

Understanding constraints on non-projectivity using novel measures

Himanshu Yadav
Jawaharlal Nehru University
Center for Linguistics
New Delhi
India
himans53_1lh@jnu.ac.in

Ashwini Vaidya
IIT Delhi
Department of
Electrical Engineering
India
ird11278@ee.iitd.ac.in

Samar Husain
IIT Delhi
Department of Humanities
and Social Sciences
India
samar@hss.iitd.ac.in

Abstract

In this work we propose certain novel measures to understand non-projectivity in various syntactic phenomena in Hindi. This is an attempt to go beyond the analysis of non-projectivity in terms of certain graphical measures such as edge degree, planarity etc. Our measures are motivated by the findings in the processing literature that have investigated the interaction between working-memory constraints and syntactic complexity. Our analysis shows that the measures pattern differently for distinct phenomena and therefore could prove to be beneficial in understanding non-projectivity in a language. We also find some interesting differences in non-projectivity between conversation and news genre.

1 Introduction

One of the main aims of the modern linguistic theories has been to understand the formal properties of the grammar and its interaction with human linguistic competence (Frazier, 1985; Chomsky and Miller, 1963). In order to represent the syntactic structure of a linguistic utterance, most current theories posit some kind of a hierarchical structure (Steedman, 2000; Chomsky, 1995; Hudson, 2010). This hierarchical structure could either be represented via the notion of constituents or through dependency relations (Rambow, 2010). It is also known that languages allow for configurations that lead to discontinuous constituents. Such configurations are known to pose a challenge to grammar formalization and, not surprisingly, they are more difficult to parse computationally (Nivre, 2009; Joshi, 1990). They are also difficult to process by native speakers (Levy et al., 2012; Husain and Vasishth, 2015).

The discontinuous constituents are termed as non-projectivity in the dependency grammar literature. Non-projectivity is characterized by a non-

canonical linear order of words in a sentence.¹ Formally, an arc $i \rightarrow j$ is non-projective if and only if there is at least one word k between i and j that i does not dominate (see Figure 1).

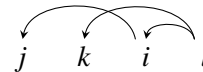


Figure 1: The dependency arc between i and its dependent j is non-projective. All other arcs are projective.

While non-projective dependencies (called discontinuous constituents in phrase structure grammar) are common in many languages that allow free word order, it is also known that not all such configurations are permitted, i.e. not all non-projective dependencies can be deemed grammatical (Joshi, 1985; Shieber, 1985). In order to describe the grammar of a language, it is therefore critical to understand the constraints on non-projectivity in that language. Understanding these constraints will throw light on the cognitive constraints that influence language comprehension and production. Needless to say, a better understanding of non-projectivity will also benefit computational parsers.

Non-projectivity occurs due to discontinuity in the yield of a node, specifically discontinuity in the head-dependent projection chain. This discontinuity in the head-dependent linear order is caused by the intervention of a constituent or sub-tree that is dependent on a head outside the current yield. The properties of this intervening element as well as the properties of the non-projective dependency (comprising a head and its dependent) can describe the constraining environ-

¹This is of course a simplification. As we will discuss later, there are some constructions that are inherently non-projective.

ment for a non-projective dependency. In this work, we will use such properties to identify constraints on non-projectivity in Hindi. In order to do this we use the Hindi-Urdu Dependency Treebank (HUTB) (Bhatt et al., 2009).

There has been some work on studying non-projectivity in Hindi. Mannem et al. (2009) carried out a preliminary study of non-projectivity in HUTB based on some widely used measures, e.g., gap degree, edge degree and planarity (Bodirsky et al., 2005; Kuhlmann and Nivre, 2006; Kuhlmann, 2007). In a similar and more elaborate work, Bhat and Sharma (2012) carried out a formal and linguistic characterization of non-projectivity for Hindi, Bengali, Telugu and Urdu. They characterized non-projectivity based on the nature of the linguistic phenomena (e.g., relativization, genitive constructions etc.) and the cause of non-canonicity (e.g., extraposition, scrambling, etc.). Similar to Mannem et al. (2009), they also used edge degree etc. to characterize non-projectivity.

In this paper, we go beyond this type of analysis to attempt a deeper linguistic understanding of non-projectivity in Hindi. In particular, we ask, what are the limiting conditions for a non-projective dependency? In other words, we attempt to uncover the kinds of non-projective configurations that are *disallowed* for a phenomenon. For example, while studying non-projectivity in genitive constructions, we attempt to identify which type of non-projectivity is not possible in such constructions. We examine the limiting conditions for a non-projective dependency with respect to hierarchical and linear distance and the nature of the intervening constituents and subtrees. A deeper understanding of non-projectivity in a language is critical for positing constraints on the generative power of a dependency grammar and understanding the interaction of working memory constraints and linguistic complexity.

Our paper is organized as follows, in Section 2 we motivate new measures for analyzing non-projectivity based on linear and hierarchical distance. In Section 3, we discuss these measures using a Hindi treebank. We conclude the paper in section 4.

