of stress in a Hindi word, the existence of lexical stress per se in Urdu/Hindi is not denied. Rather they assume that Urdu/Hindi has lexical stress and posit that syllable weight plays an important role in stress assignment. However, there is disagreement as to the exact location of lexical stress.

We follow the algorithm developed in [14], according to which stress in Urdu/Hindi is sensitive to syllable weight. Stress falls on the first non-light syllable starting from the end of the word. However, word final moras are analyzed as extrametrical and so final heavy syllables cannot carry stress. Finally, the root morpheme takes primary stress whereas affixes take secondary stress. For other, slightly different, analyses of stress assignment see [11], [12], [13], and [15].

The examples in (1) from [13] and [14] illustrate that both strong-weak and weak-strong variants of a word are possible in Urdu/Hindi. These examples support the argument by [3] about inconsistent intuitions of speakers with regard to lexical stress.[3]

(1) [ˈkəməl]∼[kəˈməl] [ˈsɑːlɑːnɑː]∼[sɑːˈlɑːnɑː]
[ˈɪbrɑni]∼[ɪbˈrɑni] [ˈpeʃɑːniː]∼[peˈʃɑːniː]

Most of the work on Urdu/Hindi stress has been based on introspective data and/or on evidence from poetic meter [17]. Work dealing with the acoustic correlates of stress in Urdu/Hindi are [4], [17] and [18]. [4] concludes that there is no significant difference in the duration of stressed and unstressed vowels and that the stressed syllable has a rising pitch followed by a fall. On the other hand, experiments conducted by [17] show that stressed syllables have a longer duration and a low f0, see also [18]. [17] also notes that [4] could not find the impact of duration because the stressed and unstressed syllables chosen in his experiment were not fully matched in their segmental structure.

Another argument brought to bear by [3] against lexical stress in South Asian languages is perceived misalignments between the low tone and the stressed syllable in Bangla. This leads her to reanalyze Bangla — a language prosodically very similiar to Hindi/Urdu — as a phrase language. In our own work on developing an annotated speech corpus for Urdu [19],

[1]https://tech.cle.org.pk/services/speech/tts

[2]Urdu and Hindi are close variants of one another, the major difference lies in the script and in Perseo-Arabic borrowings (Urdu) vs. Sanskrit (Hindi). The variants are completely mutually intelligible.

[3][16] also points out the variation in Hindi stress but attributes the variation to the fact that Hindi is spoken by a very large population who also speak some other South Asian language as their mother tongue.

we have also not observed a completely straightforward alignment pattern between the L (or H) tone and the stressed syllable. However, this might be due to a number of factors, including effects of derivational morphology and cliticization. This paper therefore aims to adduce additional experimental evidence to help resolve the question of lexical stress in Urdu.

3. Experiment 1: Matched Segments

3.1. Methods

3.1.1. Materials

To address the issue of differences in segmental material in the realization of stress as raised by [17], we selected eight pairs of frequent Urdu words. The pairs consist of identical segmental material in the target syllable and in the onset of the following syllable. The rest of the word contains similar segmental material. The target syllable is stressed in the first member of a pair, and unstressed in the second member, see Table 1.

Table 1: *Word pairs with differing stress in target syllable (bold)*

Stressed	Unstressed	Translation
ˈ**lo**.hɑ	**lo**.ˈhɑr	iron, blacksmith
ˈ**mən**.ʃɑ	**mən**.ˈʃur	aim, manifesto
ˈ**mə**.ri	**mə**.ˈriz	died, patient
ˈ**əd̪**.rək	**əd̪**.ˈrɑk	ginger, awareness
ˈ**mə**.t̪i	**mə**.ˈt̪in	obedient, soft spoken
lɑ.ˈ**gɑ**.t̪ɑ	lɑ.**gɑ**.ˈt̪ɑr	to add, continuously
mə.ˈ**zɑ**.lɪm	mə.**zɑ**.ˈmin	cruelty, subjects
ˈ**xə**.t̪ə.rɑ	**xə**.t̪ə.ˈrɑt̪	danger, dangers

3.1.2. Procedure

Participants were asked to produce the target words in two settings: 1) in isolation; 2) in the carrier sentence in (2).

(2) ɪs ləfz kɑ: mət̪ləb ___ ho: gɑ:
‘The meaning of this word will be ___.’

Randomized target words and sentences were interspersed with fillers and presented as a slide presentation in Urdu script. Participants were recorded in an anechoic chamber with a sampling rate of 48 kHz.

