Hungarian relative clause processing: Diverging results in L-maze and A-maze

Wesley Orth, Dávid Nemeskey, Eszter Ronai

Relative clause processing: A core finding predicted by several sentence processing theories is that object RCs (ORCs,1) are harder to process than subject RCs (SRCs, 2). *Memory-based accounts* predict a preference for shorter verb-argument dependencies **[1-2]**. *Expectation-based accounts* predict a preference for more frequent and predictable structures **[3-4].** In English, these accounts make converging predictions. SRCs have a shorter dependency between the extracted NP and the RC verb than ORCs, meaning memory-based accounts predict easier processing for SRCs. At the same time, SRCs are more frequent in English than ORCs, so expectation-based accounts also predict easier processing for SRCs.

Cross-linguistic studies: While these accounts converge in English, this is not always true cross-linguistically. Hungarian allows for the word order in the RC to vary independently of RC type. As in (3-4), both SRC and ORC structures can be produced with the verb (bolded) in either an early or late position in the RC. But prior work on Hungarian RCs has relied on self-paced reading (SPR, **[5]**), which has relatively poor effect localization and faces reliability issues in web-based research. SPR has also failed to detect effects which have appeared in eye-tracking experiments in Russian, which features RC structures similar to Hungarian **[6-7]**.

Maze-task: To address these methodological issues and provide new cross-linguistic evidence, we investigated Hungarian RCs with the maze-task. The maze-task has been shown to replicate effects from both SPR and eye-tracking, while providing more reliability for web-based research **[8]**. It has also been previously used to investigate RC processing in English **[9]**. Experiment 1 uses L-maze, in which foils are computationally generated pseudowords. Experiment 2 uses Hungarian A-Maze, in which the foils are computationally generated to have high surprisal given the previous context. This study makes a methodological contribution by adapting computational tools for automating foil generation for both maze-tasks in Hungarian.

Experiments: Both experiments used the same 32 items in a 2x2 design manipulating the RC type (SRC/ORC) and the position of the verb in the RC (local/non-local) as in (3-4).
Memory-based accounts predict local structures should be easier at the point of the RC verb. Expectation-based accounts would predict that more frequent structures should be easier. In Hungarian, this means that SRCs should be easier at the relative pronoun and the RC verb, and non-local structures should be easier at the RC verb. The L-maze experiment was conducted with 50 speakers of Hungarian (4 excluded), while the A-Maze experiment was conducted with 81 speakers of Hungarian (10 excluded). Both datasets were collected over the web.

Results: Data were analyzed with linear mixed effect models fit over log transformed RTs. In both experiments we observe faster response times at the relative pronoun for SRCs ($p<.001$). At the point of the RC verb, we observe different effects in L-Maze than in A-Maze. In L-maze we see faster responses for local structures ($p<.05$). In A-Maze we observe faster responses for non-local structures ($p<.001$) and faster responses for SRCs ($p<.001$).

Conclusion: These results reveal critical information about RC processing and the properties of different maze tasks. First, both experiments capture a predicted effect at the relative pronoun which was not previously observed in SPR **[5]**. This suggests that the maze-task can be an important tool as we move to investigate more languages in web-based research. Second, the experiments observe different effects at the RC verb region, each predicted by different theories. L-maze finds a locality advantage, most compatible with memory-based accounts. A-Maze instead finds a non-local advantage ("anti-locality effect", i.a. **[10]**), which is more compatible with expectation-based accounts. As A-Maze is known to require more complete processing at each step **[9,11]**, it is possible that this enhanced processing leads to sharper expectations, and thus effects compatible with expectation-based accounts are observed.

