Deeply embedded clauses in Finno-Ugric and their significance for the description of clause combining

(Edyta Jurkiewicz-Rohrbacher, Universität Regensburg, Universität Hamburg; Petar Kehayov University of Tartu; Triin Todesk University of Tartu)

A deeply embedded clause is a clause which is embedded in a clause that itself is embedded. Sentences with multiple embeddings can be formalized [C [C1 [C2 [C3 […]]]]], where C = main clause, C1 = first-order embedded clause, C2 = second-order embedded clause, C3 = third-order embedded clause, etc. Deeply embedded is any clause below C1; see (1) from Moksha Mordvin, where the deepest clause is C2.

(1) [$_{C=}$*Rasijäń praviťəĺstvaś*   *vešś,*   [$_{C1=}$*štoba ťä bańďiť*
     Russian   government:DEF.NOM   request:PST1.3SG that:IRR this bandit:DEF.ACC
     uskəĺəż   *Rasijäv,*   [$_{C2=}$*kosa ušədəĺχť*   *soń*   *ravža*
     send:CONJ:S3PL>O3SG   Russia:LAT where   begin:CONJ:3PL   he:GEN   black
     ťevńänzənkoŕas   *sletstvijä,*   *a*   *śäĺďä suďəndaləż*]]].
     deed:POSS.3SG.PL:GEN   for   investigations   and then   convict:CONJ:S3PL>O3SG
     (Moksha Mordvin, CCLM)
     'The Russian government insisted that the bandit should be sent to Russia, where court investigations will be opened for his miserable deeds, and he would be convicted.'

The properties of deeply embedded clauses (DECs) and their grammatical relations to other parts of the sentence are poorly studied because of the idea of *recursion* as a fundamental property of language. If recursive embedding leads to reproduction of the same structure (e.g. van der Hulst 2010), we cannot expect to find something new in DECs compared to first-order embedded clauses. Therefore, grammars and even special studies on complex sentences consider the issue of clausal subordination exhausted with the description of C1s.

Based on corpora of written language texts (CCLM; ENC; KKK; KORP), we study the properties of DECs in three Finno-Ugric languages: Estonian, Moksha Mordvin, and Komi. We monitor the interplay of the variables 'embedding depth', 'position of the clause relative to the higher clause', 'type of the embedded clause', 'polarity of the higher clause', 'tense of the DEC' and 'language (Estonian, Moksha, Komi)'.

We focus on two issues in this presentation. First, we discuss differences between the languages in relation to the embedding depth of their complex sentences.

Secondly, and more importantly, we argue that DECs make an independent contribution to the study of clausal subordination. We show that DECs are not entirely identical with first-order embeddings, i.e., that embedding depth is a factor influencing the grammar of subordinate clauses. For example, the use of left-embedding decreases and of center-embedding increases from first- to second-order embedded clauses. Relative clauses are more likely to be DECs than complement clauses, and successive embedding of adverbial clauses is much less likely than successive embedding of relative clauses. Higher embedding depth correlates with a more frequent use of finite clauses, lower depth with a more frequent use of non-finite clauses. We also show that some grammatical characteristics of a deeply embedded clause, e.g. its position relative to the superordinate clause and its tense, cannot always be predicted from its relation to the immediate superordinate clause, but can be governed by the ultimate main clause (cf. Letuchiy 2018; 2021: 525–570 on tense and Kehayov and Todesk 2024 on irrealis). We study the circumstances causing such long-distance effects.

References

CCLM = Arkhangelskiy, Timofey (2019), *Corpus of contemporary Literary Moksha*. <Moksha.web-corpora.net>

ENC (2019) = *Estonian National Corpus, version 2019*, accessible in SketchEngine: <https://www.sketchengine.eu>

van der Hulst, Harry. 2010. Re Recursion. In van der Hulst, Harry (ed.), *Recursion and human language* (Studies in Generative Grammar 140), xv–liii. Berlin & New York: De Gruyter Mouton.

Kehayov, Petar; Todesk, Triin 2024. Relative and absolute irrealis in deeply embedded clauses. In Björklöf, Sofia; Jantunen, Santra; Junttila, Santeri; Kuokkala, Juha; Laitinen, Saarni; Pasanen, Annika (eds.*), Itämeren kieliapajilta Volgan verkoille: Pühendusteos Riho Grünthalile 22. mail 2024* (Mémoires de la Société Finno-Ougrienne 278), 321–332. Helsinki: Société Finno-Ougrienne.

KKK = *Komi Kyv Korpus*. <https://komicorpora.ru/>

KORP = Borin, Lars, Markus Forsberg, and Johan Roxendal (2012), *Korp – the corpus infrastructure of Språkbanken* (Korp version 6.0.), Komi-Zyryan texts. <https://gtweb.uit.no/u_korp/#?lang=en&stats_reduce=word&cqp=%5B%5D>

Letuchiy, Alexandre 2018. On tense and irrealis marking in triclausal constructions (and what distinguishes them from biclausal constructions). *Linguistic*s 56–1 (2018), 163–206.

Letuchiy, Alexander B. 2021. *Russkij jazyk o situacijax. Konstrukcii s sentencial'nymi aktantami* [Russian about situations. Constructions with complement clauses]. Saint-Petersburg: Aletheia.