The learnability of center-embedded recursion : experimental studies with artificial and natural language

Lai, J.

Author: Lai, Jun

Title: The learnability of center-embedded recursion : experimental studies with artificial and natural language

Issue Date: 2013-02-26

Chapter 4

Why we do understand *the dog that the man walks barks* but struggle with *the dog walks the man that barks*:

A Semantic Memory Account for Hierarchical and Linear Linguistic Recursion (SMR)

This chapter is based on: Lai, J. & Poletiek, F. (submitted).

Abstract

Previous theoretical "locality" accounts (Gibson & Thomas, 1996; Kimball, 1973; Lewis, 1996) explain the difficulty of processing hierarchical center-embedded sentences by working memory limitations hindering accurate linking of the long distance dependencies that center embedded constructions generate. Alternately, sentences with right branching relative clauses with dependencies in nearby positions are easier to process. Although a few studies showed effects of semantic characteristics of related words in complex sentences (Blauberg & Braine, 1974; Christiansen & MacDonald, 2009; Powell & Peters, 1973; Stolz, 1967), it is unclear how positional relatedness interacts with semantic relatedness between words in linear and hierarchical constructions. We present a sentence comprehension study manipulating structure (hierarchical and linear) and the congruency between the semantic and positional pattern of word associations (match, mismatch and neutral) in the sentence. The data suggest a strong influence of semantic-syntactic pattern congruency, which occasionally even fully overshadowed difficulties caused by syntactical structure and positional distance. Moreover, this congruency effect was equally strong for linear and for hierarchical structures. We propose our semantic-memory model for processing recursive (SMR) structures to account for this effect, which can not be explained by the classical locality view. SMR also challenges the classical assumption that hierarchical structures are complex and linear not (Gibson, 1998).

Sentence complexity has been a notable focus of interest to psycholinguists. Recently, linguistic recursive complexity has been proposed to be the crucial factor distinguishing humans and nonhumans (Bloomfield, Gentner, & Margoliash, 2011; Corballis, 2007; Grainger, Dufau, Montant, Ziegler, & Fagot, 2012; Hauser, Chomsky, & Fitch, 2002; Lai & Poletiek, 2011; Rey, Perruchet, & Fagot, 2012). Recursion is a computational self referential mechanism, which allows for a finite number of rules to produce an infinite set of output (Chomsky, 1957). There are many types of recursive rules in language. However, one particularly complex type of recursion in natural language sentences has been much studied, namely, center-embedded (CE) structures, typically described formally as A^nB^n grammar (Fitch & Hauser, 2004). Assuming two word categories: A-words (e.g. nouns in natural language) and B-words (e.g. verbs in natural language), the CE A^nB^n grammar specifies a basic rule about which A words may be paired with which B words, and a recursive operation for inserting a grammatical AjBj pair within another AiBi pair to result in a new grammatical sentence. In this manner, CE sentences follow an AiAj…BjBi pattern. Since the embedding structure involves a "stack" of syntactically dependent elements possibly far away from each other in the sentence (e.g., Ai and Bi), CE structures are called "hierarchical" and non-linear and therefore require hierarchical cognitive processing (Christiansen & Chater, 1999). For example, in the natural sentence (1) with a CE structure, a higher order non-linear process of binding each A's to a specific B is required for correct comprehension.

Recursive rules can be linear, however, as well. In right branching (RB) structures of type $(AB)^n$, in which syntactically related AB pairs are close or even adjacent to each other. In the RB sentence (2), for example, the positional close distance between A and B elements is a direct cue for their syntactical relatedness, facilitating a simple linear parsing strategy.

(1) *John saw that the cat that the dog that the man walked chased ran away*. [CE]

A_1 A_2 A_3 B_3 B_2 B_1

(2) [1]*John saw that the man walked a dog that chased a cat that ran away.* [RB]

A_1 B_1 A_2 B_2 A_3 B_3

The difference between hierarchical structure and linear structure is crucial in linguistic theories on learnability of language (Poletiek & Lai, 2012), and according to recent theorising, parallels the distinction between the human and animal language faculty (Fitch & Hauser, 2004; Hauser et al., 2002). Hierarchical processing has been argued to imply cognitive control, higher order computation, consciousness and executive control, in contrast to linear processing, relying on low level memory and associative mechanisms. However, general cognitive limitations clearly affect and limit the processing of hierarchical structures, as evidenced in our difficulties to parse structures with multiple clauses in natural language, and also in experimental studies on learning complex artificial systems. Though most authors agree that there is some role for working memory mechanisms in processing hierarchical structures, it is still empirically unclear and under debate, *how* memory mechanisms and associative learning come into play, and whether they suffice to account for how human language users deal with hierarchical recursion (Friederici, Bahlmann, Friedrich, & Makuuchi, 2011; Makuuchi, Bahlmann, Anwander, & Friederici, 2009). The present study focuses on these questions. In particular, we argue that memory and associative learning mechanisms can largely explain recursive language processing, if we take into account the semantic aspects of the linguistic input and the way our memory deals with semantically rich content. By assuming memory content to be meaningful, we can push the working memory account to a clearer and more powerful explanation of complex linguistic behaviour.

Previous research has concentrated on the problem of limited memory as a quantity, i.e. memory load and computational integration effort have been argued to affect

---

[1] In the Dutch translation of the RB sentence (2) slight positional changes would occur, because the object can move in front of the verb in the relative clause (being "verb final"). However, the typical contrast between short distances for the linear RB constructions and the long distances in the CE constructions are conserved in Dutch. Also, the RB clauses are lined up in a linear sequence, over time, in both languages. See also Appendix. The Dutch translation of sentence (2), and the word-by-word translation back in English are:

(2) Jan zag dat [de man een hond uitliet] [die een kat achtervolgde] [die wegrende]
John saw that [the man a dog walked] [that a cat chased] [that ran away].
A_1 A_2 B_1 A_3 B_2 B_3

differentially CE and RB sentence comprehension, because items have to be kept in memory simultaneously and for a longer period of time in the former than in the latter structure. For example, CE sentences normally require retaining a certain subject noun in working memory until it can be associated with its further located predicate (verb), whilst in the meantime additional noun-verb pairings have to be determined. In sentence (1), “the cat” has to be encoded, stored and retrieved from memory, when “ran” appears at the end of the sentence. In the middle of the sentence, two other nouns have to be stored and retrieved, though not before the associated verb shows up.