2 A proposal for novel non-projectivity measures

Previous analyses of non-projectivity in Hindi (Mannem et al., 2009; Bhat and Sharma,

2012) and in other languages have characterized sentences that are non-projective using graph-based measures such as gap degree, edge degree, planarity and well-nestedness. While these measures have proven to be very useful, they do not explicitly capture certain information that could be used in positing constraints on non-projectivity for a particular linguistic construction. We propose three novel measures for non-projectivity in this section, based on linear word order as well as hierarchy.

2.1 Linear measures

We look at the examples in figures 3–6 to motivate the first type of linear measure. These figures show non-projective dependencies involving a genitive relation. The noun phrase (NP) *raam-kaa* ‘Ram-GEN’ is the dependent of *chashmaa* ‘spectacles’ in all these examples. Figure 2 shows the projective dependency for this phenomenon. The edge degree² in each of the non-projective structure is 1 (they also have same planarity), however, while 3 and 4 are grammatical, the sentences in figures 5 & 6 are completely ungrammatical. Critically, the type of intervening material that causes non-projectivity differs in these examples. In 3, the intervening element is an adverbial modifying the main verb; in 4, it is a non-finite clause modifying the main verb, and in 5, it is a relative clause modifying a noun outside the span of the genitive dependency arc. The example in figure 6 is ungrammatical because a negation intervenes. Note that the dependencies shown here are between chunks rather than individual words, which is in keeping with the HUTB representation.

These examples show that in order to understand the nature of non-projectivity for a phenomenon like genitive, it is important to study the *type of intervening material*. A metric like edge degree captures the number of intervening constituents spanned by a single edge (Kuhlmann and Nivre, 2006), but it is unable to capture certain linguistic nuances discussed above. Additionally, the type of intervening constituents also capture the complexity of these constituents. While both 3 and 4 are grammatical constructions, the intervening material in 3 is less complex than the one in 4. Capturing the complexity of the intervening

²Let $e=(i, j)$ be a dependency arc with ‘j’ as the head and ‘i’ as the dependent. Edge degree of an arc e is the number of connected components c in the span of arc e such that c is not dominated by ‘j’ (Nivre, 2006).

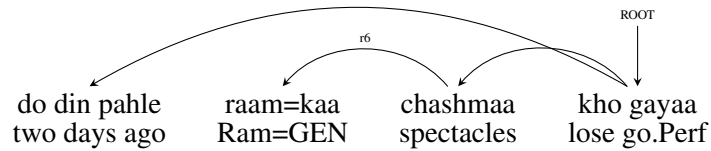


Figure 2: 'Two days ago Ram's spectacles were lost'. Projective Genitive Construction.

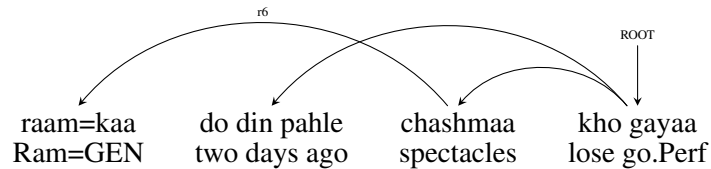


Figure 3: 'Ram's spectacles were lost two days ago'. Non-projectivity with edge degree=1, Type of intervening constituent=NP. NP: Noun chunk.

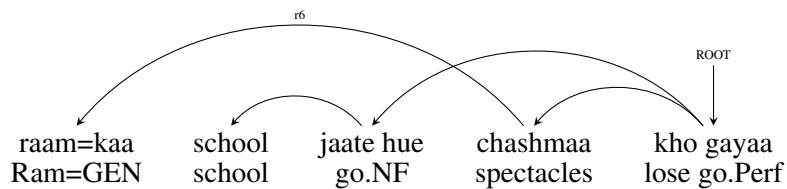


Figure 4: 'Ram's spectacles were lost while going to school'. Non-projectivity with edge degree=1, Type of intervening constituent=VGNF, Length of intervening constituent (in words)=3. VGNF: Non-finite verb chunk.

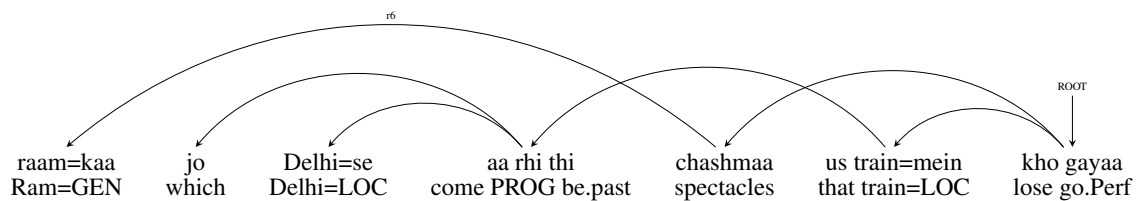


Figure 5: 'Ram's spectacles were lost in the train which was coming from Delhi'. Non-projectivity with edge degree=1, Type of intervening constituent=VGF. VGF: Finite verb chunk.

