

**The complexity of nested structures in English: Evidence for the syntactic prediction locality theory of linguistic complexity**

Edward Gibson<sup>a</sup> and James Thomas<sup>b</sup>

October 9, 1997

<sup>a</sup> Department of Brain and Cognitive Sciences  
Massachusetts Institute of Technology  
Cambridge, MA 02139

<sup>b</sup> School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213

Electronic mail: [gibson@psyche.mit.edu](mailto:gibson@psyche.mit.edu), [jthomas@cs.cmu.edu](mailto:jthomas@cs.cmu.edu)

# The complexity of nested structures in English: Evidence for the syntactic prediction locality theory of linguistic complexity

## Abstract

This paper provides results from questionnaire data in English that help to uncover some of the factors responsible for the linguistic complexity of unambiguous constructions. The constructions that were tested are variants of center-embedded and right-branching clauses, including embeddings of relative clauses, sentential complements and sentential subjects in both pre- and post-verbal positions. The most promising theory of the results is the Syntactic Prediction Locality Theory (Gibson, 1997). According to this theory, memory cost is associated with each syntactic head that is required to complete the current input string as a grammatical sentence. Furthermore, the memory cost increases the longer a required prediction has been known but has not yet been located.

# The complexity of nested structures in English: Evidence for the syntactic prediction locality theory of linguistic complexity

## 1. Introduction

It has long been known that nested structures are more difficult to understand than corresponding right- or left-branching structures (Chomsky, 1957, 1965; Yngve, 1960; Chomsky & Miller, 1963; Miller & Chomsky, 1963; Miller & Isard, 1964), where a category *A* is said to be nested (or center-embedded) within another category *B* if *A* is contained within *B* and there is lexical material on either side of *A* within *B*. For example, the nested structure in (1) is much more difficult to understand than the right-branching structure in (2), in spite of the fact that the two sentences include the same lexical items in the same thematic relationships among one another:<sup>1,2</sup>

(1) [*IP* The administrator [*CP* who [*IP* the intern [*CP* who [*IP* the nurse supervised ]]] had bothered ]]] lost the medical reports ].

(2) [*IP* The nurse supervised the intern [*CP* who [*IP* had bothered the administrator [*CP* who [*IP* lost the medical reports ]]]]]].

Note that the processing difficulty associated with (1) is not related to ambiguity confusions (e.g., a “garden-path” effect), because there is no local ambiguity within the sentence. In (1), the relative clause CP “who the intern ... had bothered” is nested within the matrix IP subject-verb dependency “the administrator ... lost the medical reports”. Furthermore, the relative clause “who the nurse supervised” is nested within the outer relative clause subject-verb dependency “the intern ... had bothered”. This type of structure is therefore doubly nested/center-embedded. On the other hand, the relative clauses are not nested within other categories in (2).

Although there are a number of theories of the factors contributing to processing complexity in unambiguous constructions, until now there has been little experimental data testing the theories.

---

<sup>1</sup>Sentences that cause extreme processing difficulty are prefixed with the symbol “#”.

<sup>2</sup>For simplicity, we will use the sentence-level phrase structure notation from Chomsky (1986) according to which a sentence is an Infl(ection) phrase (IP) which is dominated by a Comp(lementizer) phrase (CP). None of our results depends on these assumptions. In particular, more traditional S and S' categories are consistent with our results, as are phrase structure assumptions from other current syntactic frameworks (e.g., Head-driven Phrase Structure Grammar, Pollard & Sag, 1994) as well as current Minimalist syntactic assumptions, which divide the categories IP, CP and NP into further categories (Chomsky, 1993, 1995).

This paper provides an empirical investigation of some of the factors involved in the processing of unambiguous English sentence structures by means of questionnaire ratings on sentence structures like (1).

Section 2 summarizes some current theories of syntactic complexity in unambiguous constructions. In Section 3, Experiment 1 is reported, which tests a number of predictions of these theories. One theory — the Syntactic Prediction Locality Theory — appears to offer the best account of these results. An alternative account of the data from Experiment 1 is presented in Section 4. In Section 5, Experiment 2 is presented, which was designed to separate the two possible interpretations of the initial data. Once again, the Syntactic Prediction Locality Theory offers the best account of the data. A general discussion is provided in Section 6.

## 2. Current theories of processing complexity

One of the best-known theories of the structural complexity of unambiguous sentences is Kimball's (1973) principle of two sentences, a principle which associates memory cost with locally incomplete predicted clauses:

(3) The principle of two sentences:

The constituents of no more than two sentences can be parsed at one time (Kimball (1973), p. 33).

Kimball observed that unprocessable sentences like (1) generally contain a linear position in the string at which point at least three verbs are necessary to make the string a valid sentence. (1) is an example of such a sentence: At the point of processing “the nurse” the subjects of three clauses have been processed, all of which require a verb to follow. In contrast, there is no position in (2) such that more than one verb must obligatorily follow. Kimball assumes a left-corner parsing algorithm under which the presence of a sentence-initial NP triggers the prediction of an S (IP) node. If more than two IP nodes are predicted at some parse state before their heads are encountered (verbs, under Kimball's assumptions), then processing breakdown results.

Over the years, a number of extensions to the idea of associating memory load with predicted sentence nodes have been proposed. Gibson (1991; see also Hakuta, 1981; MacWhinney, 1987; Gibson, Hickok & Schütze, 1994) proposed that memory cost be associated with local thematic-role assignment violations, of which the subject-verb relationship is one.<sup>3</sup> One of the proposed

---

<sup>3</sup>See Lewis (1993, 1996) and Stabler (1994) for different extensions to Kimball's idea which share the same crucial

properties of this theory, the property of thematic reception, associates memory load with NPs which require thematic roles:<sup>4</sup>

(4) The Property of Thematic Reception:

Associate a cost of  $x_{TR}$  processing load units (PLUs) to each confirmed node 1) that is in a position that can be associated with a  $\theta$ -role in any of the structures currently under consideration; 2) that unambiguously heads a chain; and 3) whose role-assigner is not unambiguously identifiable (from Gibson, Hickok & Schütze, 1994).

At the point of processing the most embedded subject “the nurse” of a doubly-embedded RC structure like (1) there are five thematic reception violations: one corresponding to each of the initial NPs which require thematic roles, but do not yet receive roles at this point in the parse: “the administrator”, “who”, “the intern”, “who” and “the nurse”.<sup>5</sup> As a result of intuitive evidence like this, Gibson hypothesized that the capacity for these open relationships is around four thematic violations, so that five violations would cause breakdown.

A superficial advantage that the thematic-role-based complexity theory has over Kimball’s proposal is its ability to account for the acceptability of structures consisting of a relative clause embedded within either a sentential subject or a sentential complement of an NP (Cowper, 1976; Gibson, 1991):<sup>6</sup>

---

empirical difficulties as Gibson’s theory. See Gibson & Thomas (1996) for comparisons among Gibson’s, Lewis’s and Stabler’s theories.

<sup>4</sup>A second property, the property of lexical requirement, associates cost with unassigned roles. However, this property does not play a role in the examples to follow.

<sup>5</sup>It is possible to omit the relativizing pronoun in English object-extraction relative clauses, as in (5):

- (5) a. [<sub>IP</sub> The administrator [<sub>CP</sub> [<sub>IP</sub> the intern had bothered ]] lost the medical reports ].  
b. # [<sub>IP</sub> The administrator [<sub>CP</sub> [<sub>IP</sub> the intern [<sub>CP</sub> [<sub>IP</sub> the nurse supervised ]] had bothered ]] lost the medical reports ].

However, there is much syntactic evidence that a non-lexical pronoun is still present in a relative clause like that in (5b), even if it is not pronounced. Thus there are still five local thematic violations at the point of processing “the nurse” in (5b): three lexical NPs and two non-lexical *wh*-pronouns.

<sup>6</sup>The difference between a relative clause and a sentential complement is that one of the NPs inside a relative clause is missing, reflecting the presence of a relative pronoun position at the front of the clause. In contrast, there is no such empty argument position in a sentential complement. In (6) for example, the initial sequence of words “that the reporter discovered” is ambiguous between a relative clause modifier and a sentential complement of “information”:

- (6) a. [<sub>NP</sub> The information [<sub>CP</sub> that [<sub>IP</sub> the reporter discovered the tax documents ]] worried the senator.  
b. [<sub>NP</sub> The information [<sub>CP</sub> that [<sub>IP</sub> the reporter discovered ]] worried the senator.

The presence of the lexical NP as the object of “discovered” in (6a) disambiguates this clause as a sentential

(7) Sentential subject, then relative clause (SS/RC):

- a. That the employee who the manager hired stole office supplies worried the executive.
- b. That the administrator who the nurse supervised lost the medical reports had bothered the intern.

(8) Sentential complement, then relative clause (SC/RC):

- a. The fact that the employee who the manager hired stole office supplies worried the executive.
- b. The information that the administrator who the nurse supervised lost the medical reports had bothered the intern.

These structures are acceptable in spite of the fact that each contains a point at which three verbs are required in order to form a grammatical sentence. For example, after processing the most embedded subject “the manager” in (7a) and (8a), three verbs are required: one for the matrix subject (the sentential subject headed by the complementizer “that” in (7a); the NP “the fact” in (8a)); a second for the second subject “the employee” and a third for the third subject “the manager”. Thus the principle of two sentences predicts that these structures should be as unprocessable as the doubly-nested relative clause structure in (1). However, this prediction is incorrect: the structures in (7) and (8) appear to be much easier to process than the doubly-nested relative clause structure in (1).

The thematic-role-based theory offers a potential account of this contrast. In particular, the maximal complexity of processing the sentences in (7) and (8) is only four violations of the property of thematic reception. This memory cost occurs at the point of processing the most embedded subject in each. At this point, there are four phrases requiring thematic interpretation: 1) the matrix subject (the clause headed by “that” in (7a); the NP “the fact” in (8a)); 2) the first embedded subject NP (“the employee” in (7a) and (8a)); 3) the relative pronoun “who”; and the most embedded subject NP (“the manager” in (7a) and (8a)). Thus the maximal cost of processing the sentences in (7) and (8) is less than the five violations that are incurred in processing a doubly-nested relative clause, and the intuitive contrast is accounted for.

However, the advantage of the thematic-role-based theory over the sentence-based theory is not so clear when a wider range of phenomena is considered. In particular, the thematic-role-based theory predicts that the reverse embedding of relative clauses and sentential subjects/complements

---

complement. The lack of a lexical object of “discovered” in (6b) disambiguates this clause as a relative clause modifier.

— consisting of a sentential subject or sentential complement embedded within a relative clause  
— should be just as processable as the embeddings in (7) and (8). However, this prediction is not borne out: the reverse embeddings are much more complex:

(9) Relative clause, then sentential subject (RC/SS):

- a. # The executive who that the employee stole office supplies worried hired the manager.
- b. # The intern who that the administrator lost the medical reports had bothered supervised the nurse.

(10) Relative clause, then sentential complement (RC/SC):

- a. # The executive who the fact that the employee stole office supplies worried hired the manager.
- b. # The intern who the information that the administrator lost the medical reports had bothered supervised the nurse.

It is possible that the unacceptability of the embedded sentential subject structures in (9) is caused by a violation of a grammatical principle which disallows sentential subjects in any non-matrix position (Koster, 1978). Thus the acceptability difference between (7) and (9) might be due to a grammaticality difference between the two, such that the structures in (9) violate a grammatical principle, whereas the structures in (7) do not (Lewis, 1993; Stabler, 1994). (Of course, it is still possible that the difference between (7) and (9) has nothing to do with “grammaticality” *per se*, but is due to processing factors.)