It is not surprising then, that a large number of studies suggest that CE sentences are more difficult to understand than their RB counterparts (Bach, Brown, & Marslen-Wilson, 1986; Blauberg & Braine, 1974; Blumenthal & Boakes, 1967; Caplan & Hildebrandt, 1988; Christiansen & Chater, 1999; Fodor & Garrett, 1967; Gibson & Thomas, 1999; Hildebrandt, Caplan, & Evans, 1987; Larkin & Burns, 1977; Marks, 1968; Miller, 1962; Miller & Isard, 1964; Poletiek, 2011). Moreover, it is generally assumed that this difficulty increases fast with the number of levels of embedding (LoE) for hierarchical structures, since the dependencies are pushed away from each other further with each added clause. For RB sequences, LoE is thought to weakly affect difficulty or not at all (Chomsky, 1965; Church, 1982; Gibson, 1998; Marcus, 1980; Reich, 1969; Stabler, 1994). From 2-LoE on, i.e. two clauses hierarchically nested in the main clause, sentences are barely understandable (de Vries, Petersson, Geukes, Zwitserlood, & Christiansen, 2012; Foss & Cairns, 1970; Miller, 1962; Vosse & Kempen, 1991). This increasing complexity is reflected in its occurrence in actual natural languages: 2-LoE sentences are rare in written and even rarer in spoken language (Karlsson, 2010). The complexity level that an actual language user may have to deal with in natural CE sentences, therefore, ranges from 1- to 2-LoE.

Theories explaining the difficulty to process CE sentences and the relatively low accuracy in comprehending them have pointed at memory capacity constrains and computational limitations. The *structural configuration account* (Chomsky, 1965; Miller & Isard, 1964) suggests that the low acceptability of CE sentences is due to the manner in which these self-embeddings are configured. Since human parsers must apply the mirrorlike recursive operation to process each clause, they have to remember "re-entries" of each previous clause to reach the highest level (Holmes, 1973). This unique configuration of dependencies can increase in complexity beyond the human computational capacity (Johnson, 1998). Just and Carpenter's (1992) *working memory theory of comprehension* is based on a similar reasoning that complex structures require more integration and memory resources, explaining differences in processing difficulty for CE and RB constructions, but also individual differences.

A more recent account, *the processing overload account* (Gibson & Thomas, 1996; Kimball, 1973; Lewis, 1996) proposes that the integration of corresponding elements into one constituent costs more cognitive resources in CE than in RB structures, which allow for immediate integration of syntactically related elements thanks to the adjacent locations mirroring their syntactical relatedness. The *syntactic prediction locality theory* (SPLT) by Gibson (1998) provides further theoretical refining of this working memory account. The SPLT proposes that locality has a strong impact on both the integration cost and memory cost: For integrating, the computational resources needed to connect two related events increase along with the number of constituents to be related in the sentence and the distance between them. Regarding memory costs, it requires more capacity to maintain a local word in memory during a longer period of time before it can be associated with its counterpart. Summing up the common features of theories explaining differential processing difficulties for linear and hierarchical recursive constructions, it is assumed that computational and memory load increase for parsing CE sentences as compared to RB ones, because of the complex association pattern of the elements and the long distance between them in a CE sentence. For RB sentences, related elements being close or even adjacent to each other, memory and integration processes are hardly needed.

In line with these theoretical accounts, experimental studies have explored the effect of structure and level of complexity on cognitive processing, using both natural language materials (Bach et al., 1986; Blauberg & Braine, 1974; Kidd, Brandt, Lieven, & Tomasello, 2007) and artificial grammars (Conway, Ellefson, & Christiansen, 2003; de Vries, Monaghan, Knecht, & Zwitserlood, 2008; Fitch & Hauser, 2004; Lai & Poletiek, 2011; Poletiek, 2002; van den Bos & Poletiek, 2010). In Blauberg and Braine's (1974) early study, participants' comprehension of auditory presented CE and RB sentences with increasing LoE were compared. They found that RB sentences were more understandable than CE ones, and higher LoE hindered CE more than RB in comprehension. With 3-5 LoE, RB sentences were hard but still intelligible, while CE sentences became "virtually impossible" beyond 2-LoE. For 1- and 2-LoE sentences, however, accuracy of processing did not differ significantly between RB and CE sentences. Blauberg and Braine concluded that it was the unique hierarchically nested property of CE, which posed obstacles for comprehension.

Findings with the Artificial Grammar Learning (AGL) paradigm are consistent with natural language studies' findings. In the AGL procedure, typically, participants are trained with $A_1A_2A_3 \ldots B_3B_2B_1$ sentences (produced by a CE grammar) or $A_1B_1A_2B_2A_3B_3 \ldots$ (produced by a RB grammar) depending on the structure tested. After, participants give grammaticality judgments for new strings being either grammatical or ungrammatical. Accuracy of the grammaticality judgements indicates the amount of learning of the underlying grammar. Research using this paradigm suggests that CE structures are more difficult to learn than RB structures (Christiansen & Chater, 1999; Conway et al., 2003). De Vries, Monaghan, Knecht and Zwitserlood (2008) even found no learning at all of the hierarchical nested pattern of CE structures in an artificial grammar. However, recent studies have looked at extra linguistic factors that might help; for example, prosodic cues (Mueller, Bahlmann, & Friederici, 2010), frequency of occurrence of different types of CE structures (Reali & Christiansen, 2007), experience with complex grammatical constructions (MacDonald & Christiansen, 2002), animacy of the noun (Mak, Vonk, & Schriefers, 2002, 2006), and a starting small training regimen presenting the exemplars over time in increasing order of complexity, and overtraining with the simplest exemplars (without embeddings) (Lai & Poletiek, 2011; Poletiek & Chater, 2006). These studies revealed that factors external to the positional structure can help the integration and memory processes required to process these sentences.

A poorly attended but very straightforward factor that might support parsing messages with complex dependencies is simply the meaning of these dependencies. The semantic factor has hardly been considered in the discussion about the learnability of hierarchical recursion – considered to be a matter of syntax (Goldberg, 2003). The present work explores the effect of the semantic relations between syntactically dependent words in complex hierarchical sentences, and, for the first time, explores how semantic effects differentially influence non-adjacent hierarchical and linear dependencies. Especially, we model prior knowledge of language users about the semantic relations between words (e.g. A and B words) in terms of semantic distance, in analogy to positional distance. This *semantic* distance might help or hinder comprehending, depending on its congruency with the *syntactic* (i.e. positional) distance between syntactically related elements. In this manner, we explain how semantic features of the syntactically dependent elements affect cognitive processing.

