

Variable Prosodic Realization of Verb Focus in Urdu

Farhat Jabeen, Tina Bögel, Miriam Butt

University of Konstanz

firstname.lastname@uni-konstanz.de

Abstract

Urdu/Hindi has SOV default word order and an immediately preverbal default focus position [1, 2]. When focus is syntactically determined relative to the verb, verb focus is difficult to realize by position. We were therefore interested in the interplay between syntax and prosody with respect to verb focus in languages such as Urdu/Hindi and investigated the prosody associated with verb focus in Urdu declaratives. We tested SOV vs. SVO orders in contexts of broad vs. contrastive focus. In order to force the verb to be focused, we created situations of contrastive verb focus and conducted a production experiment. Our results show that the prosodic realization of verb focus is variable. The general pitch contour on prosodic phrases seems to be L*+H [3, 4]. Verb focus in SVO contexts is realized via a higher pitch span, which is consonant with existing literature [3]. In contrastive SOV contexts, however, the clause ends with L% as opposed to non-contrastive SOV, which features H%. Prosodic realization of verb focus in Urdu is thus variable according to context. We conclude that the L% in SOV is used to signal verb focus because it indicates a marked prosodic structure in contrast to the H% found otherwise on declaratives. Simply increasing the pitch span of the L*+H contour would not serve to provide a clear prosodic distinction in conjunction with a H% boundary tone. In conclusion, the prosodic realization of focus in Urdu is variable across different syntactic structures.

Index Terms: verb focus, contrastive focus, word order, Urdu/Hindi

1. Introduction

This study investigates the acoustic realization of verb focus in contrastive and non-contrastive contexts in Urdu in interaction with word order. Urdu/Hindi¹ is an SOV language in which all major constituents can be reordered. Word order is known to be tied up with the expression of information structural notions [1]. Topics tend to appear clause initially, default focus is in the immediately preverbal position and postverbal material tends to be backgrounded [5, 6, 2].² Urdu/Hindi has been described as a *wh*-in situ language; however, the default position for *wh*-phrases is the immediately preverbal position. As this is the default focus position and *wh*-phrases are generally analyzed as being inherently focused, this pattern confirms the overall view of the interaction between word order and information structure.

¹Urdu and Hindi are very closely related languages, mainly separated via script and religious issues. They are structurally so close that researchers tend to refer to them as one language. Our data comes exclusively from Pakistani speakers of Urdu.

²However, there is no simple one-to-one mapping from linear position to information structure. Postverbal material may also be placed there for prosodic reasons, cf. the heavy NP-shift phenomenon in English or to indicate particularly newsworthy items. The clause initial position is also multiply ambiguous.

In a language where the default focus position is defined with respect to the position of the verb, the question arises how verb focus is expressed. This paper presents an investigation of the expression of verb focus with respect to SOV vs. SVO word order. In order to force the verb to be focused, we created situations of contrastive verb focus and studied the resulting intonation in SOV vs. SVO orders in contexts of broad vs. contrastive focus, respectively.

Our results show that the prosodic realization of verb focus is variable. The pitch contour on prosodic phrases is generally realised as L*+H [3, 4]. Verb focus in SVO contexts is realized via a more extended pitch span, which is consonant with some of the previous literature [3]. In contrastive SOV contexts; however, the clause ends with a low boundary tone (L%) as opposed to non-contrastive SOV, which features a high boundary tone (H%). Prosodic realization of verb focus in Urdu is thus variable according to context. We conclude that the L% in SOV is used to signal verb focus precisely because it indicates a marked prosodic structure in contrast to the H% of the other declaratives. Simply increasing the pitch span of the L*+H contour would not serve to provide a clear prosodic distinction in conjunction with a H% boundary tone. In conclusion, the prosodic realization of focus in Urdu is variable across different syntactic structures.

2. Related Work

Issues surrounding the prosodic realization of focus are complex as they involve an understanding of how the expression of pragmatics, (morpho)syntax and prosody interact in any given language. Focus has been studied most extensively for Germanic languages, in particular English. SOV languages like Urdu/Hindi are to date understudied, though more research on South Asian languages is emerging. South Asian languages form a linguistic area in that there are characteristic language structures that are found in the languages of South Asia, regardless of their genetic classification. The realization of intonation is, at least impressionistically, likely to have common characteristics across South Asia as well, though this remains to be established.