However, the acceptability contrast between (10) and (8) cannot be accounted for by appeal to the same grammaticality difference, because the unacceptable structure, (10), does not contain an embedded sentential subject. Indeed, neither (10) nor (8) violates any grammatical principles, and so the acceptability difference between the two must be due to processing factors. But neither the thematic-role based theory nor the principle of two sentences predicts the contrast. According to the thematic theory, the maximal complexity of the RC/SC structures in (10) is four incomplete thematic relationships, the same maximal complexity as for the processable SC/RC structures in (8). This maximal complexity occurs at the point of processing the most embedded subject NP, at which point the four NPs requiring thematic roles in (10a) are “the executive”, “who”, “the fact” and “the employee”. The thematic-based theory therefore does not predict a contrast between the SC/RC and RC/SC structures.<sup>7</sup> On the other hand, the principle of two sentences can

---

<sup>7</sup>Gibson (1991) proposed an additional memory cost property, the property of thematic transmission, to account

account for the unacceptability of the RC/SC structure (because it contains a position at which three verbs are required to complete it as a grammatical sentence), but it does not account for the acceptability of the SC/RC structures in (8) or the sentential subject / relative clause structures in (7). Thus neither theory is satisfactory with respect to the contrast between the SC/RC and RC/SC structures.

## 2.1. The Syntactic Prediction Locality Theory

The only linguistic complexity theory that we are aware of which provides an account of the contrast between the SC/RC and RC/SC constructions is the Syntactic Prediction Locality Theory (SPLT) proposed by Gibson (1997). According to the SPLT, the memory cost associated with a (partial) syntactic structure is the sum of the costs associated with retaining each category that is required to complete the current input string as a grammatical sentence. This requires an hypothesis about the inventory of syntactic categories in a sentence. For initial purposes, a syntactic theory with a minimal number of functional categories, such as Head-driven Phrase Structure Grammar (Pollard & Sag, 1994) or Lexical Functional Grammar (Bresnan, 1982), is assumed.<sup>8</sup> Under these theories, the minimal number of syntactic head categories in a sentence is two: a head noun for the subject, and a head verb for the predicate. If words are encountered that necessitate other syntactic heads to form a grammatical sentence, then these categories are also predicted, and additional memory load is incurred. For example, after processing the words “the man who the”, there are four obligatory syntactic predictions: 1) a verb for the matrix clause; 2) a verb for the embedded clause; 3) a subject noun for the embedded clause; and 4) an empty category NP for the wh-pronoun “who”.

Furthermore, it is hypothesized that the memory cost associated with each syntactic prediction except for the highest predicate increases over distance, where distance is quantified in terms of the number of intervening new discourse referents that have been processed since the category was first predicted, where a discourse referent is an NP (an object referent) or a main verb of a VP (an event referent) (Kamp, 1981; Heim, 1982). The motivation for this hypothesis is that much

---

for the contrast between the two embeddings of RCs and SCs. Although the inclusion of this additional factor allows for an account of the contrast between RC/SC and SC/RC embeddings, it is based on a questionable syntactic assumption and it receives little independent support.

<sup>8</sup>The SPLT memory cost hypothesis is also compatible with grammars assuming a range of functional categories (e.g., Chomsky, 1995) under the assumption that memory cost indexes predicted *chains* rather than predicted categories.



computational effort is involved in building a structure for a new discourse referent (e.g., Haviland & Clark, 1974; Haliday & Hasan, 1976; Garrod & Sanford, 1977; see Garrod & Sanford, 1994 for a recent summary). Expending this effort takes resources away from maintaining (activating) syntactic and discourse representations constructed earlier. Keeping track of older predictions therefore consumes more resources because the old structures have lost activation. The assumption of increased cost for older predictions fits with what is known about short term memory recall, where it has been demonstrated that it is harder to retain items (e.g., unrelated words or digits) in short term memory as more interfering items are processed (see, e.g., Waugh & Norman, 1965; see Baddeley, 1990; Anderson, 1994; and Lewis, 1996, for recent summaries).

Motivation for the claim that there is no memory cost associated with predicting the structurally-highest predicate (usually a verb) is that all sentences in a language are headed by a predicate, so that the processor always expects a predicate. This prediction can therefore be built into the processor: The processor can make this prediction without any knowledge of the initial words in a sentence. Other syntactic heads cannot be predicted in this way, so the processor has to make these predictions on-line, as lexical items are processed.

The SPLT memory cost function is summarized in (11):

(11) Syntactic prediction memory cost:

- a. The prediction of the structurally-highest predicate,  $V_0$ , in the current memory buffer is associated with no memory cost.
- b. For each required syntactic head  $C_i$  other than  $V_0$ , associate a memory cost of  $M(n)$  memory units (MUs), where  $M(n)$  is a monotone increasing function and  $n$  is the number of new discourse referents that have been processed since  $C_i$  was initially predicted.

As a simplification, Gibson (1997) assumes that  $M(n)$  is a linear function with a zero intercept  $M(n) = K*n$ . The linear assumption is consistent with much of the evidence presented by Gibson, with one class of exceptions: examples in which memory cost distances are greatly lengthened. These kinds of examples suggest that the cost function is not linear in its limiting behavior, but rather heads asymptotically towards a maximal complexity, and so might be better modeled by a sigmoid function  $= \frac{1}{1+e^{-An+B}}$ . For simplicity, we will initially assume the simplified linear function.

It is proposed that the acceptability of a sentence which is locally and globally unambiguous (so that there is no processing difficulty due to reanalysis) is determined by the maximal quantity of memory resources that are required at any point during the parse of the sentence, such that

the more memory resources that are required, the less acceptable the sentence is. If more memory resources are required than the available capacity then processing breakdown eventually results. For example, according to the SPLT, the reason that multiply nested sentence structures like (1) (repeated here as (12)) are so difficult to process is that they require too many memory resources at some point during their processing.

(12) # The administrator who the intern who the nurse supervised had bothered lost the medical reports.

There are five predicted syntactic heads at the point of processing “who” in this sentence structure: 1) a matrix verb (eventually realized as “lost”); 2) the first embedded verb (eventually realized as “bothered”); 3) an empty-category to be coindexed with the first relative pronoun “who”; 4) the second embedded verb (eventually realized as “supervised”); and 5) an empty-category to be coindexed with the second relative pronoun “who”. According to the SPLT, maintaining the prediction of the structurally highest verb — the matrix verb in this instance — is cost-free, whereas the cost of maintaining the other predictions increases as new discourse referents are processed without realizing the predictions. Thus maintaining the predictions associated with the outer relative clause (numbers 2 and 3 above) requires  $M(1)$  MUs each, corresponding to the one new discourse referent “the intern” which has been processed since the predictions were first made. Maintaining the predictions associated with the inner relative clause (numbers 4 and 5 above) requires  $M(0)$  MUs each, because no new discourse referents have been processed since these predictions were made. Thus the total memory cost at the second occurrence of “who” is  $2M(1)+2M(0)$  MUs =  $2K$  MUs under the assumption that  $M(n) = K*n$ . After processing the most embedded subject “the nurse”, the memory cost increases for all four embedded predictions, to a total memory cost of  $2M(2)+2M(1) = 6K$  MUs. Upon processing the next word — the verb “supervised” — the two predictions associated with the most embedded relative clause are satisfied. At this point the memory cost for the other two cost-bearing predictions rises to  $2M(3) = 6K$  MUs. Both of the outer relative clause predictions are finally satisfied at the point of processing the verb “bothered”.

The point of maximal memory complexity for this structure occurs at the point of processing the most embedded subject NP “the nurse”, at which point the required memory resources are  $2M(2)+2M(1) = 6K$  MUs. It is hypothesized that this quantity of memory resources is more than the available capacity for most people, so that the sentence structure is unprocessable.<sup>9</sup>

---

<sup>9</sup>For a structure that requires more resources than the available capacity, processing breakdown does not occur

A second component of the SPLT is a theory of the computational resources involved in integrating a new category into the current structure for the input. Each integration has a syntactic component, responsible for attaching structures together, for example matching a syntactic category prediction or linking together elements in a dependency chain. Each integration also has a semantic and discourse component which assigns thematic roles and adds material to the discourse structure. It is assumed that each linguistic integration requires 1) a fixed quantity of computational resources, plus 2) additional computational resources related to the distance between the elements being integrated, to be viewed as resources necessary to reactivate the non-local element so it can be processed. The hypothesis of additional cost over longer distance is supported by evidence that local integrations are less costly than more distant integrations, leading to a locality or recency preference in instances of ambiguity (Kimball, 1973; Frazier & Fodor, 1978; Gibson, 1991; Stevenson, 1994; Gibson et al., 1996; among others). The integration cost theory also accounts for longer reading times at points when long-distance integrations are being made, independent of ambiguity. For example, the SPLT integration cost theory successfully accounts for the observation that reading times in subject-extracted relative clauses are faster than those in object-extracted relative clauses (Holmes & O'Regan, 1981; King & Just, 1991; Just & Carpenter, 1992; Gibson, 1997). Although integration cost is an important component of the SPLT, the experiments reported in this paper are acceptability judgement experiments, so that the memory cost component is the more relevant component of the SPLT.

---

as soon as more resources are required than are available. Rather, in such situation, it appears that the parser purges (or forgets) costly syntactic predictions from its memory buffer, and continues processing normally. Later, when the purged syntactic predictions are required for a grammatical sentence to result, parsing breaks down, because the incoming words cannot be grammatically attached to the current structure. Thus (i) is grammatical, but unacceptable, because it requires too many resources in its processing:

(i) # The patient [ who the nurse [ who the clinic had hired ] admitted ] met Jack.

The point of processing breakdown in (i) does not occur until the final verb “met”, as evidenced by the surprising acceptability of (ii), which is ungrammatical (Gibson & Thomas, 1997b; cf. Frazier, 1985; Gibson, 1991):

(ii) ✓ \* The patient who the nurse who the clinic had hired met Jack.

See Gibson & Thomas (1997b) for experimental evidence and discussion relevant to the claims about structural forgetting and the point of processing breakdown in complex structures.

### 2.1.1. Evidence for the discourse-based memory-cost hypothesis

Although doubly nested relative clauses like (12) are usually effectively unprocessable, if the subject of the most embedded clause is a first- or second-person pronoun, such structures become greatly more processable, to the extent that they are acceptable (Kac, 1981; Kluender, in press; Gibson & Warren, in preparation):

(13) a. Isn't it true that example sentences [ that people [ that you know ] produce ] are more likely to be accepted? (De Roeck et al., 1982)

b. A book [ that some Italian [ I've never heard of ] wrote ] will be published soon by MIT Press. (Frank, 1992)

(14) a. # The pictures which the photographer who John met yesterday took were damaged by the child.

b. The pictures which the photographer who I met yesterday took were damaged by the child.

Gibson & Warren (in preparation) conducted a questionnaire experiment in which participants rated the acceptability of doubly-nested relative clause items which varied only in the content of their most embedded subjects. Four conditions were tested in which the most embedded subject was: 1) an indexical pronoun (first- or second-person) as in (14b); 2) a short proper names (e.g., "John" in (14b)); 3) a full definite NP (e.g., "the professor"); and 4) a pronoun with no referent in the discourse. The results indicated that the indexical pronoun condition is easier to process than any of the other three conditions, with no other differences observed.

The discourse-based SPLT memory cost hypothesis accounts for this contrast based on the observation that referents for indexical pronouns are already present in the current discourse, so that processing them does not cause a memory cost increment for pending syntactic predictions. Thus the memory cost at the point of processing the most embedded subject "I" in (14b) is only  $2M(1)+2M(0) = 2K$  MUs, less than the required memory cost at the corresponding location in (14a). Furthermore, the maximal memory cost of processing (14b) is only  $2M(2) = 4K$  MUs, less than the maximal memory cost needed to process (14a), and the acceptability contrast is accounted for.

### 2.1.2. Evidence for the hypothesis that the prediction of the top-most verb is cost-free

Empirical support for the claim that the prediction of the top-most verb is cost-free comes from a number of experimental results, including a phoneme-monitoring experiment conducted by Hakes, Evans & Brannon (1976), who demonstrated that a relative clause is no more difficult to process when it modifies the subject of the matrix verb, as in (15a), than when it modifies the object of the matrix verb, as in (15b):

- (15) a. Matrix subject modifier: The reporter [ who the senator attacked ] ignored the president.  
b. Matrix object modifier: The president ignored the reporter [ who the senator attacked ] .