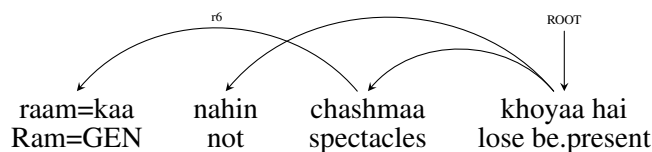


Figure 6: 'Ram's spectacles are not lost'. Non-projectivity with edge degree=1, Type of intervening constituent=NEG. NEG: Negation.

constituents becomes important when we focus on comprehension or production of non-projective structures. It is known that non-projective structures are difficult to process (Levy et al., 2012; Husain and Vasishth, 2015). In addition we also know that the type of the intervening material between a head and its dependent matters during integration stages (Levy and Keller, 2013; Safavi et al., 2016).

As mentioned in the previous section, edge degree captures the number of intervening constituents spanned by a single edge. Intervening constituents are the independent projection chains or subtrees which modify neither the dependent nor the head of a non-projective arc, rather they modify something outside the scope of the non-projective arc. The number of these intervening constituents capture the degree to which a dependent has moved from its canonical linear position.

Again, examples 4, 7, 8 have the same edge degree (1). Intuitively, we would assume 4 to be more frequent and thereby more representative of the non-projective genitive constructions. Indeed, the average length of the intervening constituents in a genitive construction is 4. We therefore expect that *the length (in words) of the intervening constituents* will be highly constrained by the type of linguistic construction in which non-projectivity occurs. Therefore, it might be beneficial to use this as a constraint in our understanding of non-projective constructions. The larger the size of intervening constituents, the more difficult it will be to process the non-projective structure for the native speaker. Indeed, this short-dependency intuition is backed by research in psycholinguistics where it has been shown that cross-linguistically dependent-head distance tends to be short (Futrell et al., 2015). More recently, Liu et al. (2017) have argued for dependency minimization as a universal cognitive constraint. This idea has also been extended to explain the occurrence of non-projectivity across multiple languages (Gómez-Rodríguez, 2017).

2.2 Hierarchical measure

The two measures discussed in the previous section, viz., the type of intervening constituent and the length (in words) of the intervening constituents do not capture an important feature of a dependency tree, i.e., the hierarchical distance. In particular we are interested in measuring the hier-

archical distance between the head of the intervening material (in a non-projective arc) and the head of the non-projective arc.

We illustrate this using examples 9–12. Examples 9 and 10 have the same type of intervening constituents and the same number of intervening constituents. But they differ with respect to the difference between the hierarchical position (or depth) of the head node of the non-projective arc (*yah* ‘this’) and the depth of the head of the intervening material (matrix verb). In 9, this depth difference is 1, while in 10, the difference is 2. In Figures 11 and 12 the depth difference is even higher. Interestingly, the sentences in figures 11 and 12 are less acceptable for Hindi native speakers.³

We propose a measure to capture the constraints on non-projectivity in terms of the *hierarchical depth difference* between the head of the non-projective arc and the head of the intervening constituent. It is evident from the examples in figures 9–12 that this measure captures the level of embedding of the non-projective arc. If the non-projective subtree is deeply embedded in the tree and the intervening constituent has a head that is higher up in the tree, we posit that the acceptability or grammaticality of the non-projective configurations will be determined by the notion of depth difference. Indeed, it has been previously shown that more embeddings in a sentence leads to processing difficulty (Gibson and Thomas, 1999).

Figure 13 shows a schematic of the environment of a non-projective dependency; X_d represents the dependent, X_h represents the head, X_i represents the intervening constituent whose head X_j is outside the span of the subtree headed by X_h . Based on the discussion in the previous sections, the constraining environment of a non-projective dependency will therefore contain the following:

- (a) Type of intervening constituent X_i
- (b) The length (in words) of the intervening constituents

³We note that the acceptability of 11 in comparison with 12 might be explained via the increased head-dependent distance in 12. However, a construction with the same head-dependent distance as 12 but with a lower depth difference (of 1) may be perfectly acceptable. An example of such a sentence would be *nalin yah do dinon se [logon ko kahte [chale jaa rahaa hai]] ki jaggu chor hai* ‘Nalin is continuously saying this to people for last two days that Jaggu is a thief’. This shows that in these cases, depth and not the linear distance is leading to lower acceptability.

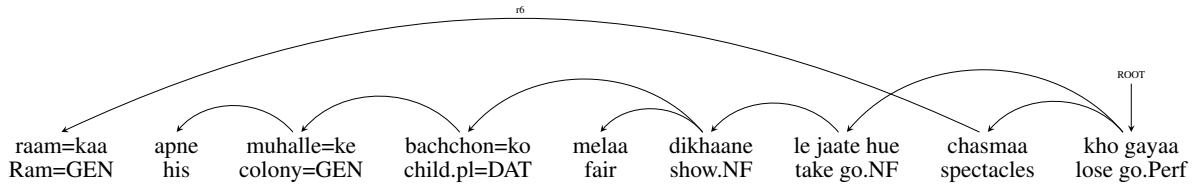


Figure 7: 'Ram's spectacles were lost while taking the children of his colony to see the fair'. Non-projectivity with edge degree=1, Length of intervening constituents (in words)=10