2.1. Focus and Prosody

We first address the issue of focus type. There are various different types of focus and the realizations of these focus types are complex [7, 8, 9]. Our study is concerned with contrastive vs. default focus, both of which have been analyzed in terms of the Alternative Semantics approach to focus [10]. Under this approach, focus is described as the selection of one option from a number of other available possible options in a given set. Focus may be realized morphosyntactically as well as prosodically by singling out certain parts of a sentence as focus elements. Both contrastive and noncontrastive focus signal a choice be-

tween a number of possible alternative answers. With contrastive focus, the possible alternatives are contextually more specified [10]. We adopt this approach and, in particular, work with Krifka’s approach to information structure, who sees all notions of information structure as expressing choices between alternatives as part of Common Ground management [11].

Previous literature has established that crosslinguistically different types of focus are realized by manipulating the available acoustic cues [12]. Word order can also be used to indicate information structure and this in turn naturally interacts with the prosodic structure of an utterance [13, 14].

Previous studies offer mixed evidence with respect to the prosodic marking of focus type. Some researchers suggest that contrastive focus is marked by a L+H* pitch accent [15], whereas others report that it aligns with a drop in F0 [16] and that the gradual or steep realization of this fall also plays a major role in the expression of contrastive focus [17]. Overall, there is a range of different acoustical cues for the expression of contrastive focus: F0 manipulation, increased duration, early peak, wider pitch span and a postfocal suppression of the pitch contour [18, 19]. Moreover, the segmental content of the postfocal elements also plays a role in the realization of the fall often associated with contrastive focus.

Focus marking has also been claimed to interact with prosodic phrasing. In Bengali, focus is shown to insert a right phonological phrase boundary [4, 20]. This is also observed in a typologically quite different language: Chicheŵa [21]. Phonetically, this boundary is marked by a low pitch accent and high phonological boundary (L* H-) in Bengali but by penultimate lengthening and retraction of tones in Chicheŵa [13].

2.2. Focus and Prosody in South Asian Languages

Since the seminal paper of Hayes and Lahiri on Bengali intonational phonology, several further studies on South Asian languages have appeared, providing suggestive information as to the interaction between focus and prosody in these languages. Hayes and Lahiri investigate yes/no questions and default focus and show that there is a L*H_p (narrow) focus accent and that H is a general marker of focus that is placed on the right boundary of the focused element. Similar results are found for Tamil [22] and Konkani, for which Féry ([9]) provides evidence that the high boundary tone concluding a focussed prosodic phrase is higher than in an unfocussed phrase.

For Urdu/Hindi, the relationship between intonation, word order and focus in Hindi has been investigated with respect to the equivalent of English cleft constructions [23]. The study concludes that focus in Hindi is realized by reduction in the postfocal pitch contour rather than by a raising of the pitch on the focused element. The patterns investigated are not equivalent to our stimuli.

A study of contrastive focus concentrating on adjectives in question-answer pairs found that contrastive focus in Hindi is realized by an increase in the pitch span and duration of the focused element [3].

A further study investigated the interplay between word order and prosody in Urdu in *wh*-constituent questions [24]. The study examined the placement of *wh*-phrases before the verb (default focus position) and in immediately postverbal position. Previous accounts of the placement of *wh*-phrases relied exclusively on syntactic constraints [25, 26] or pragmatic differences in terms of information seeking vs. rhetorical question contexts [27]. In contrast, [24] bring in prosody to explain the phenomena. They also find that focus raises the pitch span (cf. [3]).

Moreover, they suggest that the position of *wh*-phrase influences the overall prosodic structure. When the *wh*-phrase is in focus it is placed preverbally in the default focus position. However, when verb focus is to be effected, the *wh*-phrase is concomitantly defocused and placed immediately postverbally within the verbal complex.

