UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Self-Embedding Is Not a Linguistic Issue

Permalink

https://escholarship.org/uc/item/6r47c38k

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 3(0)

Author

Carroll, John M.

Publication Date

1981

Peer reviewed

Self-Embedding is Not
A Linguistic Issue*

John M Carroll
MIT Linguistics and Philosophy Dept.
and IBM Watson Research Center

Anyone familiar with language research over the last 25 years or so will be no stranger to examples like these:

The woman died.
The woman the man met died.
The woman the man the girl loved met died.

Such examples have an illustrious past, both in the formal study of language and in the psychology of language. The self-embedding (SE) property they display singularly places natural language syntax beyond the generative capacity of finite devices (Chomsky, 1959).** The interaction of SE with language comprehension entrains a remarkable phenomenon: one (or zero) levels of SE cause no noticable increment in comprehension difficulty, but two (or more) levels are typically associated with substantial impairment of comprehension (e g., Miller and Isard, 1964).

Explaining SE. This remarkable fact has maintained the study of SE as a preeminent topic in psycholinguistics. A variety of accounts have been produced The earliest accounts tended to emphasize the extent to which SE overtaxed memory resources for speakers and hearers (Miller and Chomsky, 1963: 470ff.; Yngve, 1960). Later accounts tended to focus on the kinds of operations that were invoked in the course of processing SE structures (Miller and Isard, 1964; Bever, 1970). Significantly, though, all of these accounts regarded SE as a general structural property -- whose expression in language was of special interest -- not as a property unique to language ex hypothesi. Miller and Chomsky (1963: 484) write "Self-embedding of such great theoretical significance ... that we should certainly look for occurrences of it in non-linguistic contexts." Bever (1970) cited as one of the strengths of his account of SE processing problems that analogous principles seemed to explain phenomena in visual perception.

Clearly though, there is another theoretical option. One could hypothesize that SE phenomena are special to language. On such a view, the natural explanatory mechanism for the interaction of SE with sentence comprehension would be language-specific -- not some general property of memory or perceptual process as suggested by Bever, Miller and Chomsky, etc. This view is (often implicitly) adopted by many natural language parsing theorists. A recent example is Fodor and Frazier (1980) who suggest that the interaction of SE with sentence comprehension be theoretically reconstructed as a parsing principle attaching an incoming word or phrase into its surface structure using the smallest possible number of new nonterminal nodes (Fodor and Frazier, 1980: 426-434). (See Carroll, 1981; Ford, Bresnan, and Kaplan, 1981, and Wanner, 1980; for further discussion of this model.)

The question of whether generalizations about SE phenomena are linguistic or still more general is a question of fact. But it is worth emphasizing that assumptions one way or the other lead immediately to empirical consequence. For example, the assumption that SE is strictly linguistic actually does some work in the analysis of Fodor and Frazier. The single strong distinction Fodor and Frazier (1980) are able to draw between their model and that of Wanner (1980) turns precisely on their language-specific analysis of SE phenomena.

While the resolution of this issue may ultimately be decided on theoretical grounds, I will focus chiefly on the presentation of data whose prima facie analysis entails the view that the correct level of generalization for addressing SE phenomena is something like "complex sequences" I will discuss examples from film, dance, music, and social interaction Then, I will turn back to the theoretical level and outline a proposal, following Bever's (1970) Double Function Hypothesis.

Film. An obvious candidate for SE analysis in film is the "flashback" scene: an entire scene is embedded into another scene. In the cinema of D.W. Griffith, for example, a scene (S) can consist of a long-shot (L) followed by a series of detail-shots (D) A detail-shot may itself include a scene. This has the effect of embedding a scene within a scene In Carroll (1980: 61-63) I have discussed the following grammar-fragment as the basis of a formal analysis of this sort of composition

$$S ==> L + D*$$
$$D ==> D' + S + D'$$

The right-hand side of the second rule introduces the "prime" as a notation to indicate that a return is structurally implied. This amounts to a special sort of indexing

Several examples of this structure occur in Griffith's Broken Blossoms. Richard Barthlemess, as the Chinaman, casts a misty-eyed look within a detail-shot, and immediately there is a cut away to a flashback of his youth. It would not have been possible for Griffith to cut away from Barthlemess, show the flashback, and then proceed with the story-line without first returning to Barthlemess. The return is obligatory; it is part of the structure of (Griffith's) cinema. As subsequent examples will show this is the crucial property of sequences that affords SE. If a sequential domain has a structure strong enough to afford obligatory returns, it can have SE. (This, of course, is precisely analogous to points made by Chomsky, 1959, with respect to language.)

Dance. A well-known example of dance within dance is in the Nutcracker Suite where several episodes of puppet-dance occur as part of the ballet itself. More common, but also more subtle, are examples in which a dance phrase (Lasher, 1981) is interrupted, whereupon a complete and distinct other phrase is danced, then finally the original phrase runs to completion.

One class of such examples occurs when a female dancer is lifted by a male dancer in classical ballet: the female assumes a rigid posture while the male carries her and continues moving; when she is returned to the floor, the female continues with her own dance. Another class of examples are cases in which a group of dancers assume rigid postures while a subset, usually a couple or a single dancer, perform a dance. At the completion of the embedded dance, the group resumes. These latter types of examples have analogs in music where they are easier to cite because of the available notation.