3.1.3. Participants

Ten Pakistani Urdu speakers (∅ = 22.3 years, SD = 2.40 years, 6 female) participated in the experiment and were given a small payment. All participants live in Lahore and also speak the regional language Punjabi. None had hearing/speaking disorders.

3.1.4. Data treatment

In total, we collected 480 productions (2 contexts x 16 words x 10 speakers: 5 recorded the target words twice and 5 only once). Of the 480, 40 items were rejected due to mispronunciation and glottalization. The target word was annotated using the standard annotation criteria [20] and the annotation software Praat [21]. The data were annotated via the guidelines established by [22].

Lexical stress was automatically annotated using the algorithm from [14] via a tool developed as part of our cooperative project. A further tier records lexical stress as perceived by our research team. Only examples with unanimous labelling were used for the further evaluation: a final count of 306 files. Figure 1 shows an example of the Praat annotation of a target word.

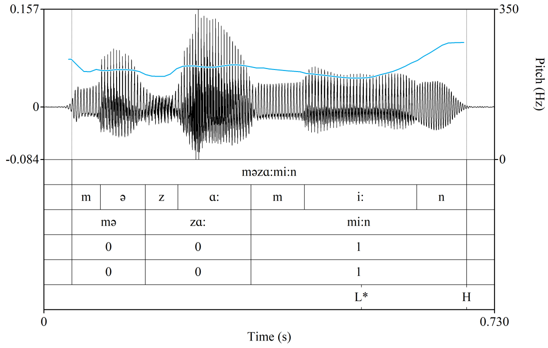

Figure 1: *Wave file and pitch contour with annotation. The tiers show the word in IPA (1), its segments (2), syllables (3), stress marked using perception (4), stress via the algorithm in [14] (5), and tone (6) in the target word [məzɑmin] ‘subjects’.*

[17] measured syllable duration and formant structure of vowels to explore the acoustic evidence for lexical stress in Hindi with respect to 3 participants. We aimed at checking the results with a larger sample size. Our data was analyzed with linear mixed effect regression (lmer) models with the stress pattern (±stress) as fixed factor and participants and syllables as crossed random factors (adjustment of intercepts and slopes). P-values were calculated using the Satterthwaite approximation.

3.2. Results

3.2.1. Syllable duration

The statistical analysis showed a strong increase in the duration of the stressed syllable (β= -0.04, SE = 0.005, t = -7.09, $p < 0.001$) as illustrated in Figure 2. Duration was also highly significant when isolated words and words pronounced in a carrier sentence were analysed individually, but the results did not differ greatly from the overall results found for duration.

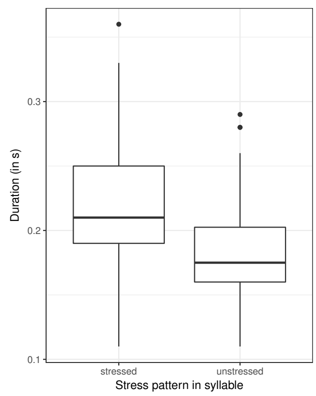

Figure 2: *Duration in stressed and unstressed syllable pairs*

3.2.2. Fundamental frequency

Another effect was found with the fundamental frequency (f0), which we converted into semitones (ST) before the analysis. The mean ST was significantly lower in the stressed syllable (β=0.66, SE=0.14, t= 4.8, $p < 0.001$) compared to the unstressed syllable. This effect was much more pronounced with the isolated words in comparison to the target words in the carrier sentences, where the effect was still significant, but less so (β = 0.61, SE = 0.24, t = 2.5, $p < 0.05$).

For a more detailed overview of the f0 patterns, we took all word pairs where the target syllable was directly followed by the syllable with the alternate stress (7 word pairs, 268 recordings) and extracted five evenly distributed f0 measurements per syllable. The resulting ten data points were converted into ST and the mean values for each point were calculated across all speakers and word pairs. Of these, the first two points — covering the onset of the first syllable (S1) — were discarded, as there was often no discernible pitch and the extracted f0 values showed great variation. The average pitch in ST is shown in Figure 3 for isolated words and in Figure 4 for the words in the carrier sentence, starting from the third data point (S1P3).

