On the phrasing properties of Hindi relative clauses

Stella Gryllia[*], Caroline Féry[+], Frank Kügler[#], Pramod Pandey[±]

Leiden University[*], Goethe-Universität Frankfurt[+], Universität Potsdam[#], Jawaharlal Nehru University[±]
s.gryllia@hum.leidenuniv.nl, caroline.fery@gmail.com, kuegler@uni-potsdam.de, pkspandey@gmail.com

ABSTRACT

This paper presents results from a production experiment in Hindi, showing that differences in attachment site of object relative clauses result in prosodic differences when the antecedent of the relative clause (RC) is part of a complex NP with the structure *N1 of N2*. In particular, based on duration and F0 data we argue that the phrasing in a matrix sentence encodes the attachment site of the object RC. When the RC attaches high, i.e. modifying the head N1 of the complex NP, N2 and N1 form together a phonological phrase, while the verb of the matrix clause forms a phonological phrase on its own. In the case of low attachment, i.e. the RC modifies the genitive N2, the N2 forms its own phonological phrase, while N1 forms a phonological phrase with the verb of the matrix clause.

Keywords: Prosody, Phrasing, RC, Hindi.

1. INTRODUCTION

Attachment preferences across constructions and languages have fuelled a long line of research, as they have been used as study cases for theorizing about the basic principles of language processing and their universal validity. For instance, Frazier [10, 11] argued for the universality of parsing mechanisms and formulated the Late Closure Principle (LCP). According to the LCP the parser prefers to attach incoming material to the clause or phrase that is currently being processed, if other grammar principles allow, and if this attachment decision does not require the formation of superfluous syntactic nodes. For relative clauses (RC), the LCP predicts a Local Attachment Preference. This means that in a sentence like (1), the RC *who was on the balcony* preferably refers to the noun *actress*. In this case, the RC modifies N2 and we speak of 'low attachment' preference.

(1) Someone shot the servant$_{N1}$ of the actress$_{N2}$ who was on the balcony.

The universal validity of the LCP was questioned by [4] arguing that speakers of Spanish show a high attachment preference (the RC modifies N1) when confronted with constructions like (1), while speakers of English show a low attachment preference.

[4]'s findings inspired numerous studies on the attachment preference of relative clauses across languages (for an overview see [1]). Taken together, these studies revealed that languages differ with respect to the preferred strategy for resolving an attachment ambiguity. Besides Spanish, other languages, like Bulgarian and Dutch, showed a preference for high attachment [29, 3], while Italian and Romanian showed a preference for low attachment [6, 7], similarly to English. This cross-linguistic variation triggered a related line of research, namely a close inspection of the principles that could explain the attachment preference within and across languages. A number of principles were proposed; syntactic principles [14], discourse-semantic principles [12], prosodic principles [8, 9, 28] and the role of frequency [3]. The prosodic principles are of particular relevance for this paper.

1.1. Prosodic principles - phrasing

Fodor [8, 9] argued that phrasing principles account for cross-linguistic variation with respect to relative clause attachment preferences. As a starting point, Fodor assumes that prosodic information is always available to the parser, even in silent reading of texts. This type of prosody is named *implicit prosody* [8, 9] and defined as follows: "In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favours the syntactic analysis associated with the most natural (default) prosodic contour for the construction." [9: 113].

In this paper we investigate whether differences with respect to the preferred attachment site of the RC result in prosodic differences; as a case study we take Hindi, a head-final language spoken primarily in South Asia. Hindi allows at least three positions for attaching a RC: prenominal, postnominal and extraposed. In this paper, we examine the structure S-N2$_{GEN}$-N1-V-RC, in which the RC is extraposed. Note that the labelling of the nouns reflects their syntactic structure; the genitive noun N2 is structurally lower than the head N1 of the NP.

