

A treebank study of clausal coordinate ellipsis in spoken and written language



Karin Harbusch¹
karbusch@uni-koblenz.de

&

Gerard Kempen^{2,3}
gerard.kempen@mpi.nl



¹Computer Science Department, University of Koblenz-Landau
²Max Planck Institute for Psycholinguistics, Nijmegen; ³Cognitive Psychology Unit, Leiden University

Incremental sentence production and ellipsis

Incremental sentence production reduces the working memory capacity needed for advance planning: The planning units can be of considerably smaller size than in case of non-incremental production.

Ellipsis preempts the need to plan the detailed shape of one or more constituents and thereby reduces the size of planning units.

Expectation: Because working memory load tends to be higher in spoken than in written language, one expects that speakers, in comparison with writers, will more frequently resort to the use of elliptical constructions.

Clausal coordinate ellipsis (CCE) forms

Forward Conjunction Reduction (FCR)

- (a) Last year, Emma lived in Nijmegen and ... worked in Amsterdam
(b) The town where Paula works and ... Harry lives, is small

Gapping

- (a) You live in Delft and your son ... in Amsterdam
(b) Conrad commutes to Leiden, and ... usually by train

Subject Gap in clauses with Finite/Fronted verbs (SGF)

Into the woods went the hunter and ... shot a hare

Right Node Raising (RNR)

Simone submitted one ... and Agnes reviewed two abstracts

Corpus study on CCE in spoken and written English

(Meyer, 1995; Greenbaum & Nelson, 1999)

In written clausal coordinations, the proportion of CCE versions was about twice as high as in spoken coordinations.

The pattern was explained in terms of **audience design**: Non-elliptical (unreduced) clauses include more repetition, thereby facilitate comprehension.

Relative frequencies of the individual CCE forms in spoken and written Dutch: A treebank study

1. **ALPINO** (written newspaper text)

2. **CGN 2.0** (text spoken in various situations)

The general data pattern for English could be verified: **In written Dutch, the percentage of elliptical versions within the set of all clausal coordinations was even three times higher than in spoken Dutch: 34% vs. 11%.**

However audience design cannot explain a remarkable shift in the relative frequencies of the two most frequent CCE forms:

	Gapping	FCR	SGF	RNR
ALPINO	10%	82%	3%	5%
CGN (2.0)	31%	61%	5%	3%

A new interpretation

The lower proportion of elliptical versions within the set of all clausal coordinations as well the frequency shift of CCE forms result from **a narrower scope (“window”) of online grammatical planning** in spoken as compared to written sentence production. More specifically, in order not to overtax online working memory load, speakers have a stronger tendency than writers to **plan the grammatical shape of each clause in isolation**, i.e., without taking the shape of coordinated clauses into account, thus overlooking many elliptical options (FCR in particular). In Gapping, however, the second clause is planned not as the projection of a verb but rather as a modification or extension of an existing (the first, non-elliptical) clause, much like a substitution repair (Kempen, 2009). **Hence, because Gapping involves one overt verb and one clause only, it is less likely to overburden online working memory.**

Conclusions The data suggest that

- (1) CCE in spontaneous speech benefits the speaker, not the listener; and
- (2) Gapping should be analyzed as a monoclausal structure-with-revisions rather than as a partly deleted biclausal structure.

References

- Greenbaum, Sidney & Nelson, Gerald (1999). Elliptical clauses in spoken and written English. In: Peter Collins & David Lee (Eds.) (1999). *The clause in English*. Amsterdam: Benjamins.
- Kempen, Gerard (2009). Clausal coordination and coordinate ellipsis in a model of the speaker. *Linguistics*, 47, 653-696.
- Meyer, Charles F. (1995). Coordination Ellipsis in Spoken and Written American English. *Language Sciences*, 17, 241-169.