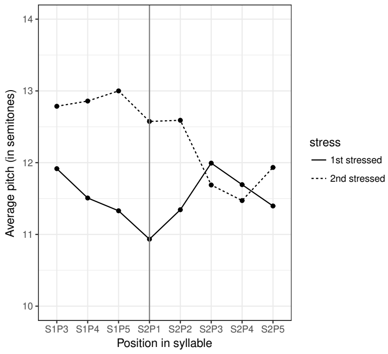

Figure 3: *Pitch in isolated words. The first syllable (S1) is target. The second syllable (S2) is the stress alternate.*

If the target S1 is stressed in isolated words, it shows a strong L towards its end, followed by a strong rise in the second syllable (S2) (Figure 3). If the target syllable is unstressed and S2 is stressed, the pitch remains high and only falls to a (less pronounced) L within S2. In carrier sentences (Figure 4), the target syllable still receives a low tone if stressed, but the fall in pitch is less pronounced. The following rise to the S2, however, is much stronger compared to the rise in the isolated word. If the target syllable is unstressed, there is only a slight lowering in the pitch. The low pitch continues in the stressed S2 before it rises towards the end.

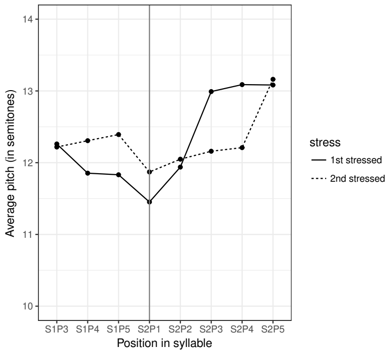

Figure 4: *Pitch development in words with carrier sentences. S1 is the target syllable. S2 the stress alternate.*

3.3. Discussion

The results support the existence of lexical stress in Urdu. Stressed syllables have a significantly longer duration in both conditions. This tallies with the results in [17] for Hindi.

The f0 results are new (as far as we know) and illustrate an interaction between lexical stress and intonational phonology. In both isolated and carrier conditions, the f0 is significantly lower in the stressed syllable. This is consistent with an association of L with a stressed syllable, as posited by the L*H analysis of the basis LH contour However, this basic LH contour is only in evidence in words produced within the carrier sentence (Figure 4), but not in words spoken in isolation (Figure 3). The placement in the carrier sentence invites speakers to emphasize the word: it is in the preverbal focal position [23, 24, 3] and it expresses new/noteworthy information. The literature has identified differing acoustic correlates for the expression of focus, robustly among these is an increased pitch span (e.g., [7]).

In Figure 4 both types of words end on a target H at a similar level, in contrast to the contours in Figure 3. This final H in Figure 4 realizes an increased pitch span in comparison to Figure 3. This is consistent with the expression of focus.

Our results thus indicate that the basic LH pattern observed consistently in the literature is an effect of phrasal intonation. Our results are consistent with an L*H analysis and we suggest that the variation or misalignments cited by scholars such as [3] in favor of a phrasal L_pH_p approach may not have taken into account in sufficient detail and depth the complexities of the interaction between lexical stress and phrasal intonation.

4. Experiment 2: Minimal Pairs

4.1. Methods

4.1.1. Materials

The effects of lexical stress are best studied via complete minimal pairs. However, we were only able to identify the two pairs in Table 2, where word meaning is distinguishable via stress [4].

Table 2: *Minimal word pairs differentiated by stress*

1st syllable stressed		2nd syllable stressed	
ˈgəlɑ	‘throat’	gəˈlɑ	‘to cook’
ˈgəʈɑ	‘thick cloud’	gəˈʈɑ	‘to reduce’

In order to elicit a strong acoustic signal for stress, we embedded them in in question-answer pairs, which elicited a production with contrastive stress on the target word, e.g., “Did you burn the meat? — No, I **cooked** the meat.", “Is your ear infected? — No, my **throat** is infected.”

4.1.2. Procedure

All questions were pre-recorded and presented to the participants via headphones. The answers containing the target words were presented on a laptop screen in Urdu script. The participants listened to the questions and then produced the answers as presented on the screen. The recording took place in an anechoic chamber at the sampling rate of 48kHz.

4.1.3. Participants

Fourteen Pakistani Urdu speakers (∅ = 20.9 years, SD = 2.33 years, 8 female) participated in Experiment 2 for a small payment. They had the same background as in Experiment 1.

4.1.4. Data treatment

A total of 56 recordings were collected in Experiment 2 (2 x 2 words x 14 speakers) and annotated as in Experiment 1.

