

## Evaluating the Limits of a Minimalist Grammar Parser on Korean Double Relative Clauses

So Young Lee (Miami University) & Aniello De Santo (University of Utah)

**Introduction:** Models explicitly considering structure building operations from formal grammars have been shown to improve fit to behavioral (Demberg & Keller 2008) and neural (Brennan et al. 2016, Stanojević, M. et al. 2023) correlates of language comprehension. However, such symbolic approaches have mostly focused on structure building operations and ignored the known contribution of factors like the interaction of syntactic/semantic features, binding, etc. Focusing on off-line behavior, an approach combining a parser for Minimalist Grammars (MGs; Stabler 2013) with metrics measuring memory usage has been particularly successful in prediction processing preferences cross-linguistically, and across a variety of constructions (Kobele et al. 2013; Graf et al. 2017; De Santo and Lee 2022; a.o). This model (MG Parser) is of interest since MGs provide a rigorous formalization of many aspects of Chomsky's (1995) Minimalist Program (MP) — which remains a predominant framework in theoretical syntax. With this in mind, this study aims to extend the evaluation of the MG parser performance by examining its predictive accuracy over behavioral results for Korean double relative clauses (DRCs).

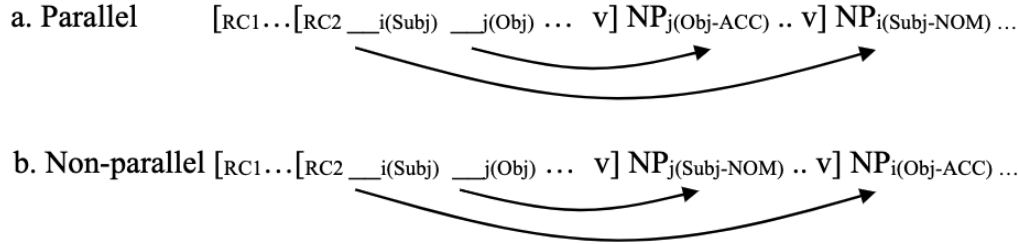
**Korean DRCs:** Korean is a head-final language with a canonical SOV order that permits flexibility through scrambling and a rich case-marking system. Relative clauses in Korean are pre-nominal without an explicit complementizer. Korean syntax permits a form of recursion within relative clauses, enabling relativization from within another clause—a phenomenon not widely observed in other languages. This nested construction (see 1) involves two fillers extracted from a single clause, each corresponding to a unique gap. Yoo & Lee (2023) demonstrate that thematic role parallelism between fillers and gaps significantly aids in resolving complex syntactic dependencies. They also identify case marking as pivotal for predicting noun phrase encoding. Findings show immediate difficulty when processing multiple nominative constructions, pointing to a selective sensitivity to case cues in language processing. According to these studies, the processing results of (2) are (2a) > (2b) > (2d) > (2c) (>: easier and faster).

**Modeling Results:** The MG model ties syntactic complexity with processing effort by mapping the memory stack states of a top-down MG parser to cognitive memory demands (Graf et al. 2017). This includes the duration that a node is retained in memory (tenure) and the length of the long distance dependencies maintained during the parse (size) — which can then be used to estimate off-line/full-sentence processing complexity. Thus, we input syntactic derivations for the sentences in (2), and evaluate the MG parser's predictions based on the complexity metrics defined by Graf et al. (2017). The model correctly predicts (2a) > (2b) and that (2b) > (2c, 2d). However, it predicts (2c) > (2d), in contrast with behavioral results.

**Discussion:** The findings accurately capture the reported parallelism effects but fail to anticipate difficulties in processing double nominative constructions. In Korean double relative clauses, this likely influences the parser's sensitivity to consecutive nominative markers. Given that the complexity of double nominative structures is more closely aligned with interference effects due to similarity rather than movement distances, the MG parser—predicated on syntactic movement and its corresponding distances—does not accurately predict these outcomes.