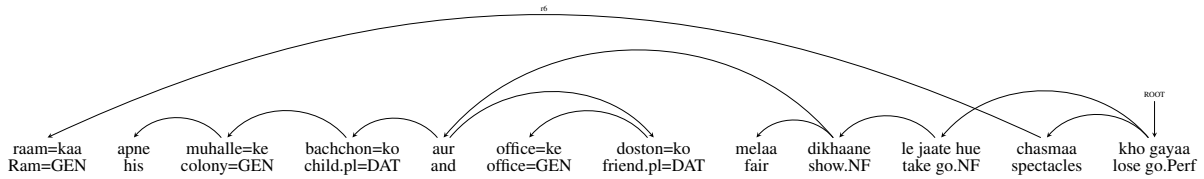


Figure 8: 'Ram's spectacles were lost while taking the children of his colony and (his) office friends to see the fair'. Non-projectivity with edge degree=1, Length of intervening constituents (in words)=15

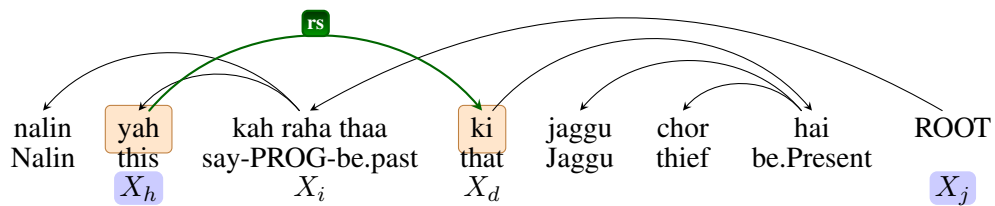


Figure 9: 'Nalin was saying that Jaagu is a thief'. Clausal complement with nominal head. Length of intervening constituents=3, Hierarchical depth difference=1

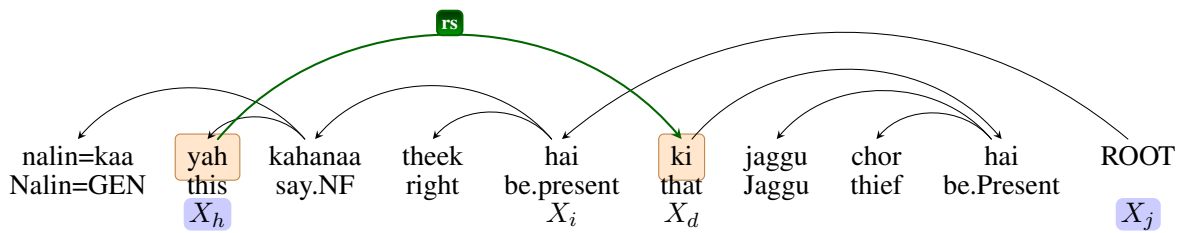


Figure 10: 'Nalin's saying that Jaggu is a thief is right'. Clausal complement with nominal head embedded in non-finite clause. Length of intervening constituents=3, Hierarchical depth difference=2

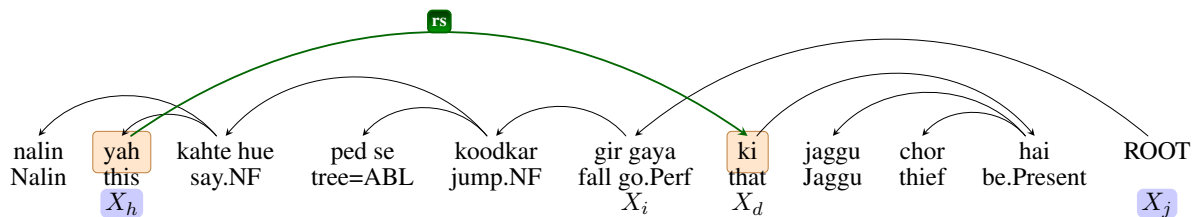


Figure 11: 'While saying that Jaggu is a thief Nalin jumped and fell from the tree'. Hierarchical depth difference=3

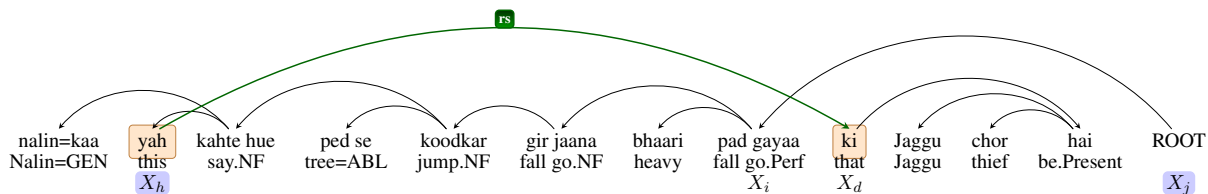


Figure 12: ‘Nalin’s jump and fall from the tree while saying that Jaggu is thief was a loss for him’. Hierarchical depth difference=4

(c) The hierarchical depth difference between X_h and X_j .

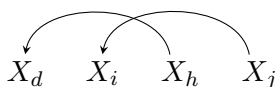


Figure 13: A non-projective configuration.

To summarize, the three measures of non-projectivity discussed above try to incorporate insights from the processing literature. Specifically, they try to reflect the notion of structural complexity and its interaction with working memory constraints. Assuming that a treebank is representative of the grammar of a language and its usage and that the proposed measures indeed reflect certain cognitive constraints, one hopes to observe some evidence for these measures while characterizing non-projectivity found in the treebank.

