

## Processing information structural ambiguity: Contrastive Topics in Estonian

Marju Kaps (University of California, Los Angeles)

mkaps@ucla.edu

**Background.** While a large literature is devoted to incremental syntactic parsing, there is considerably less work addressing the online interaction between syntactic constraints and information structure (IS). Some syntactic representations have been argued to obligatorily encode an IS component [1] and previous work on discourse-configurational languages, e.g. [2] on Finnish, indicates that comprehenders use non-canonical word order to anticipate the discourse status of upcoming referents. The present study investigates Estonian verb-third (V3) clauses, which syntactically necessitate the presence of a preverbal Contrastive Topic (CT, see 1) [3], but allow for flexible preverbal word order. I report an apparent asymmetry in which the parser commits to object contrast upon encountering a non-canonical, clause-initial object, but appears to underspecify the placement of contrast in temporarily information structurally ambiguous subject-initial clauses. I propose that the parser forms an underspecified IS representation when encountering information structurally ambiguous material, whereby general processing biases such as preferences for structural parallelism [4] have weakened influence.

**Contrastive Topics in Estonian.** Estonian is a canonically SVO language [5] but allows V3 order (e.g. SOV and OSV) when a preverbal CT is present [6]. As the use of V3 necessitates a CT [3], the parser might be expected to assign CT status to a preverbal element during online processing. CTs need a salient alternative in the discourse, which may for instance occur in a following CT remnant stripping ellipsis clause (see [7] for German), where the grammatical case of the remnant disambiguates the intended CT of the matrix clause as a subject or object. As matrix clause-initial subjects are canonical, they constrain the information structure of the clause less than initial objects. If the parser does compute contrast status incrementally, I predict that the parser makes more syntactically fine-grained commitments to information structure in OSV clauses compared to SOV clauses. Alternatively, a structural parallelism account predicts initial CT preference regardless of word order, as the CT is clause-initial in the ellipsis clauses following.

**Speeded Acceptability Task.** Native Estonian speaking participants (N=36) judged sentences in a timed forced choice RSVP task. Experimental items crossed matrix clause word order (SOV vs OSV) with CT remnant type (Subject vs Object), in order to compare the processing difficulty associated with a disambiguating CT remnant assigning CT status to a first or second position subject or object (see 2). Half of the items were followed by comprehension questions.

**Results and Discussion.** Similar asymmetries between SOV and OSV clauses were seen in all three measures used, using GLMER or LMER models. **Acceptance rates** showed an interaction between word order and remnant type ( $p < .001$ ), with an increased clause-initial CT preference in OSV clauses compared to SOV clauses. **Reaction times to “yes” responses** showed a penalty for OSV clauses ( $p < .001$ ) and subject remnants ( $p < .001$ ), and an interaction between word order and remnant type with the OSV subject remnant condition slower than the rest ( $p < .001$ ). **Comprehension question accuracy** showed a similar penalty for second-position CTs in OSV clauses. **Pairwise comparisons** revealed no remnant type effects for the temporarily canonical SOV clauses in the three measures, suggesting that in SOV clauses, comprehenders may form a syntactically underspecified representation of contrast. OSV clauses bias the reader towards an object CT, as evidenced by increased sentence rejections, slower acceptance times and poorer comprehension question accuracy when a dispreferred subject CT is encountered, in line with clause-initial objects imposing tighter restrictions on the IS of the clause (perhaps forcing the parser to commit to a CT structure early, assigning CT status to the initial object).

**Conclusion.** The present study provides evidence that the parser rapidly computes information structure, provided that the word order of the target clause is sufficiently informative. This work raises questions to be addressed by future studies on the time course and specificity of integrating syntactic, information structural and contextual cues during online comprehension. Further, our findings highlight the value of cross-linguistic research in studying syntactic parsing.

(1) **CT example.** Do John and Bill like Mary? JOHN<sub>CT</sub> likes Mary but BILL<sub>CT</sub> doesn't <like Mary>

(2) **Sample item.** 24 experimental items. 2x2, crossing matrix order (SOV, OSV) and remnant category (Subject, Object), disambiguated by case. Presented along with 84 fillers.

**SOV, S<sub>CT</sub>: Initial subject contrast**

Ants	Jaanikat	ilmselt	ei	armasta,	Margus	aga	küll
Ants.NOM	Jaanika.PART	apparently	NEG	love	Margus.NOM	but	AFF

'Apparently Ants doesn't love Jaanika, but Margus does.'