Although the task was sensitive enough to detect the difference between subject- and object-extracted relative clauses (confirming that object-extractions are more complex than subject-extractions) no complexity difference between matrix-subject-modifying and matrix-object-modifying relative was observed.<sup>10</sup> The lack of a complexity difference is not consistent with a memory function which associates memory cost with a predicted matrix verb. Such a function predicts that processing the relative clause should be harder when the relative clause modifies the subject rather than the object of a matrix verb, because the memory cost associated with predicting the matrix verb would be present when processing the subject modifier, but not when processing the object modifier. On the other hand, the zero-cost structurally-highest-verb hypothesis accounts for the lack of an observed complexity difference.

A number of the conditions in the experiments reported in this paper test this aspect of the SPLT, using acceptability judgement evidence.

### 2.1.3. Clause-based closure

An important part of the hypothesis in (11a) is that the prediction of the *top-most* predicate among the structures being retained in working memory is free of memory load. For single clause sentences, this hypothesis is equivalent to associating no cost with the matrix clause. However, the two hypotheses are distinct for multiple clause sentences. When the first predicted predicate is satisfied, it may be shunted out of working memory because the parser cannot keep an unbounded number of sites open for potential attachments, because of the bounded capacity of working memory. The

---

<sup>10</sup>Although not significant, the numerical effect actually favored the subject-modifier over the object-modifier.

parser must therefore implement some closure principle whereby some categories are shunted away, no longer to be considered as potential attachment sites. There currently exists a large amount of evidence for a clause-based closure principle like (16) (from Gibson, 1997; cf. Caplan, 1972; Jarvella, 1971, 1979; Marslen-Wilson, Tyler & Seidenberg, 1978; Frazier & Fodor, 1978; Bever & Townsend, 1979; Milsark, 1983; Gibson et al., 1996; and Roberts, Hooker & Gibson, 1997):

(16) Clause-based closure: The initiation of a new clause causes closure of an earlier clause whose obligatory requirements have been satisfied.

When such shunting occurs, a new prediction becomes the top-most predicted predicate, associated with no memory cost. For example, consider the right-branching sentence in (17):

(17) Bill said that Mary thought that Jason quit his job.

When the sentence is initiated, a predicate is immediately predicted. This prediction is satisfied by the matrix verb “said” in (2). After the word “that” is processed in (2), the obligatory immediate requirements of the verb “said” have been satisfied (an NP subject and a CP sentential complement), and so the matrix VP can be closed for further attachments by (16). Two categories are predicted at this point: an embedded predicate and its subject NP. Because the matrix VP has been closed, there is no memory cost associated with predicting the embedded predicate, because it is now the top-most predicted predicate. Thus the top-most predicted predicate is not necessarily the matrix predicate.

#### **2.1.4. The complexity of embeddings of sentential complements and relative clauses**

The SPLT offers a straightforward account of the difference between the SC/RC and RC/SC constructions discussed earlier. Consider (8a) and (10a) once again, repeated here as (18) and (19):

(18) SC/RC: The fact that the employee who the manager hired stole office supplies worried the executive.

(19) RC/SC: # The executive who the fact that the employee stole office supplies worried hired the manager.

According to the SPLT, the difference between the two constructions relies on the fact that a predicted sentential complement involves only one long incomplete category prediction — that of a verb — whereas a predicted relative clause involves two long incomplete category predictions — that of a verb and an empty category position for the wh-pronoun. For example, upon processing the

complementizer “that” in (18) two categories are required to complete the sentential complement: a noun and a verb. The noun requirement is satisfied very soon, when the noun “employee” is processed. The requirement for the verb is not met until much later downstream when the verb “stole” is encountered, after the embedded relative clause has been processed. Thus there is only one long unresolved category prediction when the sentential complement is the outer of the two embedded clauses. In contrast, when the relative clause is the outer of the two embedded clauses, as in (19), the predictions of both the verb and the empty category position for the wh-pronoun must be kept in memory during the processing of the embedded sentential complement. The two predictions are resolved far downstream, upon processing the verb “worried”.

In particular, the maximal memory complexity of the SC/RC construction occurs upon processing the word “company”, at which point the memory cost is  $M(2)+2M(1)$  MUs = 4K MUs corresponding to:

1. 0 MUs for the prediction of the matrix verb;
2.  $M(2)$  MUs for the prediction of the verb in sentential complement of “the fact that ...”;
3.  $M(1)$  MUs for the prediction of the verb in the relative clause, which was predicted after processing “the employee who ...”;
4.  $M(1)$  MUs for the prediction of a position coindexed with the wh-pronoun “who” in the relative clause.

In contrast, the maximal memory complexity of the RC/SC construction is  $2M(4)$  MUs = 8K MUs, which occurs upon processing the word “supplies”. At this point, the following unresolved predictions are present in the structure for the input:

1. 0 MUs for the prediction of the matrix verb;
2.  $M(4)$  MUs for the prediction of the verb in the relative clause, which was predicted after processing “the executive who ...”;
3.  $M(4)$  MUs for the prediction of a position coindexed with the wh-pronoun “who” in the relative clause.

Because  $2M(4) > M(3)+2M(1)$  for many monotonically increasing functions  $M(n)$ , including the linear function assumed here, the SPLT correctly predicts a memory complexity difference between the two constructions, as desired.

### 3. Experiment 1

Although the complexity contrasts discussed above seem to be substantial, until now they have not yet been evaluated experimentally. To partially alleviate the lack of experimental evidence, seven conditions were tested in Experiment 1, divided conceptually into three groups: 1) two embedding orders of relative clauses and sentential complements (SC/RC and RC/SC); 2) two embedding orders of relative clauses and sentential subjects (SS/RC and RC/SS); and 3) three conditions involving relative clauses within relative clauses. Examples of the SC/RC and RC/SC constructions tested here were given in (8b) and (10b) respectively, repeated here in (20):

- (20) a. The information that the administrator who the nurse supervised lost the medical reports had bothered the intern.
- b. The intern who the information that the administrator lost the medical reports had bothered supervised the nurse.

As discussed above, Kimball's clause-based complexity theory and Gibson's (1991) thematic-role-based theory predict no complexity difference between the SC/RC and RC/SC constructions, whereas the SPLT predicts that the SC/RC construction should be easier to process than the RC/SC construction.

Examples of the SS/RC and RC/SS constructions tested here were given in (7b) and (9b) respectively, repeated here in (21):

- (21) a. That the administrator who the nurse supervised lost the medical reports had bothered the intern.
- b. The intern who that the administrator lost the medical reports had bothered supervised the nurse.

The SPLT predicts that the RC/SS construction is more complex than the SS/RC construction, for the same reasons that it predicts that the RC/SC construction is more complex than the SC/RC construction. Although no other processing theory predicts this contrast, there is a plausible candidate for a grammatical explanation of an acceptability contrast, as discussed earlier. In particular, according to the grammatical theory of sentential subjects developed by Koster (1978),



sentential subjects appear in topic position, and so cannot appear in structures in which there is a fronted *wh*-word, such as (21b). Thus (21b) may be worse than (21a) because it violates the rules of the grammar, whereas (21a) does not (Lewis, 1993; Stabler, 1994).

The combination of the sentential complement and sentential subject conditions also allow us to test the hypothesis that the RC/SS condition is ungrammatical, in addition to being hard to process. Each of the SC/RC, RC/SC and SS/RC conditions is grammatical according to all current grammatical theories. If the RC/SS condition is ungrammatical, as Koster's theory claims, then there should be a larger difference in acceptability between the SS/RC and RC/SS conditions than between the SC/RC and RC/SC conditions. That is, the grammaticality hypothesis predicts an interaction between the factors 1) embedding ordering (RC-first, RC-second); and 2) sentential subject / sentential complement of an NP. On the other hand, if the RC/SS condition is just hard to process (but not ungrammatical), then no interaction is predicted.

The final three conditions in Experiment 1 consist of three variants of relative clauses embedded within relative clauses: 1) an object-extraction relative clause embedded within another object-extraction relative clause, in matrix subject position of a sentence; 2) the same as 1) but with a subject-extraction relative clause as the most embedded RC; and 3) the same as 1), but in a post-matrix-verb position:

(22) Relative clauses within relative clauses (RC/RC conditions):

- a. object-extraction, pre-verb: The administrator who the intern who the nurse supervised had bothered lost the medical reports.
- b. subject-extraction, pre-verb: The administrator who the intern who supervised the nurse had bothered lost the medical reports.
- c. object-extraction-RC, post-verb: The medical reports were lost by the administrator who the intern who the nurse supervised had bothered.

There are a number of comparisons involving the RC/RC conditions that are relevant to testing the syntactic complexity theories:

- Object-extraction vs. subject-extraction (both pre-verbal): The principle of two sentences predicts no difference here, because both conditions include parse states in which there are three predicted sentence nodes. The thematic theory predicts the subject-extraction RC to be less complex because there are at most four local violations in processing the subject-extraction condition, but five in processing the object-extraction condition.

The predictions of the SPLT depend on the memory-cost function being assumed. Under the linear function discussed above  $M(n) = K*n$ , no difference between the two conditions is expected, because the maximal complexity for both is  $6K$  MUs ( $= 2M(2)+2M(1)$  MUs  $= 2M(3)$  MUs). However, if the memory cost function has a non-zero intercept, so that  $M(0) = X$  MUs, then the object-extraction is predicted to be more complex than the subject extraction. Depending on the relative values of the slope  $K$  and the intercept  $X$ , the maximal memory cost of the subject-extraction RC/RC is either  $2M(3)$  MUs (at “nurse”) or  $2M(1)+2M(0)$  MUs (at “who”). Recall that one of the main ideas underlying the SPLT is locality, so that the primary memory cost associated with remembering predicted categories comes from keeping the categories in mind over the processing of intervening elements. In order to maintain the locality-based hypothesis, we will assume that the slope  $K$  is greater than the intercept  $X$ . For concreteness, we will initially assume that  $K=2X$ , so that  $M(n) = 2X*n + X$ .

Given this hypothesis, the maximal memory cost of the subject-extraction is  $2M(3) = 14X$  MUs. In contrast, the maximal complexity of the object-extraction is  $2M(2)+2M(1)$  MUs (at the point of processing the most embedded subject “the nurse”)  $= 16X$  MUs. Thus, the assumption of a non-zero intercept in the SPLT memory cost function leads to the prediction of a small complexity advantage in favor of the subject-extraction in doubly-nested RC constructions.

- Pre-verbal vs. Post-verbal (both object-extractions): Both the principle of two sentences and the thematic-role-based theory predict that the post-verbal RC/RC condition should be less complex. The principle of two sentences makes this prediction because there are only maximally two incomplete sentences in the processing of the post-verbal conditions, as compared to three in the pre-verbal condition. The thematic theory makes the same prediction based on a contrast between four and five local thematic violations.

According to the definition of the SPLT given in (11), there is no memory cost associated with the prediction of the top-level predicate in the current memory buffer. Under this hypothesis, there should be no complexity difference between pre- and post-verbal doubly-nested relative clauses, much as there is no processing difference between pre- and post-verbal singly-nested relative clauses (Hakes et al., 1976). On the other hand, if this component of the SPLT is incorrect, so that all predicted categories, including the matrix verb, are associated with an increasing memory cost, then the post-verbal condition should be much easier than the pre-verbal condition.

- Comparisons between the pre-verbal RC/RC conditions and the SC/RC, RC/SC conditions: The principle of two sentences predicts no contrasts here, since all four conditions include a parse state with three initial verb-less subjects. On the other hand, the thematic theory predicts a contrast between the object-extraction RC/RC condition (maximally five local violations) and the SC/RC and RC/SC conditions (maximally four local violations). The thematic theory predicts no contrast between the subject-extraction RC/RC condition on the one hand and the SC/RC and RC/SC conditions on the other, since each contains maximally four local violations.

The SPLT predicts that the SC/RC condition should be less complex than any of the RC/RC conditions. In particular, the maximal complexity of the SC/RC construction is  $M(2)+2M(1) = 4K$  MUs under the zero intercept assumption, or  $11X$  MUs under the non-zero intercept assumption. In contrast, the maximal complexity of the object-extracted RC/RC is  $2M(2)+2M(1) = 6K$  MUs under the zero intercept assumption, or  $= 16X$  MUs under the non-zero intercept assumption. Furthermore, the maximal complexity of the subject-extracted RC/RC is  $2M(3) = 6K$  MUs under the zero intercept assumption, or  $= 14X$  MUs under the non-zero intercept assumption.

