

Understanding constraints on non-projectivity using novel measures

Himanshu Yadav¹ Ashwini Vaidya² Samar Husain²

¹Jawaharlal Nehru University, New Delhi, India

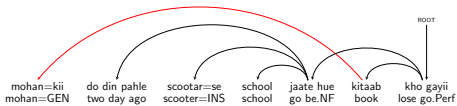
²Indian Institute of Technology Delhi, India

International Conference on Dependency Linguistics
Pisa, Italy. September 18-20, 2017

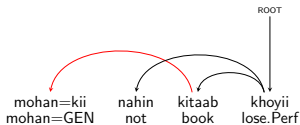
Motivation



Acceptable



Odd



Ungrammatical

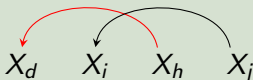


Odd

Proposed measures

Intervener Length

- Number of words between the dependent (X_d) and the head (X_h) of the non-projective arc such that they are not dominated by (X_h).



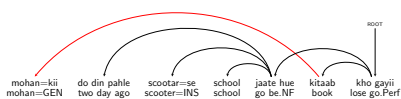
Intervener length = size of the intervening constituent (X_i).

1. Intervener type
2. **Intervener length**
3. Depth difference

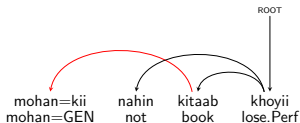
Proposed measures cont.



Acceptable (Edge degree=1, Intervener length=1, Intervener type=NP, Depth difference=1)



Odd (Edge degree=1, Intervener length=8, Intervener type=VGNF, Depth difference=1)



Ungrammatical (Edge degree=1, Intervener length=1, Intervener type=NEGP, Depth difference=1)



Odd (Edge Degree=1, Intervener length=2, Intervener type=NP, Depth difference=2)

Results


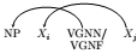



Linguistic Phenomenon	Properties of Non-projectivity			% non-proj	
	Category(X_i)	Avg-length(X_i)	Depth Diff.	within	across
Genitive 	NP(67.7%) CCP(9.3%) RBP(8.9%) VGF(7.2%) VGNN(3%) VGNF(2.4%)	4	[n=1](88.6%) [n=2](9.7%) [n=3][1.7%]	1.13%	7.2%
Non-finite Clause 	NP(83%) RBP(6.6%) CCP(2.8%) VGF(2.2%)	4	[n=1](53.5%) [n=2](38.2%) [n=3](8.3%)	1.2%	4.6%
Coordination 	NP(66.7%) CCP(33.3%)	3	[n=1](61.1%) [n=2](38.9%)	0.2%	0.5%
Finite Clause 	CCP(84.5%) VGF(7.7%) NP(7.7%)	2	[n=1](18.4%) [n=2](75.5%) [n=3](5.1%)	0.3%	2.8%
Relative Clause 	VGF(94.5%) NP(4.3%) CCP(0.5%) VGNF(0.3%) VGNN(0.1%)	3	[n=1](55.9%) [n=2](42.2%)	59.4%	23.7%

Table 1: NP: Noun chunk, CCP: Conjunction chunk, VGNF: Non-finite verb chunk, VGNN: Verbal noun chunk, VGF: Finite verb chunk.

Summary

- Although certain linguistic phenomenon can become non-projective, the range of allowed non-projective configurations is limited.
- This limitation is influenced by interaction between processing constraints and syntactic complexity.
- We have shown using three novel measures that non-projective configurations tend to be simple than complex.