The general influence of semantic effects on syntactical parsing has been shown in a number of studies (Fedor, Varga, & Szathmary, 2012). For example, in the sentence "Mary cut the bread with a knife", the syntactic pairing of "cut" and "knife" point in the direction of the correct syntactic analysis, and comprehension becomes easier (MacDonald, Pearlmutter, & Seidenberg, 1994). In an early study, Slobin (1966) showed a similar effect for parsing passive voice sentences: when the relations between two nouns were indeterminate, i.e., object and subject were reversible according to real world knowledge (as in *the girl is being held by the boy*), comprehension was more difficult than when they are irreversible (*the baby is being held by the mother*). In the same vein, Gennari and MacDonald (2008) compared processing of English sentences with objective relative- and subject relative-clauses. They denoted that semantic indeterminacy strongly caused comprehension difficulty.

A few experiments specifically looked at semantic influences in hierarchical CE constructions (Blauberg & Braine, 1974; Powell & Peters, 1973; Stolz, 1967), and more recently, computational work with Simple Recurrent Networks has been carried out (Fedor et al., 2012; Rohde & Plaut, 1999). In his early study, Stolz (1967) exposed his participants to 2-LoE CE English sentences ($A_1A_2A_3B_3B_2B_1$) and observed comprehension under different conditions of semantic relations between the A's and B's. When the semantics of the syntactically related A's and B's determined their relatedness (e.g., *dog barks*), human decoders "do very little syntactic processing" to find out the *syntactic* correspondences between individual A's and B's (Stolz, 1967). Syntactic analysis only occurs when it is highly necessary for understanding. Powell and Peters (1973) replicated Stolz's findings, and concluded that "semantically supported sentences were easier to comprehend and decode than were semantically neutral sentences" (e.g., *man walks*). Besides these early experimentational studies, a few computational and mathematical models (Poletiek & Lai, 2012; Rohde & Plaut, 1999; Weckerly & Elman, 1992) have looked at the effect of semantic biases. Rhode and Plaut (1999) found better performance for a computational model of CE pattern learning when semantic biases were present in the input. Moreover, Weckerly and Elman (1992) observed different performances for two sets of CE sentences: one set with semantic bias, i.e. verbs which were compatible with specific subjects/objects only, the other set without semantic bias. Training with a semantic biased input led to better performance.

On the basis of results reported thus far, semantically supporting content per se seems to help syntactical parsing of various syntactical constructions, including CE. However, we don't know whether semantics differentially tap into hierarchical structures as compared to "easy" linear structures. Such an interaction would be expected on the basis of a locality view on complex sentence processing. If semantic biases do not differ for RB and CE, then both long distance and short distance constructions might be controlled by semantic memory in same way; and semantic distances rather than the positional distances might determine how we deal with complex grammatical patterns.

Another open question is how *interfering* rather than supporting semantic relations affect CE and RB processing. Past research has only looked at two possible semantic biases: it compared *supporting* semantic cues (determinate) and *neutral* (indeterminate) ones. How do *negative* semantic cues affect recursive sentence processing? This third possibility, in which a syntactical analysis goes against a preferred semantic one, can crucially reveal which role is left over for syntactical analysis when a dominant semantic analysis is available, in linear RB structures versus complex CE hierarchical structures. Less interference is expected for positional easy constructions than for hierarchical constructions, by locality based views. If incongruent semantic content, however, interferes equally with hierarchical and linear constructions, then this speaks against a substantial role for linearity or hierarchy as crucial determinants of cognitive processing, in the presence of a semantic cue, even a cue that goes against the syntactic analysis. We propose a semantic-memory model of recursion (SMR) to deal with these questions. SMR makes specific predictions for processing difficulties across recursive structures, LoE, and semantic features of the clauses. In the same manner as memory for paired words is affected by the semantic relation between the to-be-memorized words, sentence processing of embedded sentences is hypothesized by SMR to be affected primarily by the *semantic* pattern of distances between the words that are to be integrated in the sentence, rather than the positional distances.

Regarding semantic "distances", ever since Craik and Tulving (1975), recall performance of word pairs has been shown to vary highly depending on their semantic "distance". Consider the following pairs of words to be memorized:

1) Dog bites / Girl cries / Bird flies

2) Dog walks / Girl runs / Bird stands

3) Dog cries / Bird bites / Girl flies

The first pattern of word pairs is plausible and determined. The second list is plausible but undetermined, since pairings could equally well be interchanged. The third list is highly implausible (going against the plausible pairings of 1) but determined (according to the alternative pairing pattern 1). We would have no difficulty to produce the second word of each pair in the first list when primed with the first word, but finding the correct pattern of matched words in list 2 and 3 poses much more difficulties. SMR assumes these strong semantic pairing effects on memory performance, rather than "locality" (distance in time and space between the items) to explain recursive sentence processing.

SMR is rooted in a usage based view on processing recursive complex language, assuming general memory and associative cognitive processes to underlie how we deal with linguistic stimuli (Christiansen & Chater, 1999; Perruchet & Rey, 2005; Tomasello, 2000). SMR, however, specifies in detail the process by which non-linguistic general memory mechanisms operate to achieve comprehension of non-linear messages.

Two crucial hypotheses of SMR are tested in our experiment: The first hypothesis is about the role of increasing complexity on processing recursive structures. Though both SMR and locality-based accounts predict that more embeddings lead to more processing difficulty, simply because the list of to be memorized pairs increases, locality theories also predict that this effect is stronger for CE (in which positional distance between dependencies increases along with the number of LoE, and thus multiple items have to be retained for a longer period of time and integrated according to an analysis of their positions) than for RB. In SMR, increased depth of embedding is predicted to affect all types of positional patterns equally, because the memory process resourced to retain and integrate the pairs relies on semantic rather than positional information. Importantly, this prediction of SMR holds for a realistic range of complexity in natural language use, i.e. no more than 2-LoE.

The second hypothesis is about the effect of semantic patterns on positional distances between syntactically related elements. Previous research suggests that when semantic associations are determinate and congruent with syntactic positional associated elements in hierarchical CE constructions, processing is facilitated (Stolz, 1967). For RB linear constructions, where the to-be-paired elements are nearby, no semantic facilitation is needed nor expected by locality-based models, and semantic interference will not substantially affect the analysis, because of the clear positional cue. In sum, if the semantic relations are clearly incongruent with the syntactical associations, the locality view predicts linear RB constructions to be less hindered than hierarchical CE constructions with long distances. SMR, however, predicts semantic facilitation or interference *independently* of positional structure. Our model predicts a strong effect of semantic bias, but no interaction with structure. It is only when semantic relations between syntactic pairs are neutral and fully indeterminate, that elements positioned close to each other might be easier to process than distant ones. We predict on the basis of SMR, this suppressing of syntactical analysis to occur for easy (linear) and "difficult" (hierarchical) constructions equally, when the semantic cues are at odds with the positional pattern.

