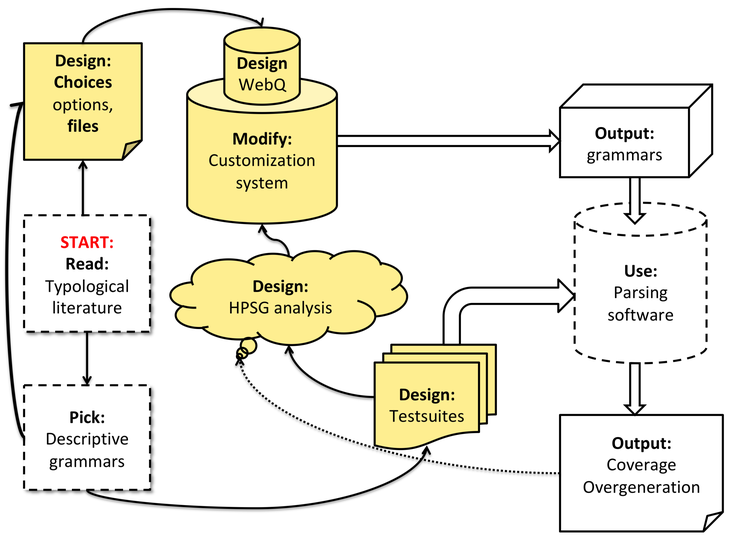
A Cross-Linguistic Account of Subordinator and Subordinate Clause Position

Olga Zamaraeva, Kristen Howell & Emily M. Bender

University of Washington, Department of Linguistics

Adding a library to the LinGO Grammar Matrix (Bender et al., 2002, 2010)



Library: Declarative objectival clausal complements

(1) Kim thinks [that Sandy left]. [eng]

Word order variation: Extraposition: Examples from Malagasy [mlg], Dryer (1980)

(2) na-mono an-dRabe Rakoto
PAST-hit ACC-Rabe Rakoto
'Rakoto hit Rabe.' [mlg]

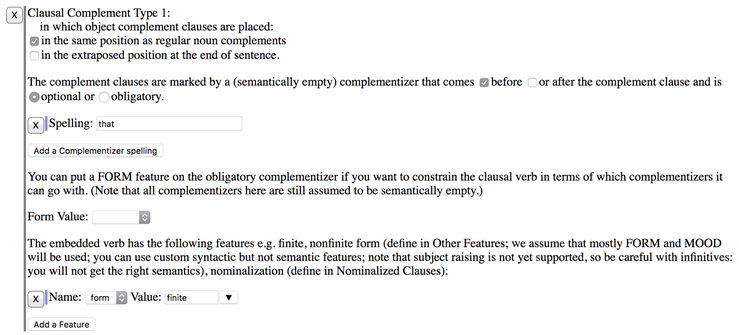
(3) Mihevitra Rabe [fa mitady ny zaza Rasoa]
thinks Rabe [COMP look.for the child Rasoa]
'Rabe thinks that Rasoa is looking for the child.' [mlg]

Order of subordinator and subordinate clause: Examples from Uzbek [uzb], Noonan (2007)

(4) Men bilamen [**ki** bu ɔdam ǰoǰa-ni oğirladi]
I know-1sg [COMP this man chicken-OBJ stole-3sg]
'I know that the man stole the chicken.' [uzb]

(5) Xɔtin [bu ɔdam ǰoǰa-ni oğirladi **deb**] dedi
woman [this man chicken-OBJ stole COMP] said.3SG
'The woman said that the man stole a chicken.' [uzb]

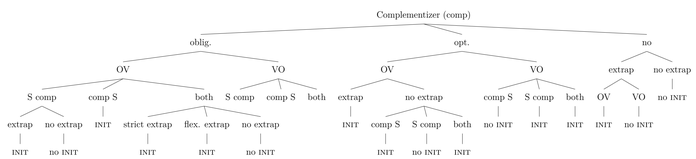
Web Questionnaire: 2300 possible combinations, 161 related to word order



Analysis

1. Rule arity and daughter order fixed in DELPH-IN formalism: need additional phrase structure rule
2. Use binary features INIT and EXTRA to avoid overgeneration
3. Main challenge: emit correct sets of feature structures and constraints in a large space of possible combinations of user choices

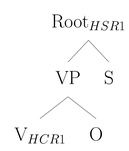
Sample customization logic: the INIT feature