1.2. Hindi prosody

We assume that Hindi has lexical stress and that syllable weight plays a major role in its location [15, 16, 17, 18, 21]. However, it must be noted that the status of lexical stress, that is whether there is always a designated syllable for stress, remains unsettled for Hindi. When lexical stress is realized in isolation, namely, when the words are pronounced out of context, then it is well perceived. In spontaneous speech, lexical stress is often perceived only weakly. This latter property is due to the absence of a systematic pitch accent corresponding to lexical stress. Intonational markers in the form of phrasal tones, are often more prominently realized than lexical stress. Hayes [18: 163] used the term "non-phonemic lexical accent"; for a thorough review of the literature on lexical stress in Hindi see [27]. [24] showed that every non-final prosodic phrase in Hindi, which he called 'foot', starts with a low tone. The prosodic phrase is defined as "one to several syllables in length, which normally is uttered with a continuously rising pitch from beginning to end." [15, 16] and [17] analysed the low part of the rising contour as a low pitch accent, annotated as L* in the tone-sequence notation system of [26], see also [22], and the high part of the rising contour as a boundary tone H%, or H-. Within an intonation phrase the individual rising contours are downstepped to each other [25].

Compared to lexical prominence, phrasing is clearer and prosodic boundaries are perceived better than lexical prominence. In this, Hindi speakers differ from speakers of English. In a perception experiment, [21] asked ten native non-linguist speakers and one linguist speaker of Hindi to identify prominent words and prosodic boundaries in ten excerpts from spoken short narratives. [21] found slight agreement among ordinary listeners and between ordinary listeners and the expert ($\kappa = 0.15$) as far as lexical prominence is concerned. The agreement improves when examining prosodic boundaries. Linguistically untrained listeners and the linguistic expert agreed moderately ($\kappa = 0.41$) with respect to the perception of prosodic boundaries. For speakers of English, [31] found agreement rates at 89.0% for boundaries and 86.0% for pitch accent. [23] found a mean kappa of 0.582 for non-expert speakers for prominence and prosodic boundaries.

As for the phonological interpretation of the data, two competing models have been proposed for Hindi. The first one assumes that Hindi is similar to English in assigning prominent tones (L*) on lexically stressed syllables, as proposed first by [17]. The other one insists on the formation of prosodic phrases on focused material, and does not necessarily associate the low and high tones with lexical prominence, see [25] and [13]. In this paper, we remain neutral as to the role of the tones. We call the initial low tone L and the final high tone H.

2. A PRODUCTION EXPERIMENT

In the production experiment, participants produced two types of object RCs, one modifying the head (N1) of the NP, and the other modifying the genitive (N2). We expected that differences in the attachment site of the object RC (high attachment in the former, and low attachment in the latter) would result in prosodic differences. In particular, we expected differences in phrasing as shown in (2a) and (2b), see [8, 9] as a function of RC attachment preferences and phrasing.

(2) a. High attachment
$[(\text{Subject})_{\varphi}(\text{N2}_{\text{GEN}}\text{N1})_{\varphi}(\text{Verb})_{\varphi}]_{\iota}\,[(\text{RC})]_{\iota}$

b. Low attachment (attach to N2)
$[(\text{Subject})_{\varphi}(\text{N2}_{\text{GEN}})_{\varphi}(\text{N1Verb})_{\varphi}]_{\iota}\,[(\text{RC})]_{\iota}$

Our expectation was that the entire complex NP forms its own phonological phrase when N1 is modified by the relative clause (2a), and that the genitive N2 forms its own phonological phrase when modified by the relative clause (2b).

2.1. Speech materials

The stimuli were 48 main clauses: 24 sentences in two variants each, see (3) for a full example of a stimulus. In (3a) the RC modifies N1, while in (3b) the RC modifies N2. Note that the genitive comes first, but is called N2, because it is structurally lower than the head N1 of the NP.