Experiment

In the present experiment, we manipulate structure (CE and RB), LoE (1- or 2-), and three conditions of congruency between the semantic and the syntactic pattern of words pairing, in sentences with one or two relative clauses. In the "*match*" condition, pairing patterns are congruent; in the "*mismatch*" condition, they are incongruent; and in the "*neutral*" condition, they are indeterminate. The three possibilities are displayed schematically in Figure 1. The following sentences illustrate the CE and RB constructions with all types of semantic-syntactic congruency conditions.

(3) *The dog that the boy pats barks.* [Match-CE]

(4) *The dog that the cat watches runs.* [Neutral-CE]

(5) *The boy that the dog pats barks*. [Mismatch-CE]

(6) *The boy pats the dog that barks.* [Match-RB]

(7) *The cat watches the dog that runs.* [Neutral-RB]

(8) *The dog pats the boy that barks.* [Mismatch-RB]

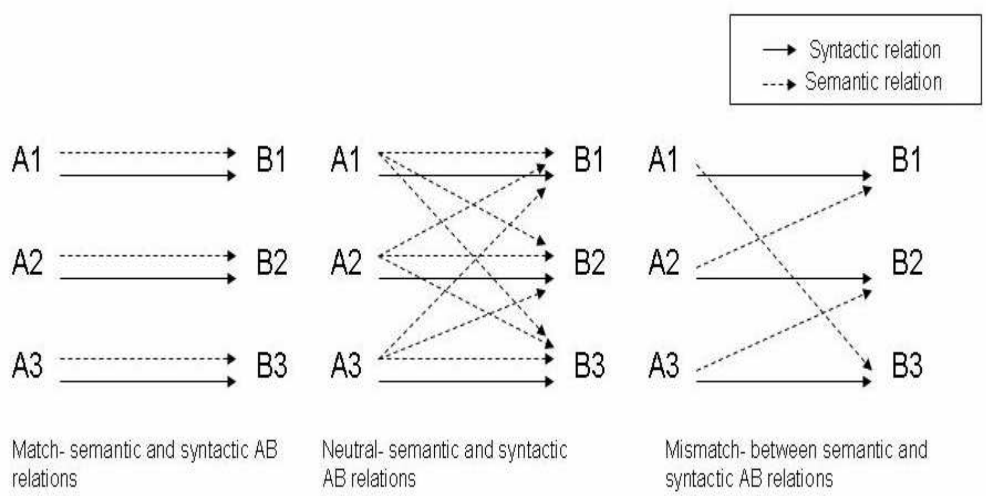


Figure 1. Schematic representation of the Syntactic-Semantic relations between A and B used in the stimulus sentences of type $A_1A_2A_3B_3B_2B_1$.

Method

Participants. Thirty-nine students (34 female), from Leiden University participated in the experiment for course credit or payment. All were native Dutch speakers. All had normal or corrected to normal vision.

Materials and design. There were 96 Dutch stimulus sentences with either one or two relative clauses (see Appendix). For each stimulus sentence, two short test sentences were constructed. Test sentences contained one subject and one predicate only. They served as test of participants' comprehension of the long stimulus sentence displayed previously. One of the short test sentences summarized an event actually described in the corresponding long stimulus sentence. The other one depicted a situation that was *not* a correct description of the content of the stimulus sentence. The incorrect test sentences contained nouns and verbs that were actually present in the stimulus sentence, but in other thematic roles than those in the stimulus sentence. For instance, in an incorrect test sentence, a subject noun could be associated with an unmatched predicate. For example, the stimulus sentence *the girl the dog bites cries* could have *the dog bites* (correct), and *the girl bites* (incorrect) as corresponding test sentences. The short summary sentences could refer to any subject in the long sentences[2]. Two counter-balanced test lists were created to ensure that each stimulus sentence had both a correct and an incorrect short test sentence. Participants were assigned randomly to one of the two counter-balanced lists, which again randomized the ordering of sentences across participants. Proportion of correct responses indicated comprehension accuracy.

The set of stimulus sentences had one of the three possible sentence structures: complex sentences with CE; complex sentences with RB; and simple sentences used as fillers. Since RB relative clauses used in the materials are verb final, subject verb pairs could be not fully adjacent, but separated by an object noun (see Appendix). For example: *Kees zag dat de man(A1) de hond(A2) uitliet(B1) die blafte(B2)* [in word-by-word translation: *Kees saw that the man the dog walked that barked.*] Overall, in our materials, RB constructions could have associated AB pairs separated by one or two words at most (short distance dependencies) and CE constructions could have AB pairs separated by eight words, for 2-LoE sentences. Furthermore, stimulus sentences had one out of three semantic types: *match*, i.e. the syntactical association pattern was congruent with the semantically most plausible association pattern; *mismatch*, i.e. the syntactical association pattern was incongruent with the semantically most plausible association pattern; and *neutral*, the

[2] However, for CE stimulus sentences with 2-LoE, A_2B_2 and A_3B_2 test sentences were excluded. This is because in the Dutch sentences used, the subject of B_2 is ambiguous. Grammatically, it can be either A_2 or A_3. See Appendix.

syntactical association pattern was unrelated to any semantic association pattern, because the semantic associations were indeterminate. In summary, the experimental stimulus sentences were manipulated orthogonally according to their structure (RB or CE), according to the match between the syntactical association pattern of A's and B's, and the semantically most plausible association pattern, and according to LoE (see Appendix for an example of each type of stimulus sentence).

Procedure. Participants were seated in front of a monitor, and were instructed that they would be exposed to pairs of Dutch sentences, visually. They would first see a long sentence, and immediately after, a short one. They had to judge whether the test sentence corresponded with the content of the stimulus sentence, or not, by pressing a YES key or a NO key. Participants were required to answer as quickly and as accurately as possible. Each trial started with a fixation cross (500ms) at the center of the screen. Each stimulus sentence began with "Kees weet dat …" (means "Kees knows that …" for 1000ms), and then appears word-by-word (800 ms per word, no interval in-between). It was followed by the short test sentence presented in the same manner. The task took approximately 35 minutes.