In our pairs, stress pattern happened to correlate with word category: the words with S1 stress are both nouns, the words with S2 stress are verbs. The target words are thus necessarily placed in different syntactic environments. This in turn means that f0 measurements are not meaningful as it is complex to distinguish between effects of sentence intonation and lexical stress. However, durational effects should only be sensitive to lexical stress and position within a prosodic unit. Duration was therefore analyzed using the same method as in Experiment 1.

4.2. Results

The statistical analysis showed an increase in duration in the S1 if the syllable was stressed (β = -0.015, SE = 0.005, t = -2.802, $p < 0.05$). The S2 also showed an increase in duration if stressed, but the effect stops short of being significant.

4.3. Discussion

The results reveal a durational effect on the lexically stressed S1, but not on the stressed S2. We attribute this to the difference in word category, which leads to a different syntactic placement in the clause, which in turn affects prosodic phrasing. The words with a stressed S1 are nouns and in our material were placed before the verbal complex. In this position they receive a typical LH pattern and the (unstressed) S2 of the noun is the last syllable of a prosodic phrase [5]. The S2 therefore in all likelihood undergoes lengthening, leading to a lack of difference between the unstressed S2 of the nouns and the stressed S2 of the verbs. The stressed syllables of the verb have a longer duration due to lexical stress; however, in our examples the verb is not phrase final (being followed by a light verb, [25]), preventing prosodic phrase final lengthening and we therefore see no statistical effect on the S2.

However, a statistical effect is observed when comparing the S1. The S1 duration of both nouns and verbs here is not confounded by boundary lengthening effects and we can see the effect of lexical stress: the stressed S1 of the noun has a longer duration in comparison to the unstressed S1 in the verb.

5. Experiment 3: Perception

We noticed that the reported examples of inconsistent native speaker intuition with respect to stress placement seemed to involve instances of weight clash (cf. (1)). We therefore designed an experiment testing speaker perception of lexical stress with respect to words that exhibit a weight clash vs. ones that do not.

5.1. Methods

5.1.1. Materials

We selected 14 frequent Urdu words, of which 7 exhibit weight clash: 3 trisyllabic and 4 bisyllabic. All the target syllables have a CVV.CVV structure. The 7 weight clash words were randomized with the 7 words which do not show weight clash.

Table 3: *List of words with and without weight clash*

Words with weight clash		**Words without weight clash**	
nɑˈrɑzi	‘upset’	əxbɑˈrɑt̪	‘newspapers’
xɑˈmoʃi	‘silence’	ehkɑˈmɑt̪	‘orders’
zəˈmɑnɑ	‘time’	mozuˈɑt̪	‘topics’
ˈgɑnɑ	‘sing’	hɑˈlɑt̪	‘condition’
ˈbeʈi	‘daughter’	meˈd̪ɑn	‘ground’
ˈgorɑ	‘fair’	ˈbʊlɑ	‘call’
ˈmiʈʰɑ	‘sweet’	ˈləgɑ	‘fix’

5.1.2. Procedure

Randomized target words with their syllabification were presented to participants on a paper in Urdu script. Participants were asked to read the word and circle the most prominent syllable. They were allowed to opt out of circling any syllable. if they thought no syllable was stressed.

5.1.3. Participants

The data was collected from 14 speakers (∅ = 27.3 years, SD = 5.13 years, 7 male) who were born and raised in Lahore. Since there is a possibility that a regional language may influence stress perception [16] we only consulted speakers who spoke the same regional language (Punjabi) besides Urdu.[4]

5.2. Results

No words were left unlabeled. Participants were not able to judge stress in bi- or trisyllabic weight clash words at a better than chance level. However, in situations where there was no weight clash, participants perceived the target syllable as stressed ca. 75% of the time in bisyllabic words, and 76% in trisyllabic words.

5.3. Discussion

Syllable weight has been tied to the determination of stress in Urdu/Hindi [13, 15, 14] so each heavy syllable can lead to stress perception, causing variability in stress assignment. This variability is reflected in our data. However, in words without weight clash, participants prefer right edge accentuation with stress on the final or penultimate syllables depending on the moraic weight of the syllable. This preference for right edge accentuation depending on the moraic weight of the syllable is in line with [13, 14], but the stark effect of ±weight clash has not been previously reported. We are currently exploring the role of on-going lexicalization of originally derivational words as a source of the variation we observe. Our results so far, however, do let us conclude that lexical stress perception is not random, but is dependent on syllable weight and weight clash.