The assumption of a linear cost function results in the prediction that the RC/SC condition should be more complex than any of the RC/RC conditions. If there is a zero intercept, then the maximal complexity of the RC/SC condition is  $2M(4) = 8K$  MUs, which is larger than the maximal complexity of the object-extracted RC/RC condition ( $= 2M(2)+2M(1) = 6K$  MUs). If there is a non-zero intercept, then the same complexity ranking is predicted, although the difference is not predicted to be as large (RC/SC maximal complexity  $= 2M(4) = 18X$  MUs; RC/RC maximal complexity  $= 2M(2)+2M(1) = 16X$  MUs). On the other hand, if  $M(n)$  is non-linear, approaching a maximal complexity, as suggested in Gibson (1997), then the maximal complexities of the RC/SC and RC/RC conditions may be similar, even if the memory cost function has a non-zero intercept.

One additional complexity hypothesis is worth mentioning here. It has been demonstrated that pragmatic distinguishability of thematic-roles and role-assignees influences the acceptability of nested structures, such that the more distinguishable the NPs and verbs are, the easier a nested structure is to comprehend (Stolz, 1967; Schlesinger, 1968). Thus one of the reasons that some RC/RC structures are hard to understand is that there is semantic interference among the NPs and/or among the verbs. It could be that the reason that the SC/RC

construction is easier to understand than the RC/RC construction is that the SC/RC construction involves more distinguishable subjects and verbs. In particular, one of the subjects is a conceptual NP (e.g., “information” or “fact”), which can only be the subject of a small subset of verbs (e.g., psychological verbs like “bothered” and verbs of causing like “caused”). This semantic restriction may make the SC/RC construction easier to understand. This hypothesis makes the following additional predictions: 1) the RC/SC construction should be of similar complexity as the SC/RC construction, because it has the same semantic interference properties as the SC/RC construction; and 2) the RC/SC construction should be less complex than the RC/RC construction.

### 3.1. Method

#### 3.1.1. Subjects

Forty-two native English speakers from the Boston, Massachusetts academic community (primarily undergraduates from MIT) participated for \$5.00 each.

#### 3.1.2. Materials

Thirty-five items, each with the seven conditions discussed above, were constructed. The compared conditions contained the same words and as much overlap of thematic relations among the NPs as was possible, with differences between conditions being matched for plausibility. The meaning differences among the conditions to be compared are summarized below, with comparisons based on the generic labelings of verbs and nouns in the SC/RC and RC/RC conditions in (23). The same shifts in relationships were maintained across each of the 35 items.

(23) a. SC/RC: NP<sub>5</sub> that NP<sub>1</sub> who NP<sub>2</sub> Verb<sub>1</sub> Verb<sub>2</sub> NP<sub>3</sub> Verb<sub>3</sub> NP<sub>4</sub>

b. RC/object-extraction-RC: NP<sub>1</sub> who NP<sub>4</sub> who NP<sub>2</sub> Verb<sub>1</sub> Verb<sub>3</sub> Verb<sub>2</sub> NP<sub>3</sub>

The meaning differences across items are summarized as follows. The items were constructed so that each of these meaning differences was similar in plausibility as judged by the experimenters.

1. SC/RC vs. RC/SC and SS/RC vs. RC/SS: No differences on Verb<sub>2</sub>, Verb<sub>3</sub>. The only difference in meaning between each of these two pairs of conditions involves the roles assigned by the most embedded verb Verb<sub>1</sub> in the SC/RC and SS/RC conditions (“supervised”):

SC/RC, SS/RC: The nurse supervised the administrator;

RC/SC, RC/SS: The intern supervised the nurse.

2. RC/RC-pre-verbal vs. RC/RC-post-verbal: There are no differences in the meanings here. However, there is a potentially confounding influence in this comparison: The items in the pre-verbal condition are all in the active voice, while the items in the post-verbal condition are all in the passive voice. Possible effects of this confound are discussed in Section 3.3.

3. RC/object-extraction-RC vs RC/subject-extraction-RC: No differences on Verb<sub>2</sub>, Verb<sub>3</sub>.

Differences on Verb<sub>1</sub>:

RC/object-extraction-RC: The nurse supervised the intern;

RC/subject-extraction-RC: The intern supervised the nurse.

4. SC/RC vs. RC/RC: No differences on Verb<sub>2</sub>. Differences on Verb<sub>1</sub>:

RC/RC: The nurse supervised the intern; SC/RC: The nurse supervised the administrator.

Differences on Verb<sub>3</sub>:

RC/RC: The intern bothered the administrator; SC/RC: The information bothered the intern.

It has been suggested by Bever (1970) that part of the problem in understanding doubly center-embedded relative clause structures has to do with the repetition of verbs which have the same tense and agreement features, as in (24):

(24) The cat that the dog that the horse kicked chased ran away.

In order to avoid this potential additional complexity influence, the items compared here were constructed so that the three verb phrases were morphologically non-uniform. In particular, either one or two of the three verb phrases contained an auxiliary verb, while the remaining verb phrases were in the simple past tense form. Most of the items were constructed so that the middle of three verb phrases in conditions with a three VP sequence was morphologically distinct from the other two.

In addition, Frazier & Fodor (1978) suggested that lengthening out the second VP in a doubly center-embedded RC structure makes such structures easier to process. As a result, the examples tested here included an adverbial phrase (e.g., a prepositional phrase, adverb etc.) that modified Verb<sub>2</sub>, which occurs as the head of the second verb phrase in the intuitively harder examples —

RC/RC, RC/SC and RC/SS — so that any difficulty in these examples is not due to short similar sentence-final VPs.

The experimental items were combined with 91 fillers (including 6 practice items) of approximately the same length and complexity as the experimental items, to form seven lists. The experimental items were counterbalanced across the lists so that each list contained exactly one version of every item. A list of materials is given in Appendix A.

### 3.1.3. Procedure

The stimuli were presented in the form of a questionnaire in which the subjects were asked to rate sentences on a scale from 1 (best) to 5 (worst) according to how easy or hard to understand the sentences were on the first reading. Six items were given as practice items on the first two pages of the questionnaire, with discussion of ratings for each of these. The remaining 120 items were presented to the subjects on a total of 12 further pages, 10 items per page. The order of the items in a list were randomized for each of the seven lists, and the pages of each survey were randomized separately. Subjects completed the task within about 30 minutes.

### 3.2. Results

The results of Experiment 1 are presented in Table 1 and Figure 1.

Table 1: Complexity ratings of embeddings of relative clauses, sentential complements and sentential subjects on a five point scale (standard errors in parentheses).

Condition	SC/RC	RC/SC	SS/RC	RC/SS
<b>Unacceptability rating</b>	1.92 (.08)	3.38 (.08)	2.24 (.09)	3.98 (.07)

Condition	RC/RC	RC/RC	RC/RC
	pre-verbal	post-verbal	pre-verbal
	obj-extract	obj-extract	subj-extract
<b>Unacceptability rating</b>	3.40 (.09)	3.24 (.08)	3.17 (.09)

The difference between the acceptability ratings of the two embedding orderings of SCs and RCs was highly significant ( $F1(1,41) = 45.57, p < .001$ ;  $F2(1,34) = 37.08, p < .001$ ). The

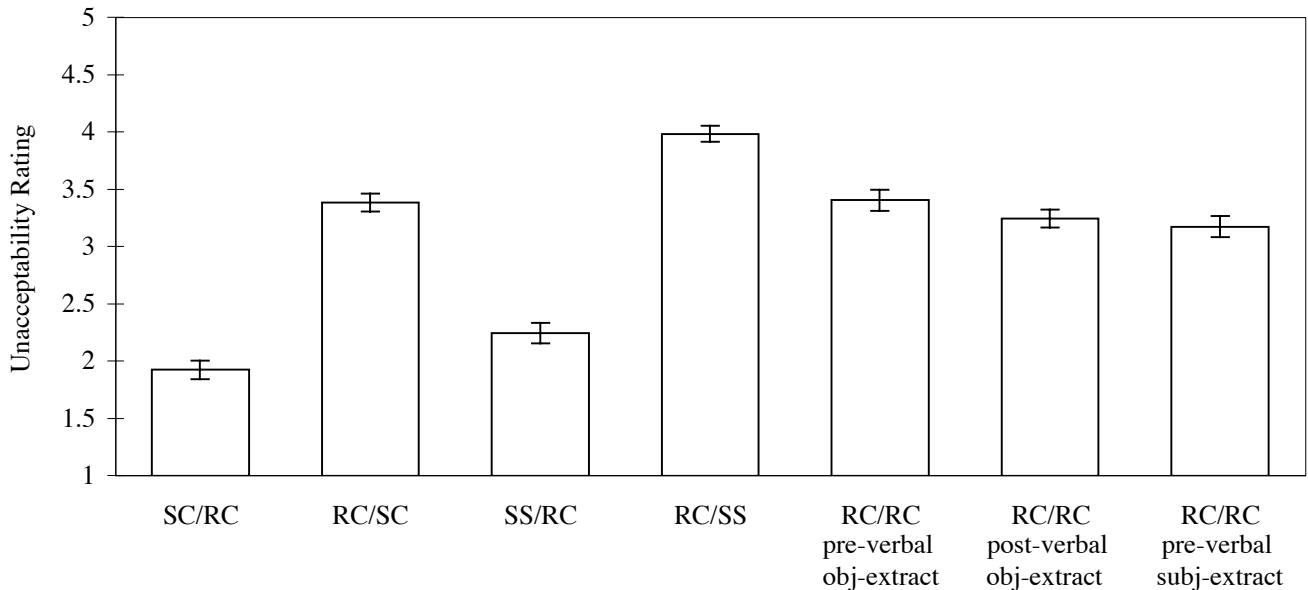


Figure 1: Complexity ratings for the conditions in Experiment 1.

difference between the relative orderings of SSs and RCs was highly significant ( $F1(1,41) = 64.71$ ,  $p < .001$ ;  $F2(1,34) = 53.45$ ,  $p < .001$ ). A significant effect of clause type (SS, SC) was also present ( $F1(1,41) = 13.93$ ,  $p < .005$ ;  $F2(1,34) = 26.19$ ,  $p < .001$ ), indicating that sentential subjects are more difficult to understand than sentential complements. There was no interaction between clause type (SS, SC) and the relative ordering of this clause with the relative clause (RC-first, RC-second) ( $F1(1,41) = 1.55$ ,  $p > .2$ ).

Within the RC/RC conditions, the effect of relative clause extraction-type (subject or object) was marginally significant ( $F1(1,41) = 3.54$ ,  $p < .10$ ;  $F2(1,34) = 3.61$ ,  $p < .10$ ), but there was no significant effect of position (pre- or post-verbal) ( $F1(1,41) = 1.73$ ,  $p > .10$ ;  $F2(1,34) = 1.99$ ,  $p > .10$ ). In addition, the ratings for the three RC/RC conditions were all highly significantly more complex than the SC/RC condition ( $ps < .001$ ). Finally, the RC/SC condition was marginally less complex than the pre-verbal subject-extraction RC/RC condition, but did not differ significantly from the pre-verbal object-extraction RC/RC condition ( $F_s < 1$ ).

### 3.3. Discussion

The results of each of the three sets of conditions are discussed in turn.

### 3.3.1. Sentential complements and relative clauses

The intuitive difference between the SC/RC and RC/SC conditions was confirmed experimentally. The result is not predicted by either the principle of two sentences or the thematic-role-based theory, or the hypothesis that pragmatic distinguishability of thematic-roles and role-assignees is what makes nested constructions easier or harder to understand. The observed complexity difference is accounted for by only the SPLT.

### 3.3.2. Sentential subjects and relative clauses

The difference between the relative orderings of the SS and RC embeddings was as predicted by the SPLT as well as Koster's theory of sentential subject phrase structure. However, the lack of an interaction between clause type (SS, SC) and embedding order (RC-first, RC-final) argues against the grammaticality account of the sentential subject results. Thus we tentatively conclude that a processing explanation of the high complexity of RC/SS structures is preferred to a grammaticality explanation. However, it is also possible that such an interaction failed to appear because of a ceiling effect in the RC/SS and RC/SC conditions, which were both rated as highly unacceptable.

