

The noncentrality of recursive self-embedding: Convergent evidence from grammar and performance

Andreas Trotzke and Markus Bader, Universität Konstanz

Since Chomsky referred to Humboldt when thinking about the ‘creative aspect of language use’ as a human-specific capacity and describing it by using recursive function theory, recursion has been a prominent issue in linguistics. Almost ten years ago, Hauser *et al.* (2002) argued that only syntactic recursion belongs to what they call FLN (Faculty of Language – Narrow Sense). Although ‘recursion’ is more general than self-embedding, the general ability that underlies the following notion of self-embedding language plays a central role in recent discussions: “A grammar is *self-embedding* [...] if it contains an A such that for some $\varphi, \psi (\varphi \neq I \neq \psi)$, $A \Rightarrow \varphi A \psi$ ” Chomsky (1959: 148). Specifically, in the recent psychological literature (e.g. Gentner *et al.* 2006, Friederici *et al.* 2006), only a grammar that can account for self-embedding languages is considered properly recursive (cf. Sauerland & Author₁ to appear). Our paper questions this centrality of recursive self-embedding both on the level of grammar and on the level of performance by focusing on recursive center-embedded relative clauses (RCs) in German. Crucially, in contrast to the divergence between grammar and performance models that has often been pointed out in the context of recursive embedding (cf. Stabler to appear), our paper points toward a unifying perspective.

On the level of grammar, we will adopt a ‘strong’ derivational view implying multiple points of Spell-Out (cf. Uriagereka 1999) and argue that center-embedded structures can be analyzed as complex forms of left-branching and, consequently, fall in one natural class with other forms of ‘left-tail recursion’ and thus do not have a special status of their own. It will be shown that our derivational analysis of center-embedded RCs is in line with recent approaches claiming that the narrow structure-building operations of grammar are not recursive at all and that recursion (in the sense of computational procedures allowing for self-embedding) crucially involves the interfaces (cf. Arsenijević & Hinzen 2010, Surányi 2010, Zwart to appear).

On the performance level, we conducted both (i) a corpus study and (ii) an experiment testing acceptability using a speeded grammaticality judgment tasks. Both studies were conducted in order to determine whether restrictions on multiple center-embedding of German RCs follow from known limitations of the human performance mechanisms that are independent from center-embedding as such, or whether it is necessary to postulate special constraints on center-embedding, as proposed in Karlsson (2007). Concerning (i), based on the *deWaC* corpus (Baroni *et al.* 2009), we show that in intraposed position, sentences with degree of center-embedding = 1 occurred about five times as often as sentences with degree of center-embedding = 2 ((2) vs. (1)). Crucially, a similar ratio turned up in extraposed position (sentences with degree of center-embedding = 0, cf. (4), versus sentences with degree of center-embedding = 1, cf. (3)). Concerning (ii), we tested the acceptability of the major structures investigated in the present corpus analysis by applying the method of speeded grammaticality judgments to syntactic issues (cf. Ferreira & Henderson, 1991). Both the corpus distribution and the performance of subjects in the judgment task dovetail well with the analysis proposed on the level of grammar: center-embedded structures fall in one natural class with other complex forms of left branch embedding and should therefore not be regarded a special – let alone a central – property of syntactic structure-building. In particular, only the lengthy dependency between the antecedent NP and the clause-final verb makes complex intraposed relative clauses difficult to process, and not recursive self-embedding as such.

(1) **RC-high center-embedded, RC-low center-embedded**

Probleme gab es, weil einige Kurse,
problems existed it because some courses
die in der schönen Broschüre, die man vorher zugeschickt bekommt, aufgelistet waren,
that in the nice brochure that one before sent got listed were
gestrichen worden waren.
canceled been were
'There were problems because some courses which were listed in the nice brochure
which was sent out in advance were canceled.'

(2) **RC-high center-embedded, RC-low extraposed**

Probleme gab es, weil einige Kurse,
die in der schönen Broschüre aufgelistet waren, die man vorher zugeschickt bekommt,
gestrichen worden waren.

(3) **RC-high extraposed, RC-low center-embedded**

Probleme gab es, weil einige Kurse gestrichen worden waren.
die in der schönen Broschüre, die man vorher zugeschickt bekommt, aufgelistet waren.

(4) **RC-high extraposed, RC-low extraposed**

Probleme gab es, weil einige Kurse gestrichen worden waren.
die in der schönen Broschüre aufgelistet waren, die man vorher zugeschickt bekommt.

References

- Arsenijević, Boban & Wolfram Hinzen. 2010. Recursion as a human universal and as a primitive. *Biolinguistics* 4, 165-173.
- Baroni, M., Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta. 2009. The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation Journal* 23, 209–226.
- Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2, 137-167.
- Ferreira, Fernanda & John M. Henderson. 1991. Recovery from misanalysis of garden-path sentences. *Journal of Memory and Language* 30, 725-745
- Friederici, Angela D., Jörg Bahlmann, Stefan Heim, Ricarda I. Schubotz & Alfred Anwander. 2006. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *PNAS* 103, 2458-2463.
- Gentner, Timothy Q., Kimberly M. Fenn, Daniel Margoliash & Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204-1207.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569-1579.
- Karlsson, Fred 2007. Constraints on multiple center-embedding of clauses. *Journal of Linguistics*, 43, 365–392.
- Sauerland, Uli & Author₁. to appear. Biolinguistic perspectives on recursion: Introduction to the special issue.
- Stabler, Edward P. to appear. Recursion in grammar and performance. In Margaret Speas & Tom Roeper (eds.), *Proceedings of the Recursion Conference (Amherst 2009)*.
- Surányi, Balázs. 2010. Toward a strongly derivational syntax. In Michael T. Putnam (ed.), *Exploring Crash-Proof Grammars*, 167-212. Amsterdam: Benjamins.
- Uriagereka, Juan. 1999. Multiple Spell-Out. In Samuel D. Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 251-282. Cambridge, Mass.: MIT Press.
- Zwart, Jan-Wouter. to appear. Recursion in language: A layered-derivation approach. *Biolinguistics*.