Results

In response to recent proposals regarding psycholinguistic data analysis accounting for both variance between participants and item simultaneously (Baayen, 2008; Brysbaert, 2007; Locker, Hoffman, & Bovaird, 2007), the analysis was carried out using a mixed-effects modelling. According to our first main hypothesis, number of LoE (one versus two) affects processing difficulty; and LoE affect processing hierarchical and linear sentences to the same extent. There was a main effect of LoE, F (1, 81) = 23.77, $p < .001$, but no main effect of sentence structure, F (1, 81) = 2.23, n.s., nor a significant interaction between LoE and structure, F (1, 81) = .07, n.s. As displayed in Figure 2, for CE sentences, performance on 1-LoE (M= .84, SE= .02) was significantly better than that on 2-LoE (M= .72, SE= .02), t (38) = 7.02, $p < .001$. Similarly, with RB sentence, performance on 1-LoE (M= .86, SE= .02) was significantly better than that on 2-LoE (M= .76, SE= .02), t (38) = 5.22, $p < .001$. At 1-LoE, performance over CE did not differ from that over RB significantly, *t* (38) = 1.72, n.s.; also, at 2-LoE, the difference between CE and RB did not reach significance, *t* (38) = 2.00, n.s.

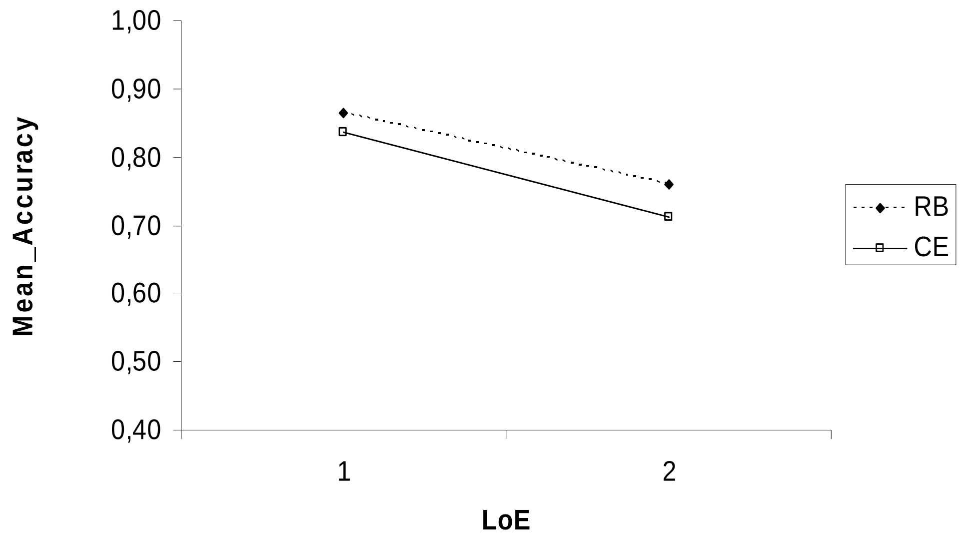


Figure 2. Mean accuracy for RB and CE sentences with 1- and 2-LoE.

Secondly, according to SMR, a strong main effect of semantic-syntactic congruency on accuracy is expected, and no interaction between congruency and structure is expected by SMR, though it is predicted by locality approaches.

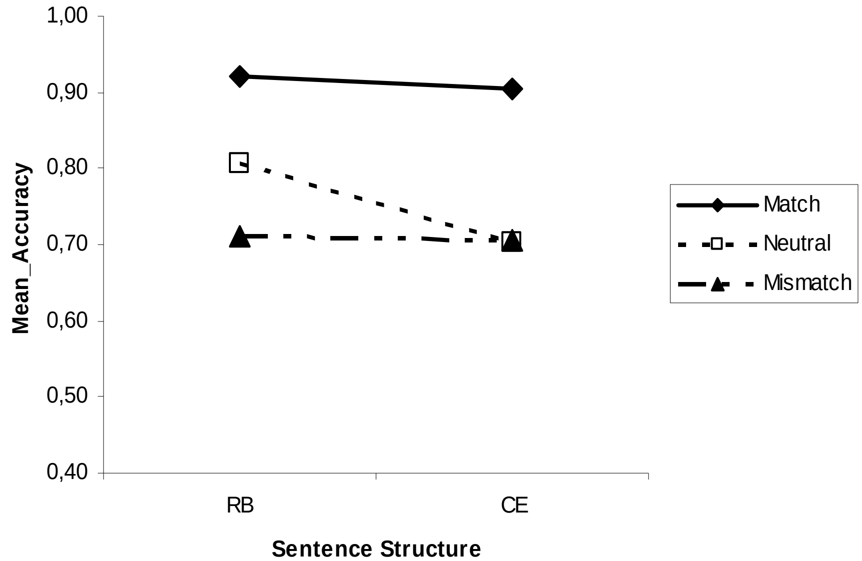


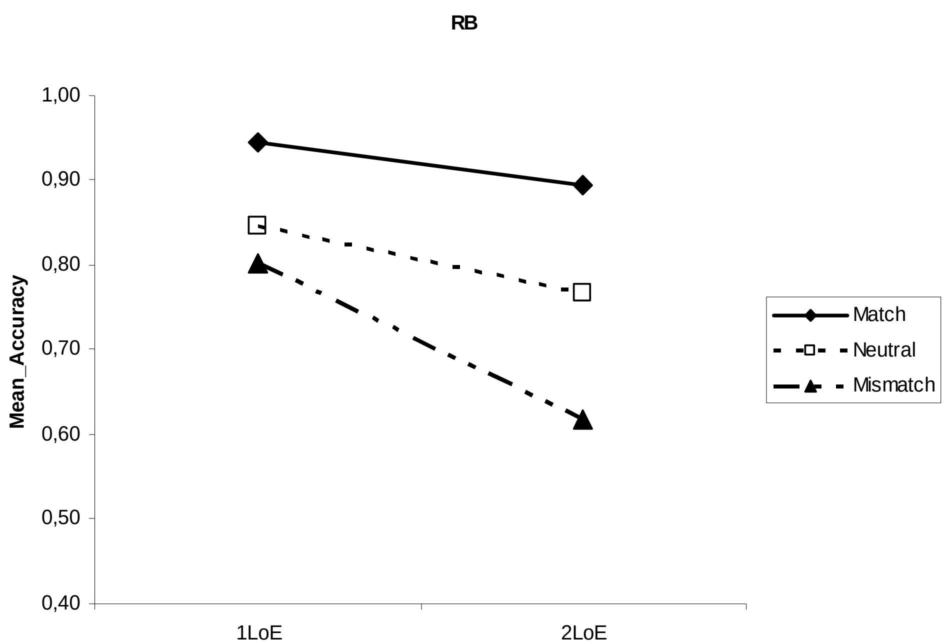
Figure 3. Mean accuracy for RB and CE over three semantic types.