There was also a main effect of clause type, such that the sentential subject versions were more complex than the corresponding versions with sentential complements of nouns. This lower acceptability of sentential subjects is not explained by the SPLT nor any other processing theory that we are aware of. It is possible that the lower acceptability might be due to lower frequency of use: sentential subjects are relatively rarely used, only in formal contexts. The lower frequency of use of sentential subjects might in turn be caused by the availability of extraposition as a mechanism to lower complexity, as in (25) (cf. Frazier & Rayner, 1988; Gibson, 1997):

- (25) a. That the administrator who the nurse supervised lost the medical reports had bothered the intern.
- b. It had bothered the intern that the administrator who the nurse supervised lost the medical reports.

### 3.3.3. Relative clauses within relative clauses

The results from the comparisons among the nested relative clause structures fail to support the two sentence theory, and provide at best weak support for the thematic theory. First, the effect



of relative clause extraction-type (subject or object) approached significance, thus supporting the thematic theory, but not the principle of two sentences. Second, although numerically in the predicted direction, the ratings for doubly embedded RCs did not vary significantly according to their position, thus disconfirming the predictions of both the thematic-based theory and the sentence-based theory.

In contrast, both of these results are easily accommodated within the SPLT. First, the fact that doubly-nested subject-extractions are (marginally) easier to process doubly-nested object-extractions follows from the SPLT if the memory cost function has a non-zero intercept, e.g., if  $M(n) = 2X \cdot n + X$  MUs. Given this assumption, the maximal complexity of the subject-extraction —  $2M(3) = 14X$  MUs — is less than the maximal complexity of the object-extraction —  $2M(2) + 2M(1) = 16X$  MUs — and the contrast between the two is accounted for. As a result of this observation, from this point on we will assume that the SPLT’s memory cost function has a non-zero intercept,  $X$ .

Second, the fact that there is no complexity contrast between pre-verbal and post-verbal RC/RC structures follows from the hypothesis that there is no memory cost associated with the prediction of the top-level predicate in the current memory buffer. However, there was a potential confound in the design of the materials which may have affected the results of the position comparison. In particular, all of the pre-verbal position examples were in the active voice, while all the post-verbal position examples were in the passive voice. The ratings of the post-verbal position RC/RC examples may have been worsened by being in the passive voice (cf. Forster & Olbrei, 1973, and the references there), thus making it harder to find a difference between the two conditions.<sup>11</sup> One of the points of Experiment 2 is to investigate processing ratings when voice information is controlled across the pre- and post-verbal conditions.

Comparisons of the RC/RC conditions with the SC/RC and RC/SC conditions provide additional evidence in favor of the SPLT over the other theories. First, both the two-sentence theory and the thematic-role-based theory predict that the SC/RC condition should be of similar complexity to the subject-extracted RC/RC condition. In contrast, the SPLT predicts that the RC/RC subject-extraction (maximal complexity =  $2M(3) = 14X$  MUs) should be more complex than the

---

<sup>11</sup>However, it is also possible that the ratings of the post-verbal position RC/RC examples were *improved* by being in the passive voice, because having one of the three verbs in passive makes the three verbs more distinguishable, which has been suggested by Bever (1970) to improve acceptability.

SC/RC construction (maximal complexity =  $M(2)+2M(1) = 11X$  MUs). The SPLT's prediction was ratified, with a highly significant difference between the two constructions being observed ( $ps < .001$ ).

The lack of a difference between the RC/SC and the RC/RC object-extraction conditions suggests that the SPLT's memory cost function is non-linear in the limit, so that  $2M(4)$  is no larger than  $2M(2)+2M(1)$ , in contrast to what the linear version of the function predicts. Although a non-linear function is more accurate, we will continue to use the simplified linear version of the function  $M(n) = 2X*n + X$  for simplicity of presentation.

Finally, note that the predictions of the semantic interference hypothesis were falsified. First, the semantic interference hypothesis predicts no difference between the SC/RC and RC/SC conditions, contrary to observation. And second, the semantic interference hypothesis predicts that the RC/SC should be easier than any of the RC/RC conditions, also contrary to observation.

### 3.3.4. Summary

In summary, the results as a whole are not generally consistent with either the two-sentence theory or the thematic-role-based theory. The following three results pose the most difficulty for these theories:

1. The large complexity difference between the two embedding orderings of sentential complements and relative clauses;
2. The large complexity difference between SC/RC structures and subject-extraction RC/RC structures;
3. The lack of a difference between pre-verbal and post-verbal positions.

In contrast, the SPLT is consistent with the results observed here. The consistency between the theory's predictions and the results is depicted graphically in Figure 2. In this graph, the maximal complexity of each construction according to the SPLT is compared with the observed unacceptability ratings. There is clearly a very high correlation between the predicted complexities and the observed complexities. (Note that the base-lines in the graph are not relevant: only the correlation between the predicted complexities and the observed ratings is important.)

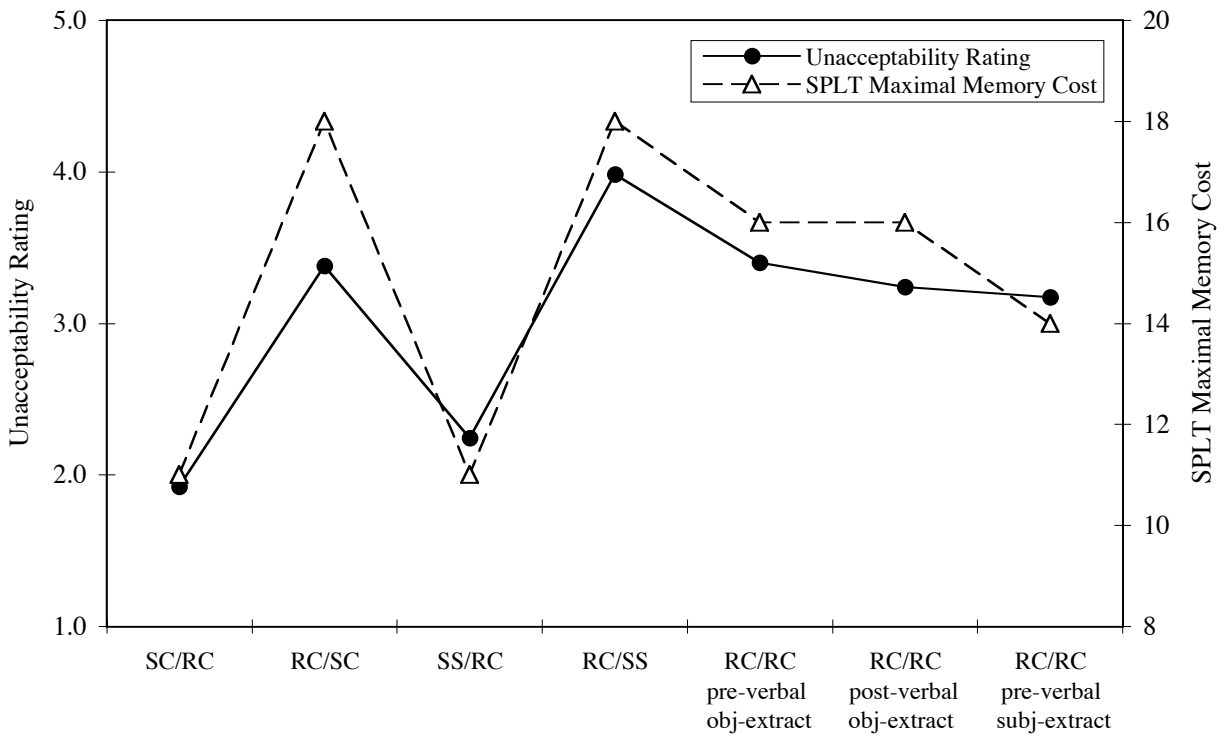


Figure 2: A comparison between the observed unacceptability ratings for the conditions in Experiment 1 and the SPLT's maximal memory complexity for these conditions as determined by the memory cost function  $M(n) = 2X*n + X$ .

Thus, according to this memory function, the borderline between acceptable and unacceptable structures occurs somewhere between 11X MUs — the maximal memory cost required to process the acceptable SC/RC and SS/RC structures — and 14X MUs — the maximal memory cost required to process the unacceptable RC/RC subject-extraction structures.

#### 4. An alternative account of the results from Experiment 1: Self-embedding

The SPLT offers one account of the data from Experiment 1. An alternative account put forward by Gibson & Thomas (1996) is the thematic-role-based theory together with one additional principle. Following Chomsky & Miller (1963) and Bever (1970), Gibson & Thomas propose that the additional memory cost property is one associated with processing a syntactic category that is *self-embedded* within another similar category (see also Thomas, 1995 and Babyonyshev & Gibson, 1995). A category  $X_1$  is said to be self-embedded within another category  $X_2$  if  $X_1$  is nested within  $X_2$  and both are of the same syntactic category. The motivation for the hypothesis that processing a self-embedded structure is associated with memory cost is that the parser may have a difficult time keeping track of two predicted categories that are extremely alike. The proposed definition of self-embedding memory cost is given in (26):

(26) The Property of self-embedding interference:

Associate a load of  $x_{SE}$  PLUs for a predicted category  $X_1$  whose head has not yet appeared which is embedded inside another predicted structure  $X_2$  whose head has also not yet appeared, when the extended projection features<sup>12</sup> of  $X_1$  are all included in the features of  $X_2$ .

The property of self-embedding interference applies when processing a relative clause structure center-embedded within another relative clause structure because at the point of predicting one relative clause CP/IP complex, a second relative clause CP/IP complex is predicted, whose features are identical. Hence, according to the definition in (26), there is an additional memory cost for processing all RC/RC structures at their most complex point.

In addition, Gibson & Thomas propose that there is a feature on a CP/IP which identifies it as either matrix or embedded (+/- matrix). The result of this claim is that the outer (first)

---

<sup>12</sup>The *extended projection* of a category is the highest functional projection of a category (Grimshaw, 1991). The extended projection of a VP therefore includes the IP and the CP, so that the features of the CP are relevant to computing the self-embedding status of a predicted IP and VP.

relative clause does not incur a self-embedding violation, because its features include a –matrix component which distinguish it from the predicted matrix CP/IP. As a result, there is exactly one self-embedding violation incurred by both pre- and post-verbal RC/RC structures.

Furthermore, Gibson & Thomas hypothesize 1) that the memory cost associated with self-embedding is near that associated with a thematic violation; and 2) that the maximal acceptable number of local violations for an acceptable sentence is four violations. Given these assumptions, the unacceptability of a RC/RC structure in post-verbal position is accounted for, because such a structure requires a parse state with five local violations (four thematic violations and one self-embedding violation), which is more than the proposed acceptable limit. The relative unacceptability of all of the RC/RC structures (with maximal complexities of five, five and six local violations respectively) as compared with the SC/RC structure is also accounted for, since the SC/RC structure does not include a self-embedded structure, and therefore has a maximal complexity of only four local violations.

In order to account for the complexity difference between SC/RC and RC/SC structures, Gibson & Thomas make one additional assumption. They assume that a relative clause CP category contains a feature that a sentential complement CP category does not: an operator feature “+operator”. Furthermore, they assume that the features of a sentential complement CP are properly included in the features of a relative clause CP. That is, the +operator feature is one of the features of a relative clause CP, but there is no corresponding –operator feature on a sentential complement clause CP. A sentential complement clause CP contains no operator feature at all, either + or –.

The result of this assumption is that there is a self-embedding interference cost associated with the processing of an RC/SC structure — because there is no distinguishing feature on a sentential complement CP — but there is no self-embedding interference cost associated with the processing of an SC/RC structure, because there is a feature, the +operator feature, that distinguishes the RC from the SC. As a result, the maximal complexity of the SC/RC structure is four violations, whereas the the maximal complexity of the RC/SC structure is five violations, thus crossing the acceptability boundary, and accounting for the complexity difference between the two.