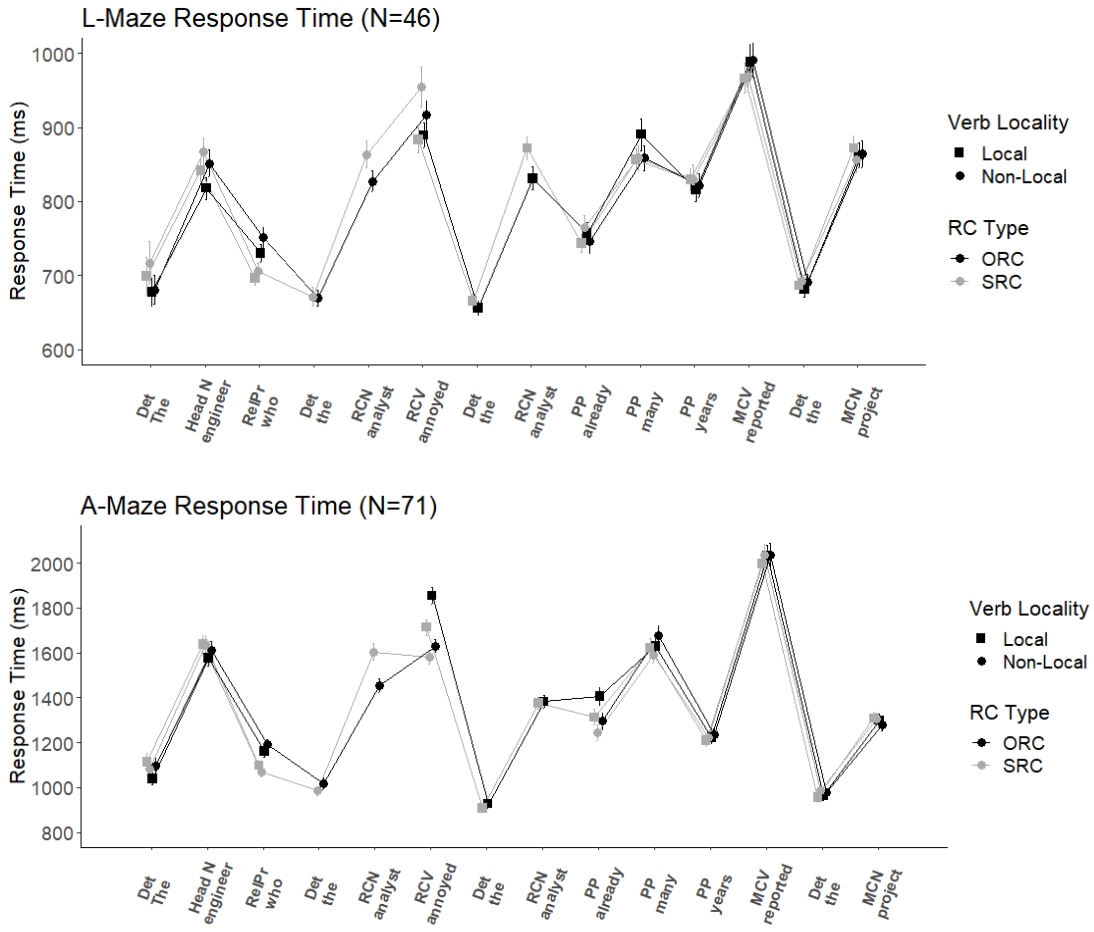
(1) The engineer who the analyst annoyed wrote a report about the project. (ORC)

(2) The engineer who annoyed the analyst wrote a report about the project. (SRC)

(3) a. A mérnök [aki **idegesítette** az elemzőt] beszámolt a projektről. (SRC|Local)
b. A mérnök [aki az elemzőt **idegesítette**] beszámolt a projektről. (SRC|Non-local)
Both: The engineer who annoyed the analyst reported on the project. (RC verb bolded)

(4) a. A mérnök [aki**t idegesített** az elemző] beszámolt a projektről. (ORC|Local)
b. A mérnök [aki**t** az elemző **idegesített**] beszámolt a projektről. (ORC|Non-local)
Both: The engineer who the analyst annoyed reported on the project. (RC verb bolded)

Figure 1 & 2: Word by word response times with English glosses.



References

[1] Gibson. 1998. Cognition. **[2]** Lewis & Vasishth. 2005. Cognitive Science. **[3]** Hale. 2001. NAACL. **[4]** Levy. 2008. Cognition. **[5]** Ronai & Xiang. 2023. Cognitive Science. **[6]** Levy, Fedorenko & Gibson. JML. **[7]** Price & Witzel. 2017. JML. **[8]** Boyce, Futrell & Levy. 2020. JML. **[9]** Forster, Guerrera & Elliot. 2009. Behavior Research Methods. **[10]** Konieczny. 2000. Journal of Psycholinguistic Research, **[11]** Witzel, Witzel & Forster. 2012. Journal of Psycholinguistic Research.