The results indeed show a main effect of semantic type on accuracy, F (2, 81) = 31.88, p < .001, but no significant interaction between semantic type and structure, F (2, 81) = 1.83, n.s. There was no significant three-way interaction (Semantic type × Structure × LoE) either, F (2, 81) = 1.16, n.s. (Figure 3). Performance on semantic-matched (congruent) items (M=.92, SE= .01) was significantly better than on semantic-neutral ones (M=.76, SE= .02), t (38) = 10.56, p < .001, which was better than performance on semantic-mismatched ones (M=.72, SE= .02), t (38) = 2.50, p < .05.

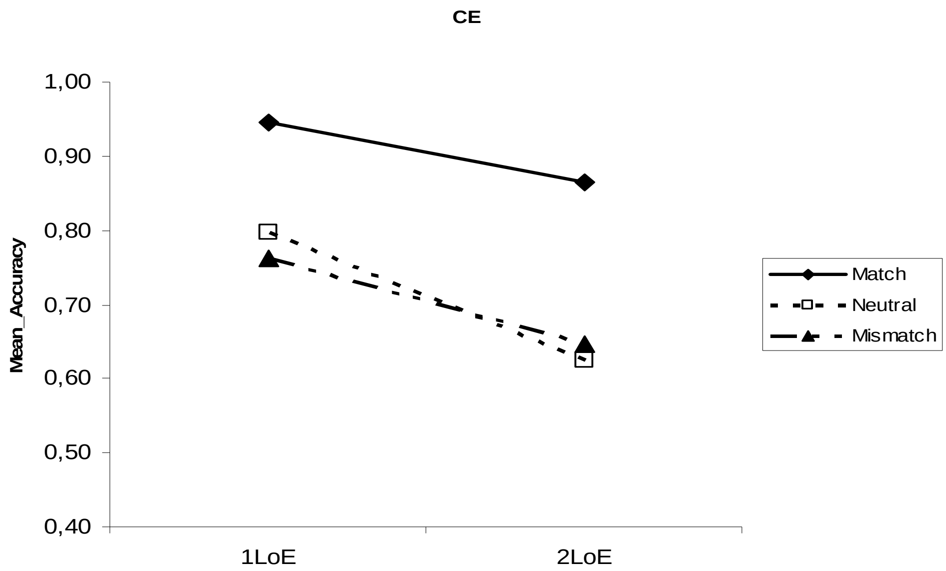
Though the interaction between semantic type and structure was not significant overall, Figure 3 shows differential performance on RB and CE for the neutral items, indicating that only these semantically neutral items were sensitive to positional organization of the pairs. This sensitivity was absent for items with either a matching or mismatching cue. For matched items, RB structures (M= .92, SE= .02) did not differ from CE structures (M= .92, SE= .01), t (38) = .04, n.s. Similarly, RB mismatched structures (M= .71, SE= .03) did not differ from CE mismatched structures (M= .72, SE= .02), t (38)

= .32, n.s. Only for the neutral items, did performance for RB structures (M= .81, SE= .02) surpass performance for CE structures (M= .71, SE= .02) significantly, *t* (38) = 3.32, *p* < .005 (see also Figure 4).

Figure 4 summarizes the effects of the manipulations taken together; only in the absence of semantic cues, RB constructions outperform CE constructions, and CE constructions are more strongly disrupted by an additional LoE than RB ones. For CE, the difference between mismatch items and neutral ones was not significant, *t* (38) = .33, n.s. Notice further two contrasts displayed in Figure 4 that are inconsistent with a locality view: for CE, 2-LoE matched sentences (M= .88, SE= .02) were scored even better than CE 1-LoE neutral (M= .79, SE= .03), *t* (38) = 2.82, *p* < .01, or CE 1-LoE mismatched ones (M= .76, SE= .03), *t* (38) = 3.86, *p* < .001. CE sentences with matching semantic-syntactic content, with both 1- and 2-LoE items, were better processed than RB sentences without semantic cue, *t* (38) = 6.42, *p* < .001.



(a)



(b)

Figure 4. (a) Mean accuracy for 1-, and 2-LoE RB over three semantic types.
(b) Mean accuracy for 1-, and 2-LoE CE over three semantic types.

Discussion

The present sentence comprehension study compares, for the first time, the effects of positional and semantic aspects of dependencies, in linear RB versus hierarchical CE structures, putting our SMR against the standard locality view. Our results show that for recursive sentences within the range of complexity that is actually present in natural language (1- or 2-LoE), sentence structure did not affect comprehension. Two levels of embedding sentences were more difficult to process than one level sentences, however. But this effect was independent of structure. Thus, the detrimental effect of additional relative clauses (that directly affects sentence length) was not larger for hierarchical structures where the dependencies are pushed apart to further positions, than for linear structures where the dependent elements remain in constant nearby positions. This data is hard to explain by a locality perspective that predicts more difficulties for multiple embeddings in hierarchically organized recursion than in linear recursion.

Furthermore, in line with SMR, there was a strong influence of the preferred semantic association pattern of dependent elements on comprehension. When a semantic cue was available to associate dependencies pair-wise, it strongly facilitated comprehension, if that cue was *congruent* with the *syntactical* association pattern. Inversely, it strongly hindered comprehension if the cue was *incongruent* with the positional association pattern. Strikingly, the semantic cue affected comprehension independently of the sentence structure. When a semantic association scheme for the words was available, it would strongly determine the sentence interpretation, whatever the positional scheme being linear or hierarchical. For example, both the linear sentence *the girl bites the dog that cries* and the hierarchical structured sentence with the same semantic content *the dog the girl bites cries* elicit an inaccurate but semantic plausible interpretation equally often. Another indication of the secondary role of positional information was that number of LoE failed to influence this semantic bias differentially for RB and CE structures. As accounted for by SMR, when there is a clear semantic pairing scheme for the elements, it strongly directs the integration of the sentence, whatever the positional distance of these elements. Positional factors also do not play a greater role with 2- than with 1-LoE.

In line with past findings, when there was no semantic cue to organize and retain in memory the pairing of elements, positional patterns mattered. The “pure” syntactical analysis then performed was remarkably poor, though, varying from 85% accuracy for the easiest linear sentences with 1-LoE to 62% for hierarchical sentences with 2-LoE. Though the latter contrast seemingly supports the locality view, the overall low performance for the neutral sentences when the structure is linear (and thus the positional conditions optimal) remains puzzling, for both locality theories and classical linguistic models assuming a predisposition for parsing the grammars of human languages (Chomsky, 1965; Church, 1982).