As discussed in Gibson & Thomas (1996), the resulting theory is highly stipulative, requiring at least the following additional unsupported assumptions: 1) the assumption that self-embedding interference *per se* is associated with memory cost; 2) the assumption that feature-subset rather than feature-identity is what triggers self-embedding cost; 3) the assumption that the features of a

sentential complement CP are properly included in those of a relative clause CP; 4) the assumption that an embedded clause CP contains a feature –matrix (or +embedded) which distinguishes it from a matrix CP; and 5) the assumption of a categorical shift in perceived complexity between structures associated with maximally four local violations of the processing principles and those associated with maximally five local violations. In spite of the serious conceptual weaknesses of the resulting theory, the theory does generally account for the data from Experiment 1, and it makes a number of predictions that contrast with those of the SPLT. Some of these predictions are tested in Experiment 2.

## 5. Experiment 2

Two of the results from Experiment 1 were not completely convincing, so one of the points of Experiment 2 was to re-test these comparisons, on better designed conditions. First, the difference between subject- and object-extraction RC/RC structures was only a marginal difference in Experiment 1. Second, the lack of a difference between pre-verbal and post-verbal RC/RC structures could have been related to the confound in voice (active vs. passive) across the examples. In Experiment 2, these comparisons and a number of other related comparisons were tested in order to further test the SPLT relative to other theories, such as the thematic/self-embedded theory. Ten conditions were tested, divided conceptually into four groups. The predictions of these theories for the conditions to be tested in Experiment 2 are summarized in Table 2. The first group of four conditions varied both the position of a doubly center-embedded RC structure (pre-verbal, post-verbal) and the extraction type of the most embedded RC (subject-extraction, object-extraction). Three of these conditions were tested in Experiment 1, with the fourth (post-verbal, subject-extraction) added here to complete the 2x2 design. In order to avoid the voice confound in Experiment 1, the items were balanced for voice (active, passive) across the conditions.

(27) a. RC/RC-object-extraction, pre-verbal:

The school board which the students who the teachers were neglecting had angered troubled the superintendent.

b. RC/RC-object-extraction, post-verbal:

The superintendent was troubled by the school board which the students who the teachers were neglecting had angered.

c. RC/RC-subject-extraction, pre-verbal:

The school board which the teachers who were neglecting the students had angered troubled the superintendent.

d. RC/RC-subject-extraction, post-verbal:

The superintendent was troubled by the school board which the teachers who were neglecting the students had angered.

The thematic / self-embedding theory predicts effects of both position and extraction type, because 1) the pre-verbal conditions involve one more thematic violation than the post-verbal conditions and 2) the subject-extraction conditions also involve one less thematic violation than the object-extraction conditions. Furthermore, if the borderline between acceptable/unacceptable structures lies somewhere around four local constraint violations, then the subject-extraction post-verbal condition is the only acceptable condition of the four to be tested. If there is a marked jump in acceptability ratings at around this point, then an interaction between the position and extraction type is also predicted, such that the difference between pre- and post-verbal subject-extractions is larger than the difference between pre- and post-verbal object-extractions.

The SPLT predicts an effect of extraction type such that subject-extractions are easier to process than object-extractions, but no effect of position type is predicted, because there is no memory cost associated with predicting the top-level clause in the current memory buffer. Furthermore, no interactions with position type are predicted by the SPLT.

A condition consisting of a VP gerund modifying the subject of the outer relative clause is to be compared to the pre-verbal subject-extraction RC/RC as an additional control:

(28) RC/VP-gerund, pre-verbal:

The school board which the teachers neglecting the students had angered troubled the superintendent.

The only difference between (28) and (27c) is the addition of the relativizing pronoun “who”

and the auxiliary verb “were” in (27c). Hence (27c) contains an extra level of self-embedding and an extra role-less wh-pronoun at the point of maximal complexity, so the VP-gerund condition should be less complex according to the both the thematic role and the self-embedding properties.

As currently defined, the SPLT’s discourse-reference-based distance metric predicts no difference between the two conditions. Like the RC/RC subject-extraction condition, the maximal complexity of the VP-gerund condition occurs at the point of processing the NP “the students”, where the memory cost is  $2M(3)$  MUs corresponding to 1) the prediction of a verb for the relative clause; and 2) an empty-category position for the relative clause. Both of these predictions are associated with a cost of  $M(3)$  MUs because three new discourse referents have been processed since the categories were first predicted: 1) “the teachers”; 2) “neglecting”; and 3) “the students”. Because neither the wh-pronoun nor the auxiliary verb in the RC/RC condition introduce a new referent into the discourse, the maximal complexity for the VP-gerund condition is predicted to be the same as that for the RC/RC subject-extraction condition, and no complexity difference is therefore predicted.

However, the discourse-based memory cost function hypothesized in (11) may be an oversimplification of the human memory cost function. A number of variants of the SPLT’s memory cost hypothesis are possible which would predict a difference between the VP-gerund and RC/RC subject-extraction conditions. First, it is possible that all words cause an increase in memory load, not just words indicating new discourse referents (although words indicating new discourse referents may be associated with more memory load than other words). The added word-length between the point of predicting the first set of relative clause categories and their realization in the RC/RC subject-extraction may therefore cause this construction to be more complex than the VP-gerund construction. Second, in keeping with the discourse-reference hypothesis, it is possible that it is the tense features of a clause — and not the main verb itself — that cause an increment in memory cost for the prediction of other categories. According to this hypothesis, the maximal memory cost for processing the VP-gerund condition is only  $M(2)$  MUs, at the point of processing “the students”. At this point only two new discourse referents have been processed since the two relative clause categories were first predicted: “the teachers” and “the students”. The verb “neglecting” does not include tense features (it is a VP and not an S), so it does not cause a memory cost increment. A complexity difference between the VP-gerund and the RC/RC subject-extraction is therefore expected under this hypothesis.

The next comparison performed in Experiment 2 is one varying the pre- and post-verbal position



of RC/SC structures:

(29) a. RC/SC, pre-verbal: The school board which the reports that the teachers were neglecting the students had angered troubled the superintendent.

b. RC/SC, post-verbal: The superintendent was troubled by the school board which the reports that the teachers were neglecting the students had angered.

The self-embedding theory as defined above applies to sentential complements inside relative clauses, so that both of these conditions contain self-embeddings of one clause inside another. As a result, the thematic-role theory enhanced with the property of self-embedding predicts a significant difference between these two conditions. The maximal complexity of the pre-verbal condition occurs at the point of processing “the teachers”, at which point there are four local thematic violations and one self-embedding violation for a total of five local violations. On the other hand, the maximal complexity of the post-verbal condition is only four local violations: three thematic violations and one self-embedding violation. In contrast, the SPLT predicts no complexity difference according to the verbal position of the RC/SC structure, again because there is no memory cost associated with predicting the top-level clause in the current memory buffer..

The last three conditions in Experiment 2 were designed to see if different kinds of self-embedded structures are associated with different processing complexity. Towards this end, pre- and post-verbal structures consisting of a sentential complement self-embedded within another sentential complement were constructed, along with a pre-verbal control consisting of a sentential complement nested within a sentential subject, which is potentially not self-embedded. These are to be compared to each other and to the RC/RC conditions discussed above.

(30) a. SC/SC, pre-verbal: The expectation that the reports that the teachers were neglecting the students might anger the school board troubled the superintendent.

b. SC/SC, post-verbal: The superintendent was troubled by the expectation that the reports that the teachers were neglecting the students might anger the school board.

c. SS/SC: That the reports that the teachers were neglecting the students might anger the school board troubled the superintendent.

The thematic-role / self-embedding theory predicts a small contrast between the pre- and post-verbal SC/SC conditions (post-verbal easier), as well as a contrast between the SC/SC and SS/SC conditions (SS/SC easier), because each of these comparisons involves structures with a maximum of four and three violations respectively. This theory also predicts that the SC/SC conditions should

be much less complex than the RC/RC conditions because the SC/SC conditions do not contain relative pronouns, and thus have two fewer local thematic violations at their point of maximal complexity.

No contrast between the pre- and post-verbal conditions is predicted by the SPLT. Furthermore, no contrast is predicted between the SC/SC and SS/SC conditions. (However, because Experiment 1 found a processing difference between sentential subjects and sentential complements favoring sentential complements, it would not be surprising to find a similar difference here also. But this difference is not predicted by the SPLT.) The SPLT also predicts a difference between the SC/SC conditions and the RC/RC conditions, because there are fewer incomplete interfering dependencies in the SC/SC conditions.

## **5.1. Method**

### **5.1.1. Subjects**

Two groups of subjects participated in this experiment, fifty subjects in the first part and a forty additional subjects in the second, all native English speakers from the Boston, Massachusetts academic community (primarily undergraduates from MIT) for \$5.00 each, none of whom participated in Experiment 1.

### **5.1.2. Materials**

Forty items were constructed, each with the ten conditions discussed above. The compared conditions contained the same words and as much overlap of thematic relations among the NPs as was possible, with differences between conditions being matched for plausibility as much as possible, as described below. The meaning differences among the conditions to be compared are summarized below. The same shifts in relationships were maintained across all 40 items.

#### **1. RC/RC conditions:**

The pre- and post-verbal conditions had exactly the same meaning, with the only difference being the voice of the matrix verb (active, passive). The voice of the matrix verb was balanced across the forty items by making 20 active and 20 passive in the pre-verbal condition and similarly for the post-verbal condition.

Table 2: Predictions of linguistic complexity theories with respect to the experimental comparisons tested in Experiment 2.

Theory	Experimental comparisons		
	RC/RC: pre- vs. post-verbal	RC/RC: subject- vs. object-extraction	RC/RC-subject-extraction vs. RC/VP-gerund
Thematic-role and self-embedding theory	post-verbal easier especially subject-extraction	subject-extraction easier especially post-verbal	RC/VP-gerund easier
SPLT	no difference	subject-extraction easier	no difference or RC/VP-gerund easier depending on the memory cost function

Theory	Experimental comparisons			
	RC/SC: pre- vs. post-verbal	SC/SC: pre- vs. post-verbal	SC/SC vs. SS/SC	SC/SC vs. RC/RC
Thematic role and self-embedding theory	post-verbal easier	post-verbal easier	SS/SC easier	SC/SC easier
SPLT	no difference	no difference	no difference	SC/SC easier

**Key:**

RC/RC: A relative clause center-embedded within another relative clause

RC/VP-gerund: An object-extraction relative clause with a VP gerund modifying its subject

SC/SC: A sentential complement of an NP which includes a further sentential complement of its subject NP

SS/SC: A sentential subject which includes a sentential complement of its subject NP

The meaning differences between the subject- and object-extraction conditions can best be seen with reference to the general form of the conditions given in (31).

(31) a. Object-extraction:

NP<sub>1</sub> who/which NP<sub>2</sub> who/which NP<sub>3</sub> V<sub>1</sub> V<sub>2</sub> V<sub>3</sub> NP<sub>4</sub>

b. Subject-extraction:

NP<sub>1</sub> who/which NP<sub>3</sub> who/which V<sub>1</sub> NP<sub>2</sub> V<sub>2</sub> V<sub>3</sub> NP<sub>4</sub>.

The thematic relationships in the object-extraction condition consist of 1) NP<sub>3</sub> V<sub>1</sub> NP<sub>2</sub>; 2) NP<sub>2</sub> V<sub>2</sub> NP<sub>1</sub>; and 3) NP<sub>1</sub> V<sub>3</sub> NP<sub>4</sub>. The thematic relationships in the subject-extraction condition consist of 1) NP<sub>3</sub> V<sub>1</sub> NP<sub>2</sub>; 2) NP<sub>3</sub> V<sub>2</sub> NP<sub>1</sub>; and 3) NP<sub>1</sub> V<sub>3</sub> NP<sub>4</sub>. Thus the only difference between the two is that NP<sub>2</sub> is the subject of V<sub>2</sub> in the object-extraction condition, while NP<sub>3</sub> is the subject of V<sub>2</sub> in the subject-extraction condition.