3. Experiment

A crucial assumption in [24] is that the immediately postverbal ordering of a *wh*-constituent is due to verb focus. However, they do not establish how verb focus is realized in declaratives. Given that [23] also did not look at verb focus per se (but at the equivalent of clefts) and [3] concentrated on contrastive focus on adjectives, we set out to understand the interplay between prosody and word order with respect to verb focus.

3.1. Data set

Recall that Urdu/Hindi allows syntactic reordering of all major constituents. This offers six different possibilities of word order with subject, object, and verb. Out of these six possibilities, we selected SOV vs. SVO. This is motivated by the immediately preverbal vs. postverbal orders that [24] investigated. The position of the subject remained constant in the clausal initial position.

In order to force verb focus, we set up contrastive contexts for the verb. Thus, our data set consisted of declarative sentences in four conditions: Subject-Object-Verb with broad focus (SOV), Subject-Object-Verb with contrastive verb focus (SOVF), Subject-Verb-Object with broad focus (SVO), and Subject-Verb-Object with verb in contrastive focus (SVFO). All the subjects and objects were proper nouns followed by case markers, which are clitics in Urdu/Hindi [28]. The target verbs were all perfective and causative.³ All the content words in the target sentence were trisyllabic (CV.CV.CV) and contained sonorants. The same set of nouns and 12 verbs was repeated for all the conditions with a random pairing pattern. The last two syllables of the verbs ended on the same vowel /a/ whereas the first syllables in the target verbs differed. Additionally, all the nouns ended with the same vowel /a/ but differed in the first two vowels.

The sentences were presented in random order and filler sentences were added after every three target sentences. A total of 48 sentences (36 target + 12 fillers) per speaker were recorded. SOV and SVO sentences with broad focus were presented without any context whereas interrogative sentences set up a contrastive context for SOVF and SVFO sentences. The target sentences for SOVF and SVFO began with a negation. An example of the target sentences with the corresponding context-setting questions is given below:⁴

SOV (no context)

- (1) jahina=ne norina=ko rul-a-ya
 Shahina=Erg Norina=Acc cry-Caus-Perf.M.Sg
 ‘Shahina made Norina cry.’

³Most transitive verbs are underlyingly causatives and formed with the causative morpheme *-a-*. We chose to use causatives consistently in order to be able to work with high frequency verbs. Using perfective morphology along with case marked objects avoids issues of verb agreement as the verbs then always have the default masculine form *-a-*.

⁴The *kya* ‘what’ at the beginning of the questions providing is an optional marker of interrogativity in polar questions.

SOVF (contrastive context)

(2) Q: kya jahina=ne norina=ko hās-a-ya?
what Shahina=Erg Norina=Acc laugh-Caus-Perf.M.Sg
'Did Shahina make Norina laugh?'

A: nahi, jahina=ne norina=ko rul-a-ya
no Shahina=Erg Norina=Acc cry-Caus-Perf.M.Sg
'No, Shahina made Norina cry.'

SVO (no context)

(3) jahina=ne rul-a-ya norina=ko
Shahina=Erg cry-Caus-Perf.M.Sg Norina.F=Acc
'Shahina made Norina cry.'

SVFO (contrastive context)

(4) Q: kya jahina=ne hās-a-ya norina=ko?
what Shahina=Erg laugh-Caus-Perf.M.Sg Norina=Acc
'Did Shahina make Norina laugh?'

A: nahi, jahina=ne rul-a-ya norina=ko
no Shahina=Erg cry-Caus-Perf.M.Sg Norina=Acc
'No, Shahina made Norina cry.'

3.2. Participants

Ten (6 male and 4 female) speakers of Urdu were recorded. The subjects were from Pakistan, fluent speakers of Urdu and between 23 to 35 years of age. All were familiar with English and at least one further indigenous Pakistani language.⁵

3.3. Data collection

The data was recorded with a head mounted mic in a sound proof environment at the sampling frequency of 44.1 KHz. The stimuli were presented via PowerPoint. The participants were familiarized with the data before the actual recording. All of them went through the data once and read the sentences loudly. For the recording, they were required to read the context sentence loudly and then read the answer. The participants were asked to speak as naturally as possible and repeat the whole sentence in case of coughing, laughing or stuttering. They were also advised to avoid major movement while speaking the stimuli and press the button for the next slide only after they finished reading the target sentence. Each speaker could take as much time as they wanted while reading the stimuli. The average duration of the experiment was 4.24 minutes.