**SOV, O<sub>CT</sub>: Second-position object contrast**

Ants	Jaanikat	ilmselt	ei	armasta,	Hellet	aga	küll
Ants.NOM	Jaanika.PART	apparently	NEG	love	Helle.PART	but	AFF

'Apparently Ants doesn't love Jaanika, but he does love Helle.'

**OSV, S<sub>CT</sub>: Second-position subject contrast**

Jaanikat	Ants	ilmselt	ei	armasta,	Margus	aga	küll
Jaanika.PART	Ants.NOM	apparently	NEG	love	Margus.NOM	but	AFF

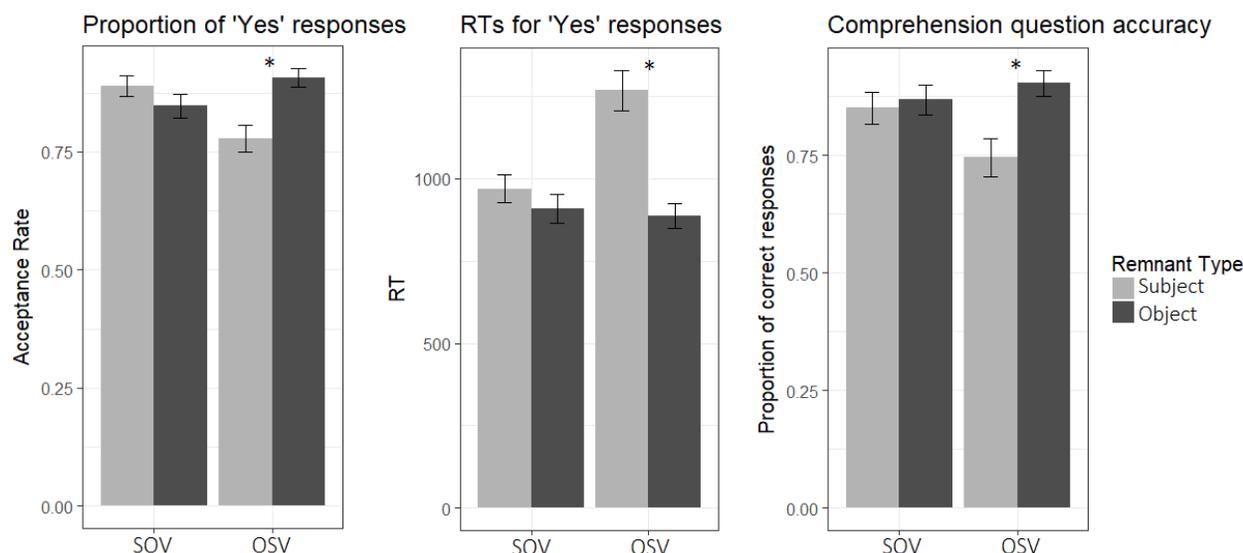
'Apparently Ants doesn't love Jaanika, but Margus does.'

**OSV, O<sub>CT</sub>: Initial object contrast**

Jaanikat	Ants	ilmselt	ei	armasta,	Hellet	aga	küll
Jaanika.PART	Ants.NOM	apparently	NEG	love	Helle.PART	but	AFF

'Apparently Ants doesn't love Jaanika, but he does love Helle.'

**Figures.** Means and standard errors shown, sig. differences in pairwise comparisons marked. The pattern indicates a penalty for object remnants after OSV clauses, due to mismatching IS.



**Selected references**

[1] Brown, M., Savova, V., & Gibson, E. (2012). Syntax encodes information structure: Evidence from on-line reading comprehension. *Journal of Memory and Language*, 66(1), 194-209. [2] Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2), 113-147. [3] Kaps, M. (in press). Licensing SOV in Estonian: A naturalness rating study, *Proceedings of CLS 54* [4] Carlson, K. (2001). The effects of parallelism and prosody in the processing of gapping structures. *Language and Speech*, 44(1), 1-26. [5] Ehala, M. (2006). The word order of Estonian: implications to universal language. *Journal of Universal Language*, 7(1), 49-89. [6] Henk, P. (2010). Information Structure of Estonian-Compared to Finnish and Hungarian. *Publications of the Institute of Cognitive Science*, 6. [7] Repp, S. (2009). *Negation in gapping* (Vol. 22). Oxford University Press.