Our experimental results challenge the view that comprehension of recursive linear RB structures is generally better than that of recursive hierarchical CE structures (Foss & Cairns, 1970; Marks, 1968; Miller & Isard, 1964), because CE structures require on the one hand the elements to be retained during a longer period of time in memory, and on the other hand, a more sophisticated computational mechanism to determine the paired association of the elements than in RB sentences (Gibson, 1998). Instead, we found that semantic "distance" between the elements actually cause friction to or alternately subserved an accurate analysis. When the semantic association scheme happens to be in line with the syntactic scheme, recursive sentences are processed easily, *whatever* the syntactic scheme. The memory processes that are resourced to achieve comprehension of complex sentences also support our memory for meaningful materials corresponding to real world knowledge, autobiographic and contextual knowledge.

A similar semantic driven mechanism for sentence comprehension was proposed in the "good enough" parsing approach (Ferreira, Bailey, & Ferraro, 2002). Human parsers build up connections between words with the help of their real-world knowledge. As long as the available semantic pairs convey "good enough" meanings for understanding, parsers rapidly take advantage of that for comprehension. Here, we compared semantic influences for hierarchical and linear constructions. The inaccuracies found for even easy recursive patterns underline that good enough considerations strongly rely on semantic analyses.

Theories studying positional effects typically treat linear RB recursion as the simple "baseline" for comparisons with other more complex varieties of recursion (Christiansen & Chater, 1999). Accordingly, RB structures are argued to be processed without any difficulty (Church, 1982; Gibson, 1998; Marcus, 1980). The present results give a new perspective on what makes recursion difficult or not. For example, RB sentences were no longer simple to process when the semantic cue was inverse to the positional cue. When there was no semantic cue, RB linear sentences with only one clause were not always accurately interpreted. Our SMR suggests that not RB (as opposed to CE) is the easy default "baseline" form of recursion for the language user, but the situation in which the semantic association scheme and the positional scheme match. All conditions that deviate from this default situation, either because there is no semantic cue (and the parser has to resource pure abstract syntactical knowledge), or the semantic cue goes against the syntactical analysis (the parser is misguided by syntactical knowledge), cause difficulties.

One consequence of SMR is therefore that RB sentences are not more “basic” than CE sentences per se. For example, the sentence *the boy walks the dog that barks* would not be more frequent, basic or easy for language users than *the dog the boy walks barks*. But *the dog the boy walks barks* is predicted by SMR to be much more frequent and easier to process than *the boy the dog walks barks*. In the SMR view, it is this contrast reflected in differential frequencies and processing difficulty, between the “default” supporting semantic scheme versus the neutral or interfering semantic scheme in recursive sentences, which guides learning and everyday usage of these constructions. It is also this contrast that explains how general cognitive low level mechanisms, such as semantic memory and associative learning, provide powerful resources to guide learning. Indeed, a human learner might be exposed to default recursive sentences only in the early stage of learning (*the girl the dog bites cries*), and therefore get prepared to understand the deviations (e.g. sentences without semantic cue) from default in a later stage (*the girl the dog sees walks*).

To evaluate the SMR model further, various types of research are needed. For example, we need to know how much language users are actually exposed to semantically supported hierarchical structures and to the other types of semantic matching patterns. If, as we hypothesize within SMR, neutral and semantically mismatching sentences are largely outnumbered by semantically supporting ones, within the set of hierarchical sentences a language user comes across, this would speak for the SMR model. Notice that an analysis of the occurrence of the different types of semantic–syntactical congruency in sentences requires more than an analysis of isolated sentences of a corpus. Indeed, the semantic plausibility of a pattern of relations in a CE sentence depends on contextual factors, like discourse context, but also of the personal background knowledge of the listener. Referring to the example above, a sentence like *the girl the dog bites cries* might be easy to parse because of its description of an actual scene in the real world, but it might also be hard to parse, in the absence of such a scene, or if it is inconsistent with what happens around.

Positioning our study in the research on the learnability of hierarchical structures, our results support the low level mechanisms explanation of how humans deal with the long distances involved in hierarchical structures. In line with statistical learning models of language learning, SMR is “usage based” (Christiansen & Chater, 1999). In contrast to statistical approaches, however, the focus of our explanation for the handling of long distance dependencies is not on mechanisms that overcome positional distances (like transitional probabilities over more than one predicting element, or changes in variability of elements in given positions) (Gomez, 2002). It is the SMR concept of semantical “distance” between elements, which explains the present new data on how we deal with recursive complex linguistic constructions. In particular, why we do easily understand the hierarchical *the dog that the man walks barks,* but struggle with the linear *the dog walks the man that barks.*

Appendix

Examples of each type of stimulus sentence used in the task (CE versus RB; Matching, Mismatching and Neutral semantic-syntactic subject (A) - verb (B) relations; and 1- and 2-LoE). AiBi pairs with the same index have a syntactical subject-verb relation according to their position in the sentence. The English translations are word-by-word translations.

Sentence structure	Semantic-Syntactic relation type	LoE	Example "Kees weet dat …[1] *"Kees knows that…*
	Match	1	"…de dokter de patiënt die kermt onderzoekt." A1 A2 B2 B1 *"..the doctor the patient who groans examines.* *A1 A2 B2 B1*

[1] The purpose of an introductory phrase in our materials was to disambiguate the thematic role of the second noun A2 in CE sentences. With the introductory phrase, A2 is always subject of B2 only (sentence (a)). In sentence (b), without introductory phrase, A2 can be subject of both B1 and B2. Using the introductory phrase reduced the number of syntactically ambiguous SV relations in our materials, especially the CE sentences, and allowed us to improve our measurement of accurate sentence comprehension.

(a) "*Keest ziet dat* de vader het meisje dat schreeuwt, ziet."
A_1 A_2 B_2 B_1

(b) "Het meisje dat de vader ziet, schreeuwt."
A_1 A_2 B_2 B_1

CE		2	"…de politie de vrouw die de hond die poept uitlaat bekeurt."[2] A1 A2 A3 B3 B2 B1 *"…the policeman the woman who the dog that poops, walks* *A1 A2 A3 B3 B2* *arrests.* *B1*
	Mismatch	1	"…de hond de man die blaft bijt." A1 A2 B2 B1 *".. the dog the man who barks bites.* *A1 A2 B2 B1*
		2	"…de bouwvakker de vrouw die de auto die ronkt nafluit A1 A2 A3 B3 B2 bestuurt" B1

[2] In the CE sentences with 2-LoE and an introductory phrase "Kees knew that…." used here, the *third* noun A_3, however, *could* either be subject of A_3 only, or it could be subject of both A_3 and A_2. This ambiguity is solved only if the numbers of A_2 and A_3 differ. Since the number of the nouns was kept constant in all stimulus sentences (singular, to avoid such differences in number to serve as semantic cues also), this syntactical ambiguity was present in all our CE sentences with 2-LoE. Therefore, test sentences of type A_2B_2 and A_3B_2 could in principle not be rated incorrectly, since A_2 being subject of B_2 and A_3 being subject of B_2 are both correct analyses of the sentence. To avoid ambiguity in the participants' responses, no test sentences with A_2B_2 or A_3B_2 were used for CE sentences with 2-LoE.