Recall that the design of the subject/object-extraction items in Experiment 1 necessitated the use of NPs and predicates that were roughly equally plausible as subject-verb-object or as object-verb-subject. In practice, it was difficult to write such items. In contrast, the design of the items in this experiment makes constructing items of equal plausibility for both extraction types much easier, since all one needs to do is select NP<sub>2</sub> and NP<sub>3</sub> from a similar semantic class. So, in (27) for example, NP<sub>2</sub> “the students” and NP<sub>3</sub> “the teachers” are roughly equally plausible as subjects of the predicate “angered the school board”.

2. RC/RC-subject-extraction vs. RC/VP-gerund. There are no differences in meaning in this comparison. All items were constructed with the most embedded verb in the progressive so that the only difference between the conditions is the addition of the words “who/which was/were”.
3. The RC/SC and SC/SC pre- vs. post-verbal comparisons. There are no meaning differences in either of these comparisons. The items were balanced for the voice of the matrix verb as discussed in the RC/RC conditions.
4. SS/SC vs. SC/SC. The only difference between these conditions is the addition of one SC-taking NP in the SC/SC condition.
5. RC/RC vs. SC/SC. The meaning differences here can best be seen with reference to (32):

(32) a. RC/RC: NP<sub>1</sub> who/which NP<sub>2</sub> who/which NP<sub>3</sub> V<sub>1</sub> V<sub>2</sub> V<sub>3</sub> NP<sub>4</sub>

b. SC/SC: NP<sub>5</sub> that NP<sub>6</sub> that NP<sub>3</sub> V<sub>1</sub> NP<sub>2</sub> V<sub>2</sub> NP<sub>1</sub> V<sub>3</sub> NP<sub>4</sub>

The thematic relationships in the RC/RC condition consist of 1) NP<sub>3</sub> V<sub>1</sub> NP<sub>2</sub>; 2) NP<sub>2</sub> V<sub>2</sub> NP<sub>1</sub>; and 3) NP<sub>1</sub> V<sub>3</sub> NP<sub>4</sub>. The thematic relationships in the SC/SC condition consist of 1) NP<sub>3</sub> V<sub>1</sub> NP<sub>2</sub>; 2) NP<sub>6</sub> V<sub>2</sub> NP<sub>1</sub>; and 3) NP<sub>5</sub> V<sub>3</sub> NP<sub>4</sub>. Thus the only differences between the two are 1) NP<sub>2</sub> is the subject of V<sub>2</sub> in the RC/RC condition, while NP<sub>6</sub> is the subject of V<sub>2</sub> in the SC/SC condition; and 2) NP<sub>1</sub> is the subject of V<sub>3</sub> in the RC/RC condition, while NP<sub>5</sub> is the subject of V<sub>3</sub> in the SC/SC condition. The nouns and verbs were selected to make these differences as equally plausible as possible.

The examples were constructed using as many different verbs as possible. Fifty-five different verbs were used across all items. Most verbs appeared in either one or two items and no verb appeared in more than three items. The items also used as many different NPs as possible. Forty-two different sentential-complement taking noun heads were used in all, with no head noun appearing in more than three items. Furthermore, the items were balanced so that no subject saw more than two instances of any of these nouns.

The experimental items were combined with 86 fillers (including 6 practice items) of approximately the same length as the experimental items, to form ten lists. The experimental items were counterbalanced across the lists so that each list contained exactly one version of every item. A list of materials is given in Appendix B.

Two separate versions of the questionnaire were constructed, which differed only in the set of filler items in each. One set of fillers was constructed to be much more complex than the other. The group of 50 subjects completed the survey with the harder fillers, while the group of 40 subjects completed the survey with the easier fillers.

### **5.1.3. Procedure**

The procedure was the same as for Experiment 1. Subjects completed the task within about 30 minutes.

## **5.2. Results**

The mean for the ratings of the fillers in the hard-filler questionnaire was 3.63 (standard error = 0.05) while the corresponding mean for the fillers in the easy-filler questionnaire was 2.11 (standard error = 0.07). Thus, as expected, the fillers in the hard-filler questionnaire were rated as much

more unacceptable than the fillers in the easy-filler questionnaire. For purposes of comparison across experiments, the fillers in Experiment 1 had a mean rating of 2.45 (standard error = 0.07), much closer to the ratings for the fillers in the easy-filler questionnaire here. In spite of the different fillers, analyses reveal that the pattern of results is almost identical across the two filler types, with the primary difference being that the ratings for the easy-filler version of the experiment were generally 0.40–0.50 rating points greater (less acceptable) than those for the hard-filler survey. This difference follows from Bever’s (1970; see also Schütze, 1996, for a recent summary) hypothesis that grammaticality/acceptability judgements are always relative to other items being queried, so that the harder the fillers, the better the ratings of the target items, and conversely. The means and standard errors for each of the target conditions in Experiment 2 are given in Table 3 for the hard filler questionnaire (see also figures 3 and 4 for a graphical depiction of these results) and Table 4 for the easy filler questionnaire (see also figures 5 and 6 for a graphical depiction of these results).

Table 3: Complexity ratings for the structures tested in Experiment 2 with the more complex set of fillers (standard errors in parentheses). Higher ratings reflect higher complexity.

	RC/RC conditions				RC/VP-gerund
	subject-extract	subject-extract	object-extract	object-extract	
	pre-verbal	post-verbal	pre-verbal	post-verbal	
<b>Rating</b>	2.52 (.11)	2.64 (.11)	3.02 (.11)	2.79 (.11)	2.14 (.10)

	RC/SC		SC/SC		SS/SC
	pre-verbal	post-verbal	pre-verbal	post-verbal	
<b>Rating</b>	2.90 (.11)	2.92 (.10)	2.19 (.09)	2.37 (.10)	2.42 (.13)

All comparisons that were non-significant in the hard-filler questionnaire pool were also non-significant in the easy-filler questionnaire pool. Furthermore, all comparisons that were significant in the hard-filler sub-experiment were also significant in the easy-filler sub-experiment. Because of the similarity of the results in the two sub-experiments, only statistics for the sub-experiment with the larger subject pool, the hard-filler sub-experiment, will be presented.

First, let us consider the RC/RC conditions. There was a significant effect of extraction type, with the object-extraction rated as more complex than the subject-extraction (mean for object-

Table 4: Complexity ratings for the structures tested in Experiment 2 with the less complex set of fillers (standard errors in parentheses). Higher ratings reflect higher complexity.

	RC/RC conditions				RC/VP-gerund
	subject-extract pre-verbal	subject-extract post-verbal	object-extract pre-verbal	object-extract post-verbal	
<b>Rating</b>	2.94 (.10)	3.06 (.12)	3.51 (.12)	3.33 (.12)	2.69 (.13)

	RC/SC		SC/SC		SS/SC
	pre-verbal	post-verbal	pre-verbal	post-verbal	
<b>Rating</b>	3.39 (.12)	3.39 (.14)	2.70 (.13)	2.70 (.12)	2.71 (.14)

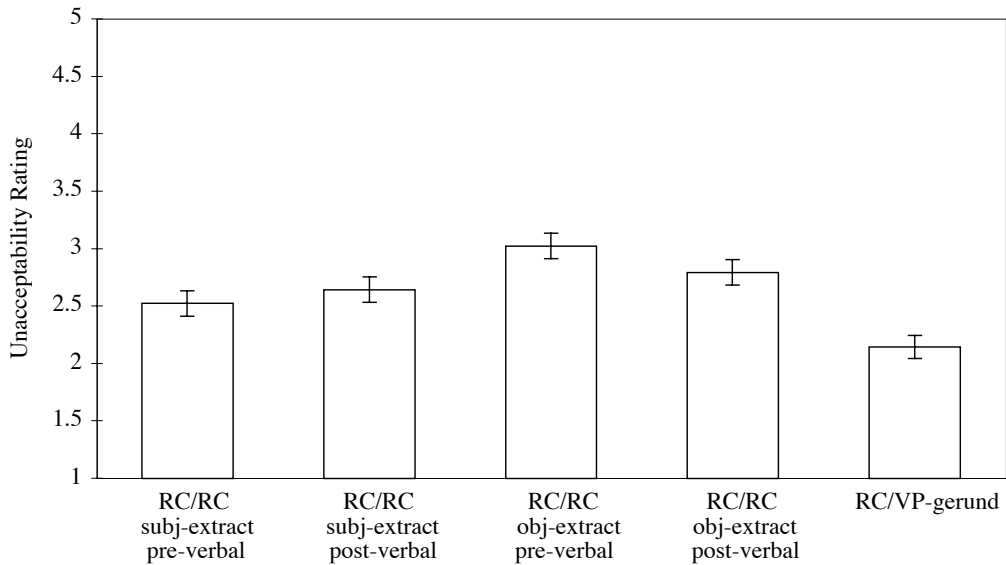


Figure 3: Complexity ratings for the first five conditions in Experiment 2. The fillers in this part of Experiment 2 were the more complex set of fillers (the “hard” fillers).

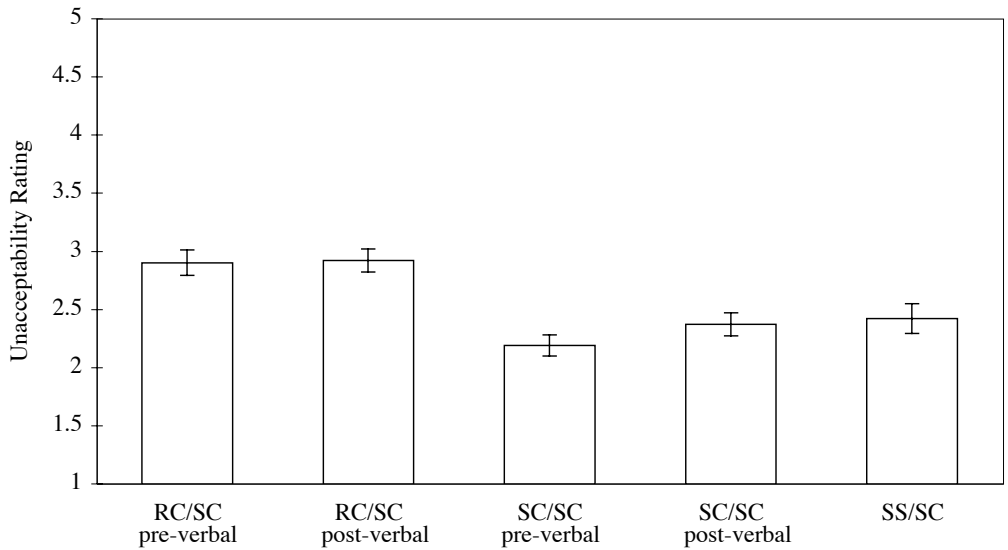


Figure 4: Complexity ratings for the last five conditions in Experiment 2. The fillers in this part of Experiment 2 were the more complex set of fillers (the “hard” fillers).

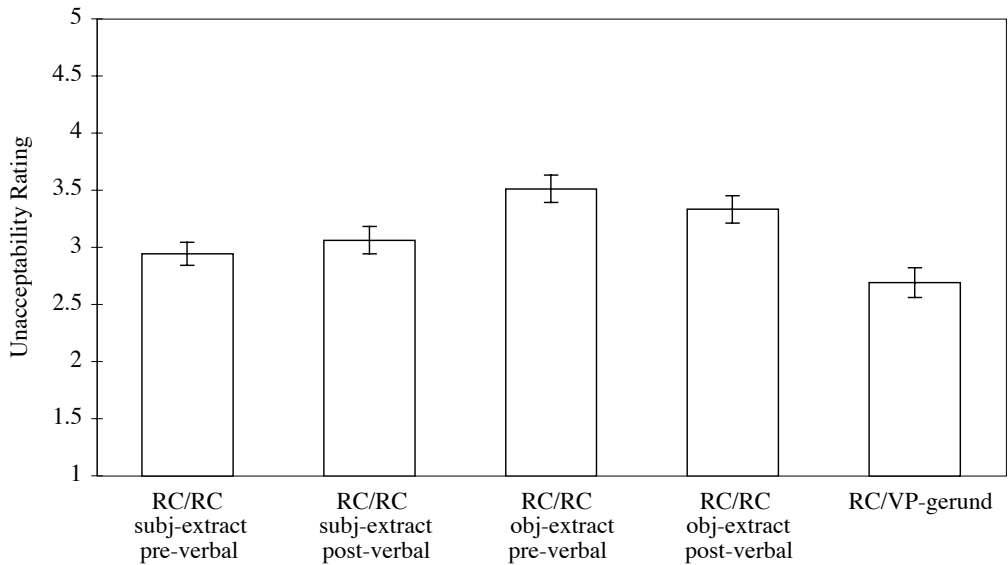


Figure 5: Complexity ratings for the first five conditions in Experiment 2. The fillers in this part of Experiment 2 were the less complex set of fillers (the “easy” fillers).