Sample analysis: VOS with extraposition

$$\begin{bmatrix} \textit{HSR1} & \\ \text{SBJ} & \boxed{1} \\ \text{HD} & \boxed{2}\begin{bmatrix}\text{SBJ} & \langle \boxed{3} \rangle \oplus \boxed{1} \\ \text{COMPS} & \langle\,\rangle\end{bmatrix} \\ \text{NHD} & \boxed{3} \\ \text{ARGS} & \langle\, \boxed{2}\,,\, \boxed{3}\,\rangle \end{bmatrix} \quad \begin{bmatrix} \textit{HCR1} & \\ \text{COMPS} & \boxed{1} \\ \text{HD} & \boxed{2}\begin{bmatrix}\text{COMPS} & \langle \boxed{3} \rangle \oplus \boxed{1}\end{bmatrix} \\ \text{NHD} & \boxed{3}\begin{bmatrix}\text{EXTRA} & -\end{bmatrix} \\ \text{ARGS} & \langle\, \boxed{2}\,,\, \boxed{3}\,\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textit{HSR2} & \\ \text{SBJ} & \boxed{1} \\ \text{HD} & \boxed{2}\begin{bmatrix}\text{SBJ} & \langle \boxed{3} \rangle \oplus \boxed{1} \\ \text{COMPS} & \langle \begin{bmatrix}\text{EXTRA} & +\end{bmatrix}\rangle\end{bmatrix} \\ \text{NHD} & \boxed{3} \\ \text{ARGS} & \langle\, \boxed{2}\,,\, \boxed{3}\,\rangle \end{bmatrix} \quad \begin{bmatrix} \textit{HCR2} & \\ \text{COMPS} & \boxed{1} \\ \text{HD} & \boxed{2}\begin{bmatrix}\text{SBJ} & \langle\,\rangle \\ \text{COMPS} & \langle \boxed{3} \rangle \oplus \boxed{1}\end{bmatrix} \\ \text{NHD} & \boxed{3}\begin{bmatrix}\text{EXTRA} & +\end{bmatrix} \\ \text{ARGS} & \langle\, \boxed{2}\,,\, \boxed{3}\,\rangle \end{bmatrix}$$



Test-driven development: 100% cov. / 0% overgen.

Language	iso639	family	WO	comp	order	extrap	# strat	pos	neg
Russian	rus	IE	free	opt	comp S	-	3	6	11
German	deu	IE	V2/V-fin	oblig[1]	comp S.	-	1	6	4
Tagalog	tgl	Astronesian	V-in	oblig	comp S	flexible	1	3	4
Lango	laj	Nilo-Saharan	SVO	oblig	comp S	-	3	4	4
Turkish	tur	Turkic	SOV	opt	both	strict	4	7	9

Evaluation (held-out): 88% cov. / 0% overgen.

Language	iso639	family	WO	comp	order	extrap	Cov.	Ovg.
Jalkunan	bxl	Nig-Cong	SOV	opt	comp S	strict	4/8	0/12
Paresi-Haliti	pab	Arawak	SOV	-	-	strict	4/4	0/6
Yakima Sahaptin	yak	Plateau-Penutian	free	-	-	-	10/10	0/6
Modern Hebrew	heb	Afro-Asiatic	SVO	oblig	comp S	-	2/2	0/9
Wangkangurru	wgg	Pama-Nyungan	free	-	-	-	10/10	0/3

Error analysis: Ex. from Jalkunan [jao] (Heath, 2017)

(6) mā ń sɔ̀ [[mā jè^{n}] sɛ̀é]
1Sg **3SgNonhObj** know.Pfv [[1Sg father] come.Pgv]
'I know that my father has come.' [jao]

Conclusions

- We present a cross-linguistic account of the order of subordinator and subordinate clause in the context of new Grammar Matrix libraries for clausal complements and clausal modifiers (the latter not included in this poster; see Howell and Zamaraeva (in press))
  - Evaluation (includes other phenomena such as nominalization and morphological marking) revealed one strategy we did not account for
- Main challenge: output streamlined grammars from a large space of typological possibilities

Future Work

- Interactions with other libraries
- Subjectival clauses
- Wh-complements

URLs

Testuites and grammars: `https://students.washington.edu/olzama/gp2.html`
Grammar Matrix: `http://matrix.delph-in.net/customize/matrix.cgi`

Acknowledgements

Based upon work supported by the National Science Foundation under Grant No. BCS-1561833 (PI Bender).

References

Emily M. Bender, Dan Flickinger, and Stephan Oepen. 2002. The Grammar Matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In John Carroll, Nelleke Oostdijk, and Richard Sutcliffe, editors, *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, pages 8–14, Taipei, Taiwan, 2002.

Emily M Bender, Scott Drellishak, Antske Fokkens, Laurie Poulson, and Safiyyah Saleem. 2010. Grammar customization. *Research on Language & Computation*, 8(1):23–72. ISSN 1570-7075. URL http://dx.doi.org/10.1007/s11168-010-9070-1. 10.1007/s11168-010-9070-1.

Matthew Synge Dryer. 1980. The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 25(2):123–196.

Jeffrey Heath. 2017. *A grammar of Jalkunan (Mande, Burkina Faso)*. Language Description Heritage Library.

Kristen Howell and Olga Zamaraeva. in press. Clausal modifiers in the LinGO Grammar Matrix. In *COLING*.

Michael Noonan. 2007. Complementation. In Timothy Shopen, editor, *Language Typology and Syntactic Description*, volume 2. Cambridge University Press, Cambridge, UK.