The measure in (a) is operationalized by using the syntactic constituent label (NP, CCP, VGNN or RBP etc.) of the intervening constituent. In order to define the property of these constituents we also see whether they are arguments or adjuncts (this information can be derived from the dependency labels).

The Hindi-Urdu Dependency Treebank (HUTB ver-0.05) was used to compute these constraints. We use the inter-chunk dependency information to extract dependency relations for a sentence. The treebank consists of 20931 sentences (Average word count per sentence: 20). The text in the treebank belongs to two genres: News (18857 sentences) and Conversation (2074 sentences). The news genre contains articles from a Hindi newspaper while the conversation has literary pieces containing dialogues.

3 Non-projectivity measures for the Hindi Treebank

In this section we try to uncover the constraining environment in which a phenomenon can occur in a non-projective configuration. In order to posit such constraints, we will use the three measures discussed in section 2: the nature of the intervening constituent, the linear distance between the head and the dependent, and the hierarchical depth difference. A constraining environment should help us in a deeper understanding of non-projectivity in a phenomenon independent of the annotation scheme. Out of the total non-projective sentences in HUTB, there are 15.4% cases that are non-projective due to annotation choices. We do not consider these cases in our analysis.

Many constructions become non-projective because of variation in word order. The word order variation could have discourse functions (Butt and King, 1996; Kidwai, 2000; Kothari, 2010). It is implied that one can projectivize these constructions by rearranging the words in their ‘canonical’ position. In our analysis we examine such non-projective constructions using the constraints shown in Table 1. As mentioned above, our analysis disregards the cases that are non-projective because of certain annotation choices in the treebank.

3.1 Type of the intervening constituent

Among GENITIVES, the most common type of intervening constituent is a nominal adjunct (67.7%). However, the intervening element in genitives can occasionally be non-nominal (like a conjunction, finite verb, non-finite verb etc.). Similarly, in NON-FINITE CLAUSE CONSTRUCTIONS, the intervening elements are nominal adjuncts (83%). The non-projective COORDINATION CONSTRUCTIONS and FINITE CLAUSE CONSTRUCTIONS are quite constrained with respect to the nature of intervening element. A coordination

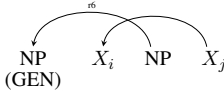
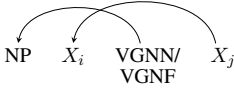
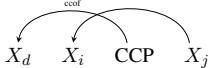
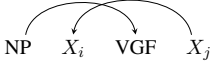
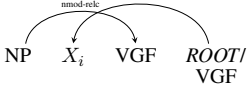
Linguistic Phenomenon	Properties of the Intervening Constituents X_i			% non-proj	
	Category(X_i)	Avg-length(X_i)	Arg-Adj(X_i)	within	across
Genitive 	NP(67.7%) CCP(9.3%) RBP(8.9%) VGF(7.2%) VGNN(3%) VGNF(2.4%)	4	Argument(25.7%) Adjunct(74.3%)	1.13%	7.2%
Non-finite Clause 	NP(83%) RBP(6.6%) CCP(2.8%) VGF(2.2%)	4	Argument(31.5%) Adjunct(68.5%)	1.2%	4.6%
Coordination 	NP(66.7%) CCP(33.3%)	3	Argument(38.5%) Adjunct(61.5%)	0.2%	0.5%
Finite Clause 	CCP(84.5%) VGF(7.7%) NP(7.7%)	2	Argument(1.9%) Adjunct(98.1%)	0.3%	2.8%
Relative Clause 	VGF(94.5%) NP(4.3%) CCP(0.5%) VGNF(0.3%) VGNN(0.1%)	3	Argument(2.9%) Adjunct(97.1%)	59.4%	23.7%

Table 1: Constraining environment for non-projectivity due to non-canonical word order. The data is taken from the News genre. Here $\text{Category}(X_i)$ represents the phrasal category of the intervening constituents, $\text{Arg-Adj}(X_i)$ represents whether an intervening element is either an argument or an adjunct and $\text{Avg-length}(X_i)$ is the average length of intervening constituent(s). The % non-proj within construction means the percentage of non-projective constructions out of total constructions of a specific type say Genitive. The % non-proj across all constructions means the percentage of non-projective cases of a specific construction type out of total non-projective cases in the treebank. NP: Noun chunk, CCP: Conjunction chunk, VGNF: Non-finite verb chunk, VGNN: Verbal noun chunk, VGF: Finite verb chunk.

subtree becomes discontinuous because of a noun (66.7%) or noun-noun conjunction (33.3%). A finite clause becomes non-projective due to a paired connective (84.5%). This happens when the connective *agar* ‘if’ moves from its canonical sentence-initial position and intervenes between the finite verb and its modifiers. RELATIVE CLAUSE CONSTRUCTIONS have finite verbs as the intervening element due to right extraposition of the relative clause (94.5%), other types of intervening elements like noun, conjunctions, non-finite verb are rarely found. The dominant pattern that emerges from this is that when something intervenes within a dependency span to make it non-

projective, it is more likely for it to be simple (e.g. noun phrase) than complex (e.g. clause).