3.4. Data analysis

The recordings were analysed using Praat [30]. After the automatic separation of the target sentences from the fillers, the resulting 360 sentences were manually divided into syllables and labelled accordingly. The F0 was measured in the middle of each vowel in all the syllables manually. Moreover, the duration of syllables was also measured. 27 sentences had to be discarded for two reasons: 1) the speakers showed no discernible pitch on the verb; 2) the insertion of an extra vowel into certain verbs, which turned our target three syllable verbs into four.⁶ The final count thus showed 333 items.

⁵In Pakistan Urdu is the national language, but is mainly spoken as a second language. The L2 users of Urdu are more than nine times in number than the native speakers [29]. It is next to impossible to find monolingual speakers of Urdu.

⁶A hypothesis for this vowel insertion is the influence of Punjabi. Some of the respondents grew up in a Punjabi speaking environment. It is claimed that Punjabi phonotactic constraints do not allow consonant clusters, leading to vowel epenthesis especially at the coda position.

The statistical analysis was performed via a linear mixed effects regression model (lmer) with the condition (SOV, SOVF, SVO, SVFO) as fixed factor and subjects and items as crossed random factors.

4. Results

4.1. Duration

The results indicate that there are no systematic duration differences in the acoustic realisation of contrastive and noncontrastive verb focus. Urdu has been described as a duration sensitive language [31] in that vowels are distinguished by duration (long vs. short). Vowel length interacts with the placement of lexical stress, which is sensitive to syllable weight. It would thus make sense that durational differences are kept to a minimum in the expression of focus.

4.2. Pitch

Figures 1 and 2 represent the time-normalized pitch contours for the SOV and SVO conditions; the dashed line indicating the pitch of the contrastive construction. S₁-1-3, O₁-1-3, and V₁-1-3 refer to the three syllables of the subject, object and verb, respectively. C1 is the case marker following the subject and C2 is the second case marker following the object. The (syntactic) constituents are indicated by the vertical grey lines. While focus in Urdu does not seem to be expressed by means of duration, the evaluation of pitch showed several clear differences between the 4 conditions.

In the SOV/SOVF condition, the rise of F0 from the second to the third syllable of the object is steeper with the target syllable being significantly higher (O₃: SE=2.2, t=-2.9). Another significant difference can be found on the last syllable of the verb. Here, the final syllable of the focussed verb is significantly lower than in the unfocussed variant (V₃: SE=6.2, t=-2.1).

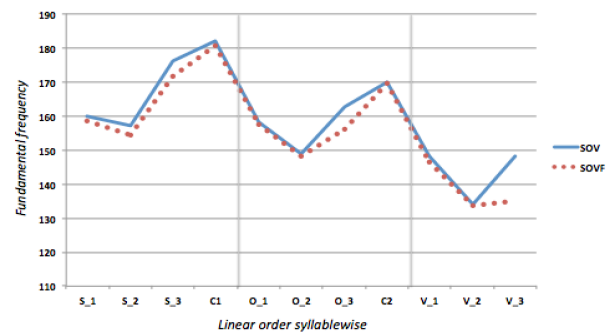


Figure 1: Pitch Contour of SOV vs SOVF

In the SVO/SVFO variant (Figure 2) all syllables in the subject and the verb are significantly higher in the SVFO condition (Verb: V₁: SE=1.6, t=2.8; V₂: SE=1.5, t=2.1; V₃: SE=2.8, t=3.1); no significant differences were found for the object, but a prominent rise can be found on the case marker of the object (C2, but see discussion below).