Music. Music has an extremely articulated structure and affords several classes of SE structures. One of these is the cadenza: a strong cadence is interrupted by a solo and then completed. For example, the cadence I-6/4 -- V7 -- I, closing a section of a violin concerto, can be interrupted by a violin solo: I-6/4 -- V7 -- Solo -- I

When the interruption amounts to a delay in the matrix structure, it is often called "parenthesis" (Meyer, 1973). Meyer discusses an example from Haydn's String Quartet in Eb Major, Opus 50 No. 3: a melodic patterning of thirds Eb-G, F-Ab, implying G-Bb, and reinforced by harmonic and rhythmic patterns, is interrupted by a four measure repeating a-b-a-b structure: ". . the real melody is characterized by goal-directed motion; but the parenthesis is static."(p. 241)

Social interaction. For some years now, Sacks (1973) and associates have been developing a theory of the structure of conversational interactions based on the "adjacency pair" unit. The paradigm example is question/answer:

A: Do you know what an adjacency pair is?
B: No.

The implication of the "second pair part" entailed by the occurrence of the "first pair part" is sufficient structure to support SE Indeed, Goffman (1979: 258-9) noticed such examples.

A1: Can I borrow your hose?
B2: Do you need it this very moment?
A2: No.
B1: Yes.

Instead of answering A's initial question (and completeing an adjacency pair structure), B initiates a second question-answer adjacency pair. When this second pair is complete, there is a return to the first structure. Goffman also noted an example of multiple SE (B is a trainman in a station)·

A1: Have you got the time?
B2 Standard or Daylight Saving?
A3: What time are you running on?
B3: Standard.
A2· Standard then.
B1: It's five o'clock

Jefferson (n.d.) has shown however that such cases of multiple SE are extremely rare, although noniterative SE is quite common.

Indeed, it is striking that, as in the case of language, multiple SE structures simply do not obtain One certainly has flashbacks, but never flashbacks within flashbacks When such structures are employed (as by Alain Resnais in Je T'aime, Je T'aime), they are employed precisely in order to confuse the viewer with respect to temporal sequence Analogously, one cannot imagine straightforward contexts in which a dance phrase within a dance phrase within a dance phrase could appear in a ballet One cannot embed a theme within a parenthesis within a theme One cannot embed a play within a play within a play Etc. In the final section of this paper I want to return to the theoretical analysis of SE phenomena and to suggest a general account of this intermodal restriction.

Double function. The double function principle can be put quite simply as: "The same stimulus cannot be perceived in two incompatible ways at the same time." The prima facie force of the principle in visual perception and language comprehension (Bever, 1970) and cinema perception (Carroll, 1980: 190-193), exclusive of SE phenomena, has been reviewed elsewhere. As applied to SE, the line one would want to run is that the categories "embedder" and "embeddee" are perceptually significant and that accordingly an object of perception cannot simultaneously be both embedder and embeddee. From this the usual facts regarding SE follow.

This remains an empirical hypothesis, of course, but the plausibility of the premises for the analysis is considerable. In contrast, appeals to memory (Miller and Chomsky, 1963; Yngve, 1960) are less compelling in that people typically can memorize multiple SE sequences -- they just can't understand them. Appeals to processing constraint are either too vague to assess (Miller and Isard, 1964) or tantamount to the proposal under discussion (Bever, 1970). Finally, the cross-modal evidence of SE and the iterative SE constraint strongly indicate that language-specific analyses are missing the true generalization.

References

Bever, T G. The cognitive basis for linguistic structures. in J.R. Hayes (Ed.) Cognition and language learning. New York: Wiley, 1970.

Carroll, J.M. Toward a structural psychology of cinema. The Hague: Mouton, 1980.

Carroll, J.M. On fallen horses racing past barns. In Papers from the parasession on language and behavior. Chicago: Chicago Linguistic Society, 1981.

Chomsky, N. On certain formal properties of grammars. Information and Control, 1959, 2, 137-167.

Chomsky, N. and Miller, G.A. Introduction to the formal analysis of natural languages. In Luce et al. (1963).

Fodor, J D. and Frazier, L. Is the human sentence parser an ATN? Cognition, 1980, 8, 417-459.

Ford, M., Bresnan, J., and Kaplan, R. A competence-based theory of syntactic closure. In J Bresnan (Ed.) The mental representation of grammatical relations. Cambridge: MIT Press, 1981.

Goffman, E. Replies and responses. Language in Society, 1979, 5, 257-313.

Jefferson, G. Unpublished research, n.d.

Lasher, M.D. The cognitive representation of an event involving human motion. Cognitive Psychology, 1981.

Luce, R.D., Bush, R.R., and Galanter, E. (Eds.) Handbook of mathematical psychology. New York: Wiley, 1963.

Meyer, L.B. Explaining music. Berkeley: University of California Press, 1973.

Miller, G.A. and Chomsky, N. Finitary models of language users. In Luce et al. (1963).

Miller, G.A. and Isard, S. Free recall of self-embedded English sentences. Information and Control, 1964, 7, 292-303.

Sacks, H. Lecture notes. Summer Institute of Linguistics, Ann Arbor, Michigan, 1973.

Wanner, E. The ATN and the sausage machine: Which one is balony? Cognition, 1980, 8, 209-225.

Yngve, V.H. A model and an hypothesis for language structure. Proceedings of the American Philosophical Society, 1960, 104, 444-466.