3.2 Length of the intervening constituents

The head-dependent distance i.e. the length of the intervening constituents (in words) will vary across linguistic phenomena. The head-dependent distance is contingent on ‘the size of the projection chain of an intervener’. The GENITIVE and NON-FINITE CLAUSE CONSTRUCTIONS have an average head-dependent distance of 4 words. However, in cases where a genitive construction allows an embedded non-finite clause and coordinated non-finite clause as intervening elements,

the distance between the head and dependent can get quite large (up to 15 intervening words) as compared to the average of 4 words. The FINITE CLAUSE CONSTRUCTION has the average head-dependent distance of 2 words as they become non-projective due to a paired connective (which is just a single word).

3.3 Rightward scrambling & extraposition

Although the leftward scrambling of genitive noun (i.e. the genitive noun still remains to the left of its head) is more common among genitive constructions, rightward scrambling of dependent genitive noun is also observed in the treebank. Example (a) in Figure 14 shows the genitive marked noun *raam=kaa* ‘Ram GEN’ appearing after the copula *hai* ‘is’. A similar kind of rightward scrambling causing non-projectivity is observed in case of NON-FINITE CLAUSES, where a modifier of non-finite verb is scrambled to the right of the main verb (see example (b) of Figure 14). In both cases, the scrambling could happen because the subtree headed by this noun is ‘heavy’ due to a relative clause modification. Such a heavy NP shift should be seen whenever the noun subtree becomes large. Non-projectivity due to right extraposition is very common in relative clause constructions in the treebank (see example (c) of Figure 14). Recent work in processing suggests that extraposition of Hindi relative clauses is highly constrained (Kothari, 2010). Together, these rightward scrambling and right-extraposition support the influence of working memory constraints during processing (Wasow, 1997; Gibson, 2000; Lewis and Vasishth, 2005).

For the construction types discussed above, we assumed that their projective counterparts had the canonical word order; the non-canonical word order in such constructions led to non-projectivity. There are some clausal complement constructions that are ‘inherently’ non-projective, i.e., there are no projective counterparts to these constructions. The complementizer is headed not by a finite verb, but a noun or a pronoun; an example – *mohan ne yah bataayaa ki aaj masterji school nahin aayenge* ‘Mohan said that the teacher will not come to school today’, where *yah* ‘this’ is the head of the clausal complement headed by *ki* ‘that’. Out of all the clausal complements in the treebank 67.3% are of this type.

A few linguistic phenomena in the treebank are

non-projective due to certain annotation choices. One such construction is the conditional or paired connective. Certain types of argument structure alternations with respect to complex predicates also become non-projective due to annotation choices. We do not include these cases in our analysis or in the computation of the non-projectivity measures. Such cases make up 15.4% of the total non-projective sentences in the treebank.

So far, we have been discussing non-projectivity using two of the three constraints that were introduced in Section 2. We will now discuss non-projectivity with respect to our third constraint, hierarchical depth difference.

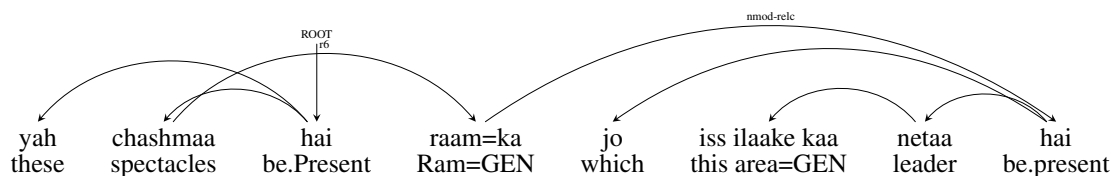
3.4 Hierarchical depth difference

Linguistic Phenomenon	Depth Difference (no. of heads)			
	n=1	n=2	n=3	n>3
Genitive	88.6%	9.7%	1.7%	-
Non-finite clause	53.5%	38.2%	8.3%	-
Coordination	61.1%	38.9%	-	-
Finite Clause	18.4%	75.5%	5.1%	1%
Relative Clause	55.9%	42.2%	0.1%	1.6%
Clausal Complement	64.7%	34.9%	-	0.3%

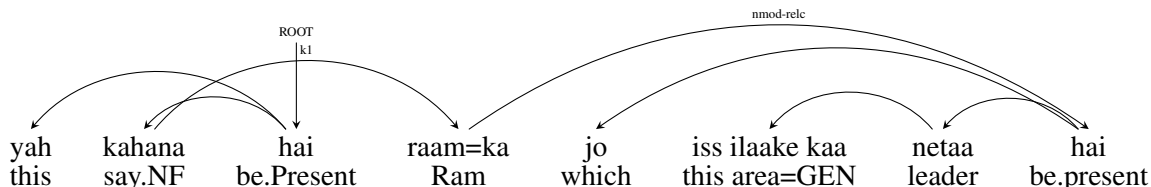
Table 2: The depth difference constraint on non-projectivity across constructions (in the news genre)

As shown in table 2, as the depth difference increases, the no. of non-projective constructions decreases. Recall that this measure captures the level of embedding of the non-projective arc in the dependency tree. If the non-projective subtree is deeply embedded in the tree and the intervening constituent has a head that is higher up in the tree, we posit that the acceptability or grammaticality of the non-projective configurations will be determined by the notion of depth difference. This seems to be validated by the data and is consistent with previous work that has shown the cost of embedding during processing (Gibson and Thomas, 1999), also see Yngve (1960). In fact, since non-projectivity is costly, we could predict that non-projectivity at a larger depth difference will be extremely difficult to process.