5. Discussion

The results for the SVO variant have to be relativized. First, there is no obvious reason as to why the subject should be sig-

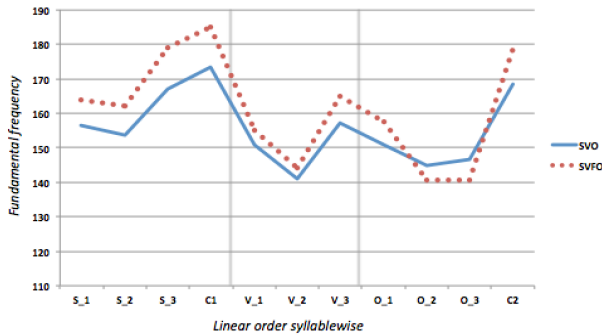


Figure 2: Pitch Contour of SVO vs SVFO

nificantly higher in the SVFO condition. A possible conditioning factor could be that the initial *nahī* ‘no’ in the contrastive context forces the pitch onto a higher level to allow for a fall to the focussed L* and a following steep rise on the verb. In the SOV condition, on the other hand, such a forced pitch raising would not be necessary as there is an intermediate constituent between the subject and the verb. Further research needs to be conducted to verify this hypothesis.

Another data point that requires relativization is the final prominent rise on the case marker of the object (C2). While there is a general rise on the final syllable in declaratives, the data for the case marker given in Figure 2 had to be reduced significantly, as most of the case markers in this condition were highly glottalized. This led to a 62% reduction of the C2 pitch data — the results given in Figure 2 for the case marker were therefore computed with very small numbers and are not reliable.

While glottalization reduced the measurable pitch on the object case marker in the SVO/SVFO variant, it is important to note that glottalization did not occur in the SOV/SOVF variant. Speakers might use glottalization to deaccentuate constituents and mark an intonational phrase boundary [32]. This observation is in line with theories of postfocal deaccentuation, also visible in Figure 2, where the postverbal object is lower when following a contrastively focussed verb than a noncontrastive verb.

Turning to verb focus, in the SVO ordering the verb is found with an extended pitch span. This is consonant with the previous literature [3]. The results for the SOV ordering, in contrast, raise several questions. While sentences of this type generally end in a high intonational phrase boundary tone in our data, the final syllable of the focussed verb in the SOVF condition is significantly lower (albeit still rising) than its non-focussed counterpart (SOV). There are two possible explanations for this. The first involves issues of markedness: depending on the relative position in the sentence, focus can be realized differently. While sentence-medial focus is indicated by extended pitch range, focus at the sentence-final position is signaled via a lower boundary tone to differentiate it from the high intonational phrase boundary.

The second involves a rethinking of the L*+H pattern for phonological phrases that has been assumed. Rather than an L*+H, phonological phrases have an L*, followed by a high prosodic phrase boundary tone: L* H- [4]. It is difficult to distinguish L*+H from L* H- at a sentence-medial position, but in the sentence final conditions such as the SOV/SOVF variant a

difference would become apparent due to the interplay with the boundary tone of the intonational phrase. If the boundary tone is L, as established for Bengali [4], then the lower final pitch in the SOVF condition follows from an interaction with the boundary tone. If the boundary tone is high, as indicated by our data, variable focus marking must again be at play.

6. Conclusion

Our study of the expression of verb focus yielded two different patterns. In SVO structures, verb focus followed the attested pattern of extending the pitch range of the L*+H (or L* H-) intonational contour associated with phonological phrases. In SOV contexts, however, a different strategy could be identified. Here verb focus is expressed via a low boundary tone, contrasting with the high boundary we otherwise found with declaratives in Urdu/Hindi. The acoustic expression of focus is thus variable, depending on the syntactic structure of the utterance.

Further research is needed to determine which of the two alternative explanations for the marked structure provides the more comprehensive explanation for understanding the intonational structure of Urdu/Hindi.