**Conclusion:** We show that, while the MG parser is adept at predicting processing difficulties arising from structural parallelism, it falls short when it comes to the specific challenges presented by double nominative constructions in Korean. These results are consistent with the intuition that the DRC asymmetries are due to interference effects caused by similar grammatical structures (like adjacent nominative case markers) more than to the distance over which syntactic elements move. On the way to developing comprehensive models of human sentence comprehension, approaches exploiting symbolic grammars will need to be augmented with measures of factors such as feature sensitivity (De Santo 2021) and similarity-based interference (Lewis et al. 2006).

(1) Yoo & Lee (2023; p.2)



(2) A sample set of experiment sentences (Yoo & Lee 2023)

	Role of a distractor (low head noun in low RC: object)	Role of a target (high head noun in low RC: subject)	Case mismatch	
(a)	Parallel (ACC-object)	Parallel (NOM-subject)	Mismatch (ACC-NOM)	[RC1[RC2 __i phyenci-lo kyeklyeha-yesste-n] ai <sub>i</sub> -lul machimnay manna-n] hwuwenca-ka ... letter-with encourage-PST-ADN kid-ACC finally meet-ADN sponsor-NOM...
(b)	Parallel (ACC-object)	Non-Parallel (ACC-object)	Match (ACC-ACC)	[RC1[RC2 __i phyenci-lo kyeklyeha-yesste-n] ai <sub>i</sub> -lul machimnay manna-n] hwuwenca-lul ... letter-with encourage-PST-ADN kid-ACC finally meet-ADN sponsor-ACC... 'the sponsor <sub>i</sub> [who finally met the kid <sub>j</sub> [who (the sponsor) <sub>i</sub> encouraged __j with letters]]'
(c)	Non-Parallel (NOM-subject)	Parallel (NOM-subject)	Match (NOM-NOM)	[RC1[RC2 __i phyenci-lo kyeklyeha-yesste-n] ai <sub>i</sub> -ka machimnay manna-n] hwuwenca-ka ... letter-with encourage-PST-ADN kid-NOM finally meet-ADN sponsor-NOM...
(d)	Non-Parallel (NOM-subject)	Non-Parallel (ACC-subject)	Mismatch (NOM-ACC)	[RC1[RC2 __i phyenci-lo kyeklyeha-yesste-n] ai <sub>i</sub> -ka machimnay manna-n] hwuwenca-lul ... letter-with encourage-PST-ADN kid-NOM finally meet-ADN sponsor-ACC... 'the sponsor <sub>i</sub> [who the kid <sub>j</sub> [who (the sponsor) <sub>i</sub> encouraged __j with letters] finally met]'

## References:

[1] Brennan, J. R., Stabler, E. P., Van Wagenen, S. E., Luh, W. M., & Hale, J. T. (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and language*. [2] Demberg, V. & Keller, F., 2008. Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*. [3] De Santo, A., 2021. A Minimalist Approach to Facilitatory Effects in Stacked Relative Clauses. *Proceedings of the Society for Computation in Linguistics*. [4] De Santo, A. & Lee, S.Y., 2022. Evaluating Structural Economy Claims in Relative Clause Attachment. In *Proceedings of the Society for Computation in Linguistics 2022*. [5] Graf, T., Monette, J., & Zhang, C., 2017. Relative clauses as a benchmark for Minimalist parsing. *Journal of Language Modelling*. [6] Kobele, G.M., Gerth, S. & Hale, J., 2012, August. Memory resource allocation in top-down minimalist parsing. In *International Conference on Formal Grammar*. [7] Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in cognitive sciences*. [8] Stabler, E.P., 2013. Two models of minimalist, incremental syntactic analysis. *Topics in cognitive science*. [9] Stanojević, M., Brennan, J.R., Dunagan, D., Steedman, M., & Hale, J.T., 2023. Modeling Structure-Building in the Brain With CCG Parsing and Large Language Models. *Cognitive science*. [10] Yoo, M.H. & Lee, S.Y., 2023. Interference effects on the processing of Korean double relative clauses. *Proceedings of the Linguistic Society of America*.