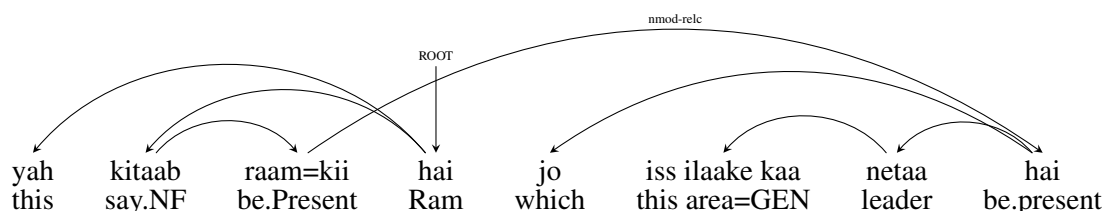
It is interesting to note that the difference in percentage of non-projectivity across various depths (cf. table 2) is not the same. While the no. of non-projective constructions reduce dramatically as depth difference increases in the case of genitives, this is not true for relative clauses. Non-finite clause constructions frequently have depth differ-



Example (a): ‘These spectacles are Ram’s, who is the leader of this area’



Example (b): ‘This is Ram’s saying, (the one) who is the leader of this area’



Example (c): ‘This is Ram’s book, (the one) who is the leader of this area’

Figure 14: Non-projectivity due to extraposition/scrambling of a dependent to the right of the verb in (a) Genitive, (b) Non-finite clause, (c) Relative clause construction.

ence of 2. Such constructions allow shared arguments to be embedded inside a non-finite clause, which is in turn within another non-finite clause. Interestingly, there is a considerable number of non-projective cases at $n=3$ for the non-finite and finite clause constructions. It is very rare to have non-projectivity for depth >3 . Finally, clausal complements allow a depth difference of up to 5. They allow a chain of embedded non-finite clauses inside the main clause, which increases the depth of embedded head of non-projective subtree.

3.5 Differences across genre

The news data has 18.36% non-projective sentences (3457 sentences) while conversation data has 11.14% cases of non-projectivity (231 sentences). This is surprising since one would assume conversation data to allow for more word order variation. While this requires further research, we found a considerable difference between the two genres (News vs Conversation) with respect to non-projectivity for some of the linguistic phenomena. In case of NON-FINITE CLAUSE CONSTRUCTIONS, it is more common for the intervening constituent to be an argument in the conversation data (71%) compared to the news data

(31.5%). The rightward scrambling of a genitive noun is highly productive in the conversation section of the treebank, making up 33% of all non-projective genitive constructions. This implies that speaker tends to move large phrases rightward (heavy NP shift) to minimize the dependency length in a sentence (Wasow, 1997).

Also, the maximum depth difference for genitives in the conversation data was 1, while in the news data this was 3 (cf. table 2). This points to a possibility that non-projectivity of this kind is simpler in conversation data.

Interestingly, the total number of non-projective RELATIVE CLAUSE CONSTRUCTIONS is half the amount in conversation (26.6%) as compared to news (59.4%). This is due to the frequent occurrence of relative-correlative constructions in the conversation data which are projective. E.g. (i) *ye dost jinse tumhe nafrat hai, vahi ek din tumhare kaam aayenge* ‘These friends whom you hate, they will help you one day’ (ii) *jisko kal tumne kitaab di thi, vah ladkaa aaj skool nahin aayaa* ‘To whom you gave the book yesterday, that boy did not come to school today’. Also, embedded relative clauses, which are projective, are frequent

in the conversation data. Together, these patterns support a well known claim in the production literature that syntactic choices are predominantly determined by production ease (MacDonald, 2013, amongst others), also see Arnold (2011).

4 Summary and Conclusion

This paper was an attempt to use certain novel measures to understand non-projectivity in Hindi. These measures were informed by the processing literature that has tried to formalize the notion of linguistic complexity using working memory constraints. The three measures, namely, the type of intervening constituent, its length, and the hierarchical depth, tried to capture and characterize the nature and complexity of non-projectivity in various phenomena. One would assume that overall non-projective structures will be less complex. These measures show that this is indeed true; on average the nature of intervening phrase is simple, the length of this phrase is not very large and the depth difference is small. In addition we also find support for the role of production ease in the data of the conversational genre compared to that of the news genre. It would be interesting to see the efficacy of the proposed measures across multiple languages. We intend to do this in the near future. We also hope to investigate if the proposed measures have any relevance for computational parsing.