(3) a. High attachment
Lata-ne gaaRi kii [caabii]$_{\text{N1}}$ DhoonDi [jo
Lata-erg car- gen key search rel
naukar-ne kho dii thii]$_{\text{RC}}$
servant-erg lost
"Lata searched for the keys of the car that the servant had lost."

b. Low attachment
Lata-ne [gaaRi kii]$_{\text{N2}}$ caabii DhoonDi [jo
Lata-erg car- gen key search rel
naukar-ne saaf kii thii]$_{\text{RC}}$
servant-erg clean
"Lata searched for the keys of the car that the servant had cleaned."

The attachment preferences of the object RCs were determined by a paper-pencil questionnaire that was run independently. Participants were asked to answer a forced choice question formed on the basis of the object RC. For instance for (3a) the question in the paper-pencil questionnaire was: 'What had the servant lost? i) the car, ii) the key)'. The paper-pencil questionnaire was conducted at the Jawaharlal Nehru University (JNU) in New Delhi.

In constructing the stimuli, a number of variables were taken into consideration. In particular, we kept the length of N1 and N2 constant, as [19] and [20] have shown that the length of NPs affects prosodic phrasing, so N1 and N2 were always di-syllabic. Moreover, N1 and N2 were always trochaic. We also controlled animacy (see [5] for an effect of animacy on the attachment preferences of RCs); N1 and N2 were inanimate. The length of the object RC was also kept constant. The 48 target sentences (24 main clauses × 2 types of object RCs) were intermingled with 24 fillers. All 72 sentences were presented to each participant in a pseudo-randomized manner. To avoid any ordering effects, two pseudo-randomized lists were used.

2.2. Recording procedures

A self-paced stimulus presentation was used. The utterances were directly recorded via a head-mounted close taking microphone (Shure SM10A) on a computer disk using Audacity Software in a quiet room at JNU. Participants were aware of the attachment site of the object RC. When the object RC modified N1, N1 was underlined, while when the object RC modified N2, N2 was underlined. Participants were instructed to utter the sentence displayed on the screen as naturally as possible.

2.3. Participants

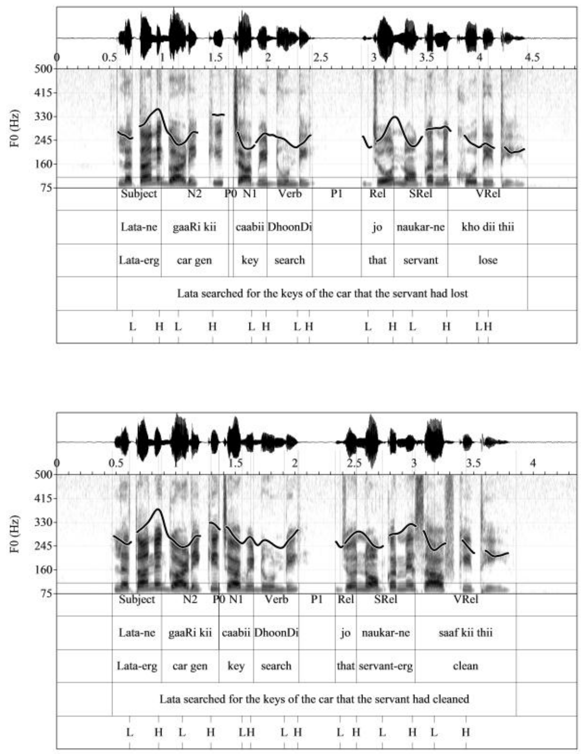
Fourteen native speakers of Hindi as spoken in the New Delhi region participated in the experiment. Each speaker was reimbursed for participation. The experiment lasted 20 minutes.

2.4. Analysis

The productions of all fourteen participants were analysed, 672 utterances in total (48 targets × 14 participants). The data were labelled manually at the word level using Praat [2]. For labelling the data we used conventional segmentation guidelines [30]. Furthermore, we marked all realized pauses (P0) between N2 and N1 and all pauses (P1) between the verb and the relative pronoun *jo* 'that' (see Fig. 1). The durations of N1, N2, P0 and P1 were extracted using a Praat script; the data were analysed with linear mixed models; in all cases, varying slopes were fit by subject and by item, but no correlation parameters were estimated. The contrast coding was always sum contrasts (+/−1 coding). The R package lme4 (version 1.1-8) was used.