6. Conclusion

We have presented three experiments that were designed to probe the question of lexical stress in Urdu. Our results support the existence of lexical stress with duration being a strong indicator. We also found that lexical stress interacts with phrasal intonation and that our results are compatible with a L*H analysis of the basic LH pattern in Urdu. We also found that speakers indeed vary in their perception of stress, as has been reported previously in the literature, but that this variation is related to syllable weight and weight clash.

7. Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for funding within project BU 1806/9-2 “Information Structure and Questions in Urdu/Hindi” of the FOR2111 *Questions at the Interfaces* and the DAAD (German Academic Exchange Office) for funding as part of a German-Pakistan Cooperation scheme.

[4]Monolingual speakers of Urdu are exceedingly rare.

8. References

[1] E. Nöth, A. Batliner, V. Warnke, J. Haas, M. Boros, J. Buckow, R. Huber, F. Gallwitz, M. Nutt, and H. Niemann, "On the use of prosody in automatic dialogue understanding," *Speech Communication*, vol. 36, no. 1-2, pp. 45–62, 2002.

[2] J. Stergar and Ç. Erdem, "Adapting prosody in a text-to-speech system," in *Products and Services; from R&D to Final Solutions*, I. Fuerstner, Ed. IntechOpen, 2010.

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

[4] M. Ohala, "A search for the phonetic correlates of Hindi stress," in *South Asian languages: structure, convergence, and diglossia*, B. Krishnamurti, C. Masica, and A. Sinha, Eds. Motilal Banarsidass, 1986.

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

[6] F. Jabeen and B. Braun, "Production and perception of prosodic cues in narrow & corrective focus in Urdu/Hindi," in *9th International Conference on Speech Prosody*, 2018, pp. 30–34.

[7] 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, 2008.

[8] F. Kügler, "Post-focal compression as a prosodic cue for focus perception in Hindi," *Journal of South Asian Linguistics*, vol. 10, pp. 38–59, 2019.

[9] B. Hayes and A. Lahiri, "Bengali intonational phonology," *Natural Language & Linguistic Theory*, vol. 9, no. 1, pp. 47–96, 1991.

[10] S. Khan, "Intonation Phonology and Focus Prosody of Bengali," Ph.D. dissertation, University of California, 2008.

[11] A. R. Kelkar, *Studies in Hindi-Urdu*. Postgraduate and Research Institute, Deccan College, 1968, vol. 35.

[12] R. Dixit, "The segmental phonemes of contemporary standard hindi," Master's thesis, University of Texas, 1963.

[13] P. K. Pandey, "Word accentuation in Hindi," *Lingua*, vol. 77, no. 1, pp. 37–73, 1989.

[14] S. Hussain, "Phonological processing for Urdu text to speech system," *Contemporary Issues in Nepalese Linguistics*, 2005.

[15] M. Ohala, "Stress in Hindi," in *Studies in stress and accent*, L. M. Hyman, Ed. Los Angeles: Southern California Occasional Papers in Linguistics, 1977, vol. 4, pp. 327–338.

[16] B. Hayes, *Metrical stress theory: Principles and case studies*. University of Chicago Press, 1995.

[17] R. Nair, "Acoustic correlates of lexical stress in Hindi," in *Linguistic Structure and Language Dynamics in South Asia – Papers from the proceedings of SALA XVIII roundtable*, 2001.

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

[19] B. Mumtaz, S. Urooj, S. Hussain, and E. U. Haq, "Break Index (BI) annotated speech corpus for Urdu TTS," in *Conference of The Oriental Chapter of International Committee for Coordination and Standardization of Speech Databases and Assessment Techniques (O-COCOSDA)*. IEEE, 2016, pp. 22–27.

[20] A. Turk, S. Nakai, and M. Sugahara, "Acoustic segment durations in prosodic research: a practical guide," in *Methods in Empirical Prosody Research*, S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, and J. Schliesser, Eds. Berlin, New York: De Gruyter, 2006, pp. 1–28.

[21] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program, Version 5.3.56]," 2013, available at http://www.praat.org/ [retrieved 15.09.2013].

[22] B. Mumtaz, A. Hussain, S. Hussain, A. Mahmood, R. Bhatti, M. Farooq, and S. Rauf, "Multitier annotation of Urdu speech corpus," in *Conference on Language and Technology (CLT14), Karachi, Pakistan*, 2014.

[23] 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.

[24] A. Kidwai, *XP-adjunction in Universal Grammar: Scrambling and binding in Hindi-Urdu*. Oxford: Oxford University Press, 2000.

[25] M. Butt, *The Structure of Complex Predicates in Urdu*. Stanford, California: CSLI Publications, 1995.