7. References

- [1] V. Gambhir, “Syntactic restrictions and discourse functions of word order in standard Hindi,” Ph.D. dissertation, University of Pennsylvania, Philadelphia, 1981.
- [2] A. Kidwai, *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. Oxford: Oxford University Press, 2000.
- [3] S. Genzel and F. Kügler, “The prosodic expression of contrast in Hindi,” in *The Proceedings of Speech Prosody 2010*, Chicago, 2010.
- [4] B. Hayes and A. Lahiri, “Bengali intonational phonology,” *Natural Language and Linguistic Theory*, vol. 9, pp. 47–96, 1991.
- [5] 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.
- [6] —, “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>.
- [7] C. Gussenhoven, *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris, 1984.
- [8] D. Büring, *Intonation and Meaning*. Oxford: Oxford University Press, 2015.
- [9] C. Féry, “Focus as prosodic alignment,” *Natural Language and Linguistic Theory*, vol. 31, no. 3, pp. 683–734, 2013.
- [10] M. Rooth, “Association with focus,” Ph.D. dissertation, University of Massachusetts, Amherst, 1985.
- [11] M. Krifka, “Basic notions of information structure,” *Acta Linguistica Hungarica*, vol. 55, no. 3–4, pp. 243–276, 2008.
- [12] M. Breen, E. Fedorenko, M. Wagner, and E. Gibson, “Acoustic correlates of information structure,” *Language and Cognitive Processes*, vol. 25: 7, pp. 1044–1098, 2010.
- [13] J. Costa and N. Kula, “Focus at the interface: Evidence from romance and bantu,” in *The Bantu-Romance Connection*, C. D. Cat and K. Demuth, Eds. John Benjamins, 2008.
- [14] J. Kahn and J. Arnold, “A processing-centered look at the contribution of givenness to durational reduction,” *Journal of Memory and Language*, vol. 67, pp. 311–325, 2012.
- [15] C. Bartels and J. Kingston, “Salient pitch cues in the perception of contrastive focus,” in *Focus and natural language processing [IBM Working Papers on Logic and Linguistics, v. 6]*, I. P. Boach and R. V. der Sandt, Eds., 1994, vol. 6.

- [16] E. Cooper, S. Eady, and P. Mueller, "Acoustical aspects of contrastive stress in question answer contexts," *Journal of Acoustical Society of America*, vol. 77(6), 1985.
- [17] E. Couper-Kuhlen, "A new look at contrastive intonation," in *Modes of interpretation: Essays presented to Ernst Leisi*, R. Watts and U. Weidman, Eds. Tübingen, Germany: Gunter Narr Verlag, 1984.
- [18] A. Chen, *The prosodic investigation of information structure*. Mouton de Gruyter, 2012, ch. The expression of information structure, pp. 249–286.
- [19] J. Hanssen, J. Peters, and C. Gussenhoven, "Prosodic effect of focus in dutch declaratives," in *Proceedings of Speech Prosody*, P. A. Barbosa, S. Madureira, and C. Reis, Eds., 2008, pp. 609–612.
- [20] S. ud Dowla Khan, "Intonational phonology and focus prosody of bengali," Ph.D. dissertation, University of California, 2008.
- [21] H. Truckenbrodt, "The syntax-phonology interface," in *The Cambridge Handbook of Phonology*, P. de Lacy, Ed. Cambridge: Cambridge University Press, 2007, pp. 435–456.
- [22] 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.
- [23] U. Patil, G. Kentner, A. Gollrad, F. Kügler, C. Féry, and S. Vasisht, "Focus, word order and intonation in Hindi," *Journal of South Asian Linguistics*, vol. 1, no. 1, pp. 55–72, 2008.
- [24] F. Jabeen, T. Bögel, and M. Butt, "Immediately postverbal questions in urdu," in *InterSpeech*, 2015.
- [25] R. Bhatt and V. Dayal, "Rightward scrambling as rightward movement," *Linguistics Inquiry*, vol. 38, no. 2, pp. 287–301, 2007.
- [26] E. Manetta, "Reconsidering rightward scrambling: Postverbal constituents in Hindi-Urdu," *Linguistic Inquiry*, vol. 43, no. 1, pp. 43–74, 2012.
- [27] 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.
- [28] M. Butt and T. H. King, "The status of case," in *Clause Structure in South Asian Languages*, V. Dayal and A. Mahajan, Eds. Berlin: Springer Verlag, 2005.
- [29] 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>.
- [30] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program, Version 6.0.05]," 2015, available at <http://www.praat.org/> [retrieved 08.11.2015].
- [31] S. Hussain, "Acoustic correlates of lexical stress in urdu," Ph.D. dissertation, Northwestern University, 1997.
- [32] L. Redi and S. Shattuck-Hufnagel, "Variation in the realization of glottalization in normal speakers," *Journal of Phonetics*, vol. 29, pp. 407–429, 2001.