We annotated a high (H) and a low (L) tone in every prosodic word using a Praat script which identified first the last H ($F0_{max}$) of the prosodic word and then its closest preceding L ($F0_{min}$) (Fig.1). We also examined the pitch reset of H in N2 to H in N1 subtracting $F0_{max}N1$ from $F0_{max}N2$.

Figure 1: Waveform, spectrogram and F0 of a RC modifying N1(upper) and of a RC modifying N2

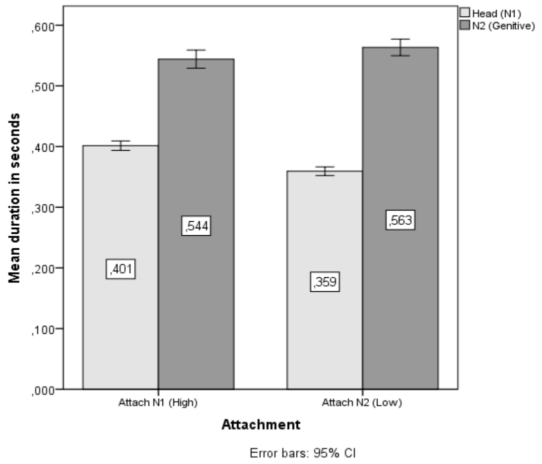


3. RESULTS

3.1. Duration

Figure 2 shows the mean duration of N1 and N2 in seconds for High and Low attachment. On average the duration of N1 when it was modified by the RC (0.401 sec, *S.E.* = 0.004) was significantly longer than when it was not modified by the RC (0.359 sec, *S.E.* = 0.004), the linear mixed model analysis showed that this difference was statistically significant (coef = − 0.02, SE = 0.01, t = − 4.07). Furthermore, on average the duration of N2 when it was modified by the RC (0.563 sec, *S.E.* = 0.006) was numerically longer than when it was not modified by the RC (0.543 sec, *S.E.* = 0.007), but this difference did not reach statistical significance (coef = 0.01, SE = 0.01, t = 0.98), (see Fig. 2).

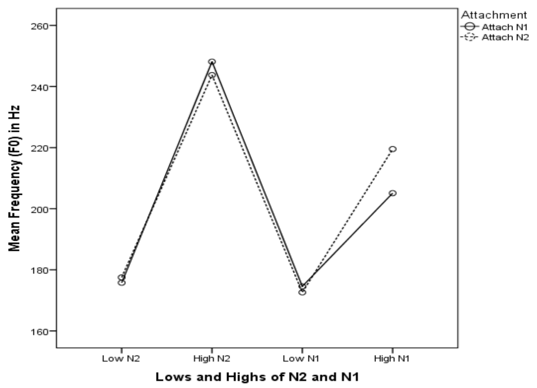
Figure 2: Mean duration of N1 and N2 in seconds.



2.5. Frequency (F0)

On average the $F0_{max}$ of N1 when it was modified by the RC (205.1 Hz, *S.E.* = 3.2) was significantly lower than when it was not modified by the RC (219.5 Hz, *S.E.* = 3.2), the linear mixed models estimates are as follows: coef = − 8.37, SE = 2.55, t = − 3.28. In this analysis, two extreme data points, one less than 100 Hz and one more than 400 Hz was removed. The result does not depend on these points being removed; the difference is significant even with the complete data-set. The $F0_{max}$ of N2 when it was not modified by the RC (248.1 Hz, *S.E.* = 3.8) was on average significantly higher than when it was modified by the RC (243.7 Hz, *S.E.* = 3.7), coef = 13.3, SE = 2.6, t = 5.112, (Fig. 3). Here, 14 data points (out of 671) were removed as they had extreme values (>350 Hz) and were skewing the residuals. However, the result does not depend on removing these values; the difference is statistically significant even when all data points are retained.

Figure 3: Lows and Highs of N2 and N1.