References

- JE Arnold. 2011. Ordering choices in production: For the speaker or for the listener. *Language from a cognitive perspective: Grammar, usage, and processing*, pages 199–222.
- Riyaz Ahmad Bhat and Dipti Misra Sharma. 2012. Non-projective structures in indian language treebanks. In *Proceedings of TL11*, pages 25–30.
- Rajesh Bhatt, Bhuvana Narasimhan, Martha Palmer, Owen Rambow, Dipti Misra Sharma, and Fei Xia. 2009. A multi-representational and multi-layered treebank for hindi/urdu. In *Proceedings of the Third LAW*, pages 186–189.
- Manuel Bodirsky, Marco Kuhlmann, and Mathias Möhl. 2005. Well-nested drawings as models of syntactic structure. In *In Tenth Conference on Formal Grammar and Ninth Meeting on Mathematics of Language*, pages 88–1. University Press.
- M. Butt and T. C. King. 1996. Structural topic and focus without movement. In *M. Butt and T. H. King, eds., The First LFG Conference. CSLI Publications*.
- N. Chomsky and G. A. Miller. 1963. Introduction to the formal analysis of natural languages. In *R. D. Luce, R. R. Bush, E. Galanter (Eds.), Handbook of Mathematical Psychology*, volume 2, pages 269–321. Wiley, New York.
- N. Chomsky. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- L. Frazier. 1985. Syntactic complexity. In *D. Dowty, L. Karttunen, A. Zwicky (Eds.), Natural Language Processing: Psychological, Computational and Theoretical Perspectives*, volume 37, pages 129–189. Cambridge University Press, Cambridge, UKWiley Online Library.
- R. Futrell, K. Mahowald, and E. Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, 112(33):10336–10341.
- Edward Gibson and James Thomas. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes*, 14(3):225–248.
- Edward Gibson. 2000. Dependency locality theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita, and Wayne O’Neil, editors, *Image, Language, brain: Papers from the First Mind Articulation Project Symposium*. MIT Press, Cambridge, MA.
- Carlos Gómez-Rodríguez. 2017. On the relation between dependency distance, crossing dependencies, and parsing: Comment on dependency distance: a new perspective on syntactic patterns in natural languages by haitao liu et al. *Physics of Life Reviews*.
- R. Hudson. 2010. *An introduction to Word Grammar*. Cambridge University Press.
- S. Husain and S. Vasisht. 2015. Non-projectivity and processing constraints: Insights from hindi. In *Proceedings of the Third Depling*, pages 141–150.
- Aravind K. Joshi. 1985. Tree adjoining grammars: how much context-sensitivity is required to provide reasonable structural descriptions? In David R. Dowty, Lauri Karttunen, and Arnold Zwicky, editors, *Natural Language Parsing*, pages 206–250. Cambridge University Press, Cambridge.
- Aravind K. Joshi. 1990. Processing crossed and nested dependencies: An automaton perspective on the psycholinguistic results. *Language and Cognitive Processes*, 5:1–27.
- A. Kidwai. 2000. *XP-Adjunction in universal grammar: Scrambling and binding in Hindi- Urdu*. Oxford University Press, New York.
- A. Kothari. 2010. *Processing Constraints And Word Order Variation In Hindi Relative Clauses*. Ph.D. thesis, Stanford University.

- Marco Kuhlmann and Joakim Nivre. 2006. Mildly non-projective dependency structures. In *Proceedings of COLING-ACL*, pages 507–514.
- Marco Kuhlmann. 2007. *Dependency Structures and Lexicalized Grammars*. Ph.D. thesis, Saarland University.
- Roger Levy and Frank Keller. 2013. Expectation and Locality Effects in German Verb-final Structures. *Journal of Memory and Language*, 68(2):199–222.
- R. Levy, E. Fedorenko, M. Breen, and E. Gibson. 2012. The processing of extraposed structures in English. *Cognition*, 122(1):12–36.
- R. L. Lewis and S. Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29:1–45.
- Haitao Liu, Chunshan Xu, and Junying Liang. 2017. Dependency distance: a new perspective on syntactic patterns in natural languages. *Physics of Life Reviews*.
- Maryellen MacDonald. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology*, 4:226.
- P. Mannem, H. Chaudhry, and A. Bharati. 2009. Insights into non-projectivity in hindi. In *Proceedings of ACL-IJCNLP 2009 SRW*, pages 10–17.
- Joakim Nivre. 2006. Constraints on non-projective dependency parsing. In *EACL*.
- Joakim Nivre. 2009. Non-projective dependency parsing in expected linear time. In *Proceedings of ACL and IJCNLP, ACL '09*, pages 351–359.
- O. Rambow. 2010. The simple truth about dependency and phrase structure representations: An opinion piece. In *Proceedings of ACL*, pages 337–340.
- Molood S. Safavi, Samar Husain, and Shravan Vasishth. 2016. Dependency resolution difficulty increases with distance in persian separable complex predicates: Evidence for expectation and memory-based accounts. *Frontiers in Psychology*, 7:403.
- S.M. Shieber. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8:333–343.
- M. Steedman. 2000. *The Syntactic Process*. Cambridge, MA: MIT Press.
- T. Wasow. 1997. Remarks on grammatical weight. *Language Variation and Change*, 9:81–105.
- Victor H Yngve. 1960. A model and an hypothesis for language structure. *Proceedings of the American philosophical society*, 104(5):444–466.