		“…the worker who the woman who the car that throbs hails *A1 A2 A3 B3 B2* *drives” .* *B1*
Neutral	1	“…de jongen de vriend die valt helpt.” A1 A2 B2 B1 *“…the boy the friend who falls helps”.* *A1 A2 B2 B1*
	2	“…de vader het meisje dat de jongen die valt ziet volgt.” A1 A2 A3 B3 B2 B1 *“…the father the girl the boy who falls sees follows.* *A1 A2 A3 B3 B2 B1*
Match	1	“…de bakker het brood bakt dat rijst.”[3] A1 A2 B1 B2

[3] Another feature of Dutch related to embedded sentences with an introductory phrase like “Kees knew that…” used here is that the clauses are verb final. Therefore, the object in the relative clause precedes the verb (SOV), in contrast to English where the sequence within a relative clause is SVO. In RB sentences with multiple clauses, this results in SV pairs (AB) that are not adjacent, but separated by the object that is also subject of the next clause. In our materials this results in sequences with related AB pairs being separated by one other word. Our RB sentences with 1-LoE would have sequence $A_1A_2B_1B_2$, and sentences with 2-LoE $A_1A_2B_1B_2A_3B_3$. As a result, the AB (Subject Verb) pairs in the RB sentences were either separated by one element, or they were adjacent. To keep all sentences as similar as possible, we used the same introductory phrase for every stimulus sentence, at the cost of this disadvantage for RB sentences. Both the maximum distance between A and B’s in CE sentences and the mean distance were higher in CE than in RB sentences.

RB			“…*the baker the bread bakes that rises*”. *A1 A2 B1 B2*
		2	“…de groenteman de klant helpt die vraagt om de bananen die A1 A2 B1 B2 A3 rijp zijn. B3 *“…the greengrocers the customer helps who asks for banana’s* *A1 A2 B1 B2 A3* *that are ripe* *B3* .
	Mismatch	1	“…de baby de moeder troost die huilt.” A1 A2 B1 B2 “… *the baby the mother comforts who cries.* *A1 A2 B1 B2*

		2	“…de muziek de DJ aanzet die klinkt in de zaal A1 A2 B1 B2 A3 die groot lijkt.” B3 *“…the music the DJ turns on that echoes in the hall* *A1 A2 B1 B2 A3* *that looks big.* *B3*
	Neutral	1	“…het kind de oma omhelst die puzzelt.” A1 A2 B1 B2 *“…the child the grandmother hugs who puzzles”*. *A1 A2 B2 B1*
		2	“…de pastoor de man begroet die zwaait naar de bakker die A1 A2 B1 B2 A3

			fietst." B3 " …*the priest the man greets who waves at the baker* *A1 A2 B1 B2 A3* *who cycles.* *B3*

Footnote

[1] In the Dutch translation of the RB sentence (2) slight positional changes would occur, because the object can move in front of the verb in the relative clause (being "verb final"). However, the typical contrast between short distances for the linear RB constructions and the long distances in the CE constructions are conserved in Dutch. Also, the RB clauses are lined up in a linear sequence, over time, in both languages. See also Appendix. The Dutch translation of sentence (2) and the word-by-word translation back in English are:

(2) Jan zag dat [de man een hond uitliet] [die een kat achtervolgde] [die wegrende].

John saw that [the man a dog walked] [that a cat chased] [that ran away].

A_1 A_2 B_1 A_3 B_2 B_3

[2] However, for CE stimulus sentences with 2-LoE, A_2B_2 and A_3B_2 test sentences were excluded. This is because in the Dutch sentences used, the subject of B_2 is ambiguous. Grammatically, it can be either A_2 or A_3. See Appendix.

[3] The purpose of an introductory phrase in our materials was to disambiguate the thematic role of the second noun A_2 in CE sentences. With the introductory phrase, A_2 is always subject of B_2 only (sentence (a)). In sentence (b), without introductory phrase, A_2 can be subject of both B_1 and B_2. Using the introductory phrase reduced the number of syntactically ambiguous SV relations in our materials, especially the CE sentences, and allowed us to improve our measurement of accurate sentence comprehension.

(a) "*Keest ziet dat* de vader het meisje dat schreeuwt, ziet."

A_1 A_2 B_2 B_1

(b) "Het meisje dat de vader ziet, schreeuwt."

A_1 A_2 B_2 B_1

[4] In the CE sentences with 2-LoE and an introductory phrase "Kees knew that…." used here, the *third* noun A_3, however, *could* either be subject of A_3 only, or it could be subject of both A_3 and A_2. This ambiguity is solved only if the numbers of A_2 and A_3 differ.

Since the number of the nouns was kept constant in all stimulus sentences (singular, to avoid such differences in number to serve as semantic cues also), this syntactical ambiguity was present in all our CE sentences with 2-LoE. Therefore, test sentences of type A_2B_2 and A_3B_2 could in principle not be rated incorrectly, since A_2 being subject of B_2 and A_3 being subject of B_2 are both correct analyses of the sentence. To avoid ambiguity in the participants' responses, no test sentences with A_2B_2 or A_3B_2 were used for CE sentences with 2-LoE.

5 Another feature of Dutch related to embedded sentences with an introductory phrase like "Kees knew that…" used here is that the clauses are verb final. Therefore, the object in the relative clause precedes the verb (SOV), in contrast to English where the sequence within a relative clause is SVO. In RB sentences with multiple clauses, this results in SV pairs (AB) that are not adjacent, but separated by the object that is also subject of the next clause. In our materials this results in sequences with related AB pairs being separated by one other word. Our RB sentences with 1-LoE would have sequence $A_1A_2B_1B_2$, and sentences with 2-LoE $A_1A_2B_1B_2A_3B_3$. As a result, the AB (Subject Verb) pairs in the RB sentences were either separated by one element, or they were adjacent. To keep all sentences as similar as possible, we used the same introductory phrase for every stimulus sentence, at the cost of this disadvantage for RB sentences. Both the maximum distance between A and B's in CE sentences and the mean distance were higher in CE than in RB sentences.