On average the pitch downstep ($F0_{max}N2 – F0_{max}N1$) when the RC was modifying N2 (43.0 Hz, S.E. = 2.5) was significantly shorter than the pitch downstep when the RC was not modifying N2 (24.2 Hz, S.E. = 1.6), coef = 9.4, SE = 3.13, t = 3.0.

4. DISCUSSION

Our expectations that the attachment site of a RC is prosodically signalled were partly met. The two types of phrasing are repeated in (4). We see significant differences in duration of N1 but no significant differences in duration of N2. The null result for N2 is uninformative; it is possible that a study with higher power is necessary to determine whether N2 durations are also affected by RC attachment. We also see differences in $F0_{max}$ of N2 and N1, and in pitch downstep.

In duration, the noun that is modified by the RC is longer compared to when it is not modified. Thus, longer duration on the noun indicates the presence of a phrase boundary, see (4). In F0max we find a higher scaling of the H tone on N1 when the RC modifies N2, while we find a lower scaling of the H tone on N1 when the RC modifies N1. Crucially, the difference of the H tones between N2 and N1 reveals an indication of phrasing. When the RC modifies N2, the difference between the H tone on N2 and N1 is smaller than when the RC modifies N1. In both cases, the H tones on N2 and N1 show a downstep. However, this relation is smaller when the RC modifies N2. This means that a phrase break after N2 induces a pitch reset of the scaling of the following H tone. When no phrase break follows N2, there is no pitch reset, and the two tones show a downstep. The RC forms its own intonational phrase (ι) due to its extraposed postverbal position.

(4) a. High attachment (attach to N1)
$[(\text{Subject})_{\varphi}(\text{N2}_{\text{GEN}}\text{N1})_{\varphi}(\text{Verb})_{\varphi}]_{\iota}\,[(\text{RC})]_{\iota}$

b. Low attachment (attach to N2)
$[(\text{Subject})_{\varphi}(\text{N2}_{\text{GEN}})_{\varphi}(\text{N1Verb})_{\varphi}]_{\iota}\,[(\text{RC})]_{\iota}$

5. CONCLUSION

Our results partly confirmed that differences in attachment result in prosodic differences. The duration and F0 of the head of the complex noun phrase (N1) and the genitive (N2) partially support our hypothesis that object RCs that modify N1 and object RCs that modify N2 differ in phrasing.

Acknowledgements
This research was funded by the DFG with the SPP-1234 project 'Prosody in Parsing' (KU 2323/1-2) at Potsdam University and the Research Unit 1783 on Relative Clauses at Goethe-University Frankfurt. We are also grateful to Prof. P. Pandey and to S. Sudhanshu, graduate student for support at JNU.

7. REFERENCES

[1] Augurzky, P., 2005. *Attaching Relative Clauses in German. The Role of Implicit and Explicit Prosody in Sentence Processing*. PhD dissertation, Leipzig University.

[2] Boersma, P., Weenink, D. 2014. Praat: doing phonetics by computer (Version 5.4.04), http://www.praat.org/.

[3] Brysbaert, M., Mitchell, D.C. 1996. Modifier attachment in sentence processing: Evidence from Dutch. *The Quarterly Journal of Experimental Psychology* 49(A), 664-695.

[4] Cuetos, F., Mitchell, D.C. 1988. Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure Strategy in Spanish. *Cognition* 30, 73-105.

[5] Desmet, T., Brysbaert, M., De Baecke, C. 2002. The correspondence between sentence production and corpus frequencies in modifier attachment. *Quarterly Journal of Experimental Psychology*, 55A, 879-896.

[6] De Vincenzi, M. & Job, R. (1993) Some observations on the universality of the Late Closure strategy: evidences from Italian. *Journal of Psycholinguistic Research*, 22, 189-206.

[7] Ehrling, K., Fernández, E., Fodor, J.D., Stenshoel, E., Vinereanu, M. 1999. Low attachment of relative clauses: New data from Swedish, Norwegian and Romanian. Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing, New York, NY, March 18-120.

[8] Fodor, J. 1998. Learning to parse? *Anniversary Issue of the Journal of Psycholinguistic Research* 27(2), 285-318.

[9] Fodor, J. 2002. Prosodic disambiguation in silent reading. *Proc.NELS 32* New York, 113-132.

[10] Frazier, L. 1978. On comprehending sentences: Syntactic parsing strategies. Unpublished PhD dissertation, University of Connecticut, Storrs, CT.

[11] Frazier, L. 1987. Sentence processing: A tutorial review. In: Coltheart, M. (ed), *Attention and performance XII*. Hillsdale, NJ: Erlbaum.

[12] Frazier, L., Clifton, J.C. 1996. *Construal*. Cambridge, MA: MIT Press.

[13] Genzel, S., Kügler, F. 2010. The prosodic expression of contrast in Hindi, *Speech Prosody* 2010.

[14] Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E. & Hickok, G. 1996. Recency preference in the human sentence processing mechanism. *Cognition* 59, 23-59.

[15] Harnsberger, J.D. 1994. Towards an intonational phonology of Hindi. Ms., University of Florida.

[16] Harnsberger, J.D. 1999. The role of metrical structure in Hindi intonation. In South Asian Language. Analysis Roundtable, University of Illinois, Urbana-Champaign.

[17] Harnsberger, J.D., Judge, J. 1996. Pitch range and focus in Hindi. *The Journal of the Acoustical Society of America* 99. 2493.

[18] Hayes, B. 1995. *Metrical Stress Theory. Principles and Case Studies*. Chicago: University of Chicago Press.

[19] Jun, S.A. 2003. Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research* 32(2), 219-249.

[20] Jun, S.A., Koike, C. 2008. Default prosody and relative clause attachment in Japanese. *Japanese Korean Linguistics* 13 Stanford, 41-53.

[21] Jyothi, P., Cole, J., Hasegawa-Johnson, M., Puri, V. 2014. An Investigation of Prosody in Hindi Narrative Speech, Proceedings of Speech Prosody.

[22] Ladd, D.R. 1996/2008. Intonational Phonology. Cambridge: Cambridge University Press.

[23] Mo, Y., Cole, J., Lee, E-K. 2008. Naïve listeners' prominence and boundary perception. Proceedings of Fourth Speech Prosody. 735-736. http://www.isca-speech.org/archive/sp2008.

[24] Moore, R.R. 1965. A study of Hindi intonation. PhD dissertation, University of Michigan.

[25] Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., Vasishth, S. 2008. Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics* 1. 53-70.

[26] Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. PhD dissertation, MIT. Distributed by Indiana University Linguistics Club, Bloomington.

[27] Puri, V. 2013. Intonation in Indian English and Hindi late and simultaneous bilinguals. PhD dissertation, University of Illinois at Urbana-Champaign.

[28] Schafer, A.J., Carter, J., Clifton, C., Jr., Frazier, L. 1996. Focus in relative clause construal. *Language and Cognitive Processes* 11, 135-163.

[29] Sekerina, I.A., Fernández, E.M., Petrova, K.A. 2004. Relative Clause Attachment in Bulgarian. In: Arnaudova, O., Browne, W., Rivero, M. L., Stojanović, D. (eds.), *The Proceedings of the 12th Annual Workshop on Formal Approaches to Slavic Linguistics. The Ottawa Meeting 2003*, 375-394.

[30] Turk, A., Nakai, S., Sugahara, M. 2006. Acoustic segment durations in prosodic research: A practical guide. In: Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N., Schließer, J. (eds), *Methods in empirical prosody research*. Berlin: Mouton de Gruyter, 1–27.

[31] Yoon, T-J., Chavarría, S., Cole, J., Hasegawa-Johnson, M. 2004. Intertranscriber reliability of prosodic labeling on telephone conversation using ToBI. *Interspeech 2004*. 2729-2732. http://www.iscaspeech.org/archive/interspeech_2004.