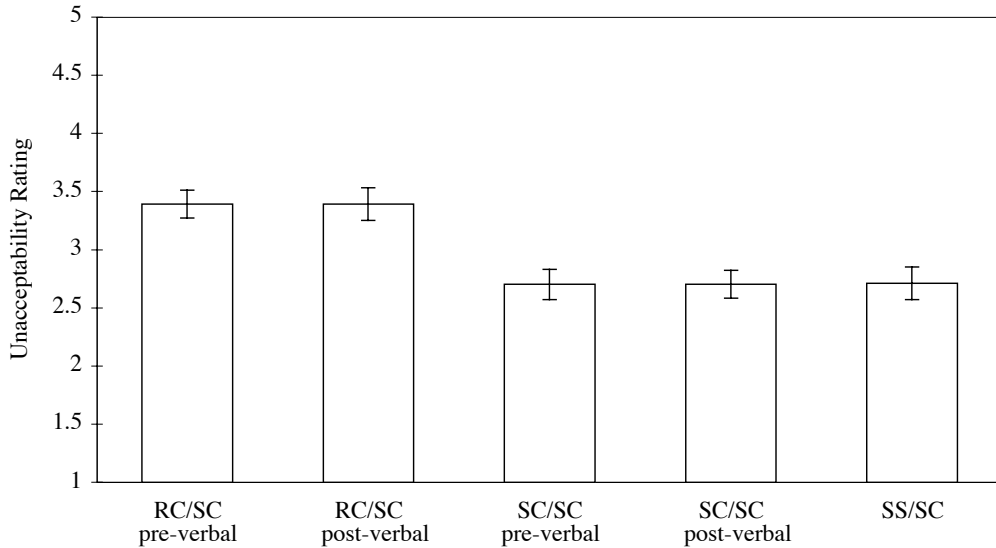


Figure 6: Complexity ratings for the last five conditions in Experiment 2. The fillers in this part of Experiment 2 were the less complex set of fillers (the “easy” fillers).

extraction = 2.91; mean for subject-extraction = 2.58;  $F1(1, 49) = 22.0, p < .001$ ;  $F2(1, 39) = 11.3, p < .005$ ). However, there was no effect of position (mean for pre-verbal = 2.77; mean for post-verbal = 2.72;  $F_s < 1$ ) and no interaction. Thus the results from Experiment 2 replicate those from Experiment 1 in finding an effect of extraction type, but not of position. Furthermore, none of the other comparisons of pre-verbal and post-verbal position show an effect of position. Thus the hint of a difference between the pre- and post-verbal positions observed in Experiment 1 may have been due to the confounding influence of voice, such that the distinguishability of the voice on the verb may have improved acceptability.

The comparison between the RC/RC-subject-extraction and RC/VP-gerund revealed a significant difference ( $F1(1, 49) = 13.2, p < .001$ ;  $F2(1, 39) = 8.8, p < .01$ ). The position comparisons in the RC/SC items revealed no significant differences ( $F_s < 1$ ). The position comparisons in the SC/SC items also showed no differences.

The SS/SC vs. SC/SC comparison revealed a marginal difference, such that the SC/SC was favored among the hard-filler questionnaire items ( $F1(1, 49) = 3.0, p < .10$ ;  $F2(1, 39) = 5.9, p < .05$ ) with no differences among the easy-filler questionnaire items ( $F_s < 1$ ). Finally, there was a significant difference between the SC/SC and RC/RC conditions in both sets of fillers (statistics for the hard-filler questionnaire:  $F1(1, 49) = 44.5, p < .001$ ;  $F2(1, 39) = 51.0, p < .001$ ).

### 5.3. Discussion

The results of Experiment 2 are not consistent with the thematic-role / self-embedding theory discussed in Section 4. Most critically, none of the pre-verbal / post-verbal position effects predicted by this theory were observed. Furthermore, the predicted contrast between the SS/SC and SC/SC was not observed. In addition to making these empirically incorrect predictions, the theory itself is extremely stipulative. This theory is therefore abandoned.

In contrast, the results of Experiment 2 are generally consistent with the SPLT. In particular, the SPLT correctly predicts the following results from Experiment 2:

1. The lack of pre-verbal / post-verbal contrasts, in accordance with the hypothesis that there is no memory cost associated with maintaining the top-level clause prediction.
2. Subject-extractions are easier to process than object-extraction in RC/RC structures.
3. RC/RC structures are more complex than SC/SC structures.

The only result from Experiment 2 which is not predicted by the SPLT in its original form is the result that the RC/VP-gerund construction is easier to process than the subject-extracted RC/RC. Two possible accounts of this result within the SPLT framework were discussed above: 1) that there is memory cost associated with processing all words between a category prediction and its realization, not only those representing new discourse referents; and 2) that the memory cost function for each predicted category is incremented when processing the tense features of a clause, not the main verb of that clause. We will adopt the second approach, because it does not affect any of the other predictions made by the SPLT that have been discussed thus far. Under this hypothesis, the maximal complexity of the RC/VP-gerund construction is  $2M(2)$  MUs, which is less than the maximal complexity of the subject-extracted RC/RC construction, at  $2M(3)$  MUs.

Figure 7 presents a comparison between the observed ratings of the easy-filler experimental conditions in Experiment 2 and their predicted maximal complexities according to the SPLT. As in Experiment 1, the predictions of the SPLT are highly correlated with the observed complexity results. As in the comparison between the SPLT and the results of Experiment 1, the base-lines in the graph are not relevant, because people's complexity ratings are relative to the complexity of the filler sentences in the experiment. The correlation between the predicted complexities and the observed ratings is the relevant feature of the graph to observe.

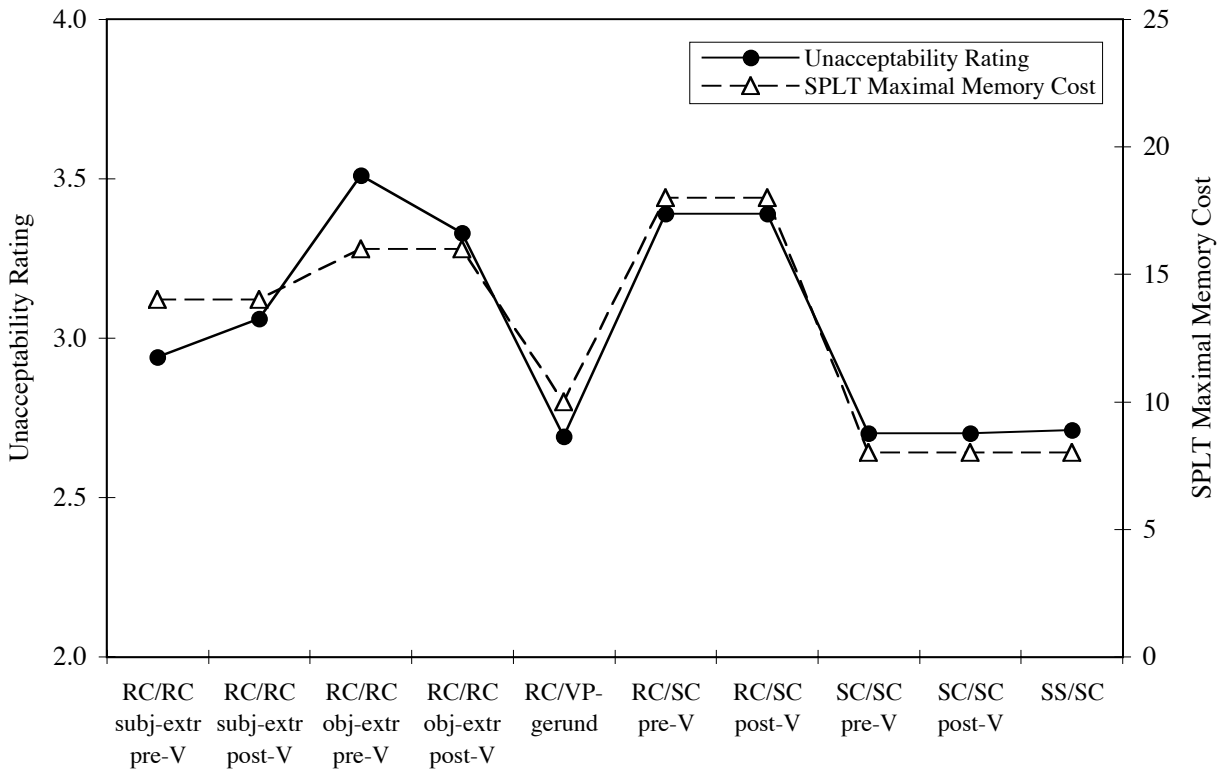


Figure 7: A comparison between the observed unacceptability ratings for the conditions in Experiment 2 and the SPLT’s maximal memory complexities for these constructions.

## 6. General Discussion

The SPLT works very well in accounting for the results of the experiments presented here. The SPLT also accounts for a wide range of further linguistic complexity phenomena which are superficially unrelated to the ones examined here. Four representative classes of acceptability phenomena from Gibson (1997) are summarized in this section, because of their relevance to the theories evaluated here. See Gibson (1997) for a discussion of how the SPLT accounts for other kinds of linguistic complexity phenomena, including preferences in ambiguous structures (Gibson, Grodner & Tunstall, 1997) and “heaviness” effects, such that sentences are easier to understand when larger phrases are placed later (Bever, 1970; Hawkins, 1990; 1994).

### 6.1. Complex structures consisting of only two clauses

According to the principle of two sentences, a nested structure is unacceptable if and only if its parse includes a state with three incomplete sentence nodes. We have seen a number of counterexamples to this claim. First of all, we have seen a number of constructions which are processable in spite of the fact that each contains a state with three incomplete sentence nodes. Examples of this kind of falsification of the principle of two sentences are the SC/RC, SS/RC, SC/SC and SS/SC constructions. We have also seen an example of a construction — the post-verbal object-extracted RC/RC — which is *not* processable in spite of the fact that its parse involves only at most two incomplete sentence nodes at any point.

If the SPLT is on the right track, then it should be possible to construct further example constructions of this second type. This turns out to be the case, as the construction in (33c) demonstrates:

- (33) a. Who did the artist who the woman met at the party talk to?  
b. ?# Who did the artist who the woman from the college met at the party talk to?  
c. # Who did the artist who the woman from the college in the suburbs met at the party talk to?

Each of the examples in (33) consists of a relative clause modifying the subject of an object-extracted wh-question, so each contains exactly two clauses. The relevant difference among the three is that the distance between the predicted categories associated with the relative clause and the wh-question increases over the three sentences. In particular, the maximal memory complexity

of (33a) is  $M(3)+2M(1)$  MUs, which occurs upon processing the most embedded subject noun “woman”. Hence this structure should be roughly comparable in perceived complexity to an SC/RC structure. Adding the PP “from the college” following this subject as in (33b) pushes the maximal complexity up to  $M(4)+2M(2)$  MUs at “college”. Adding an additional PP pushes the maximal complexity up to  $M(5)+2M(3)$  MUs following the third PP, as in (33c), causing a clear case of unacceptability.

Thus, unacceptable structures like (33c) provide another kind of counterexample for the principle of two sentences. The unprocessability of this construction is also a counterexample to the thematic-role-based theory. In particular, there are at most four local thematic violations during its processing, which means that it should be acceptable, but it is unacceptable.

## 6.2. Extractions from prepositional phrases

Another unprocessable construction which is not accounted for by either the principle of two sentences or the thematic-role-based theory is the nested stranded prepositional phrase construction in (34) (Pickering & Barry, 1991):

- (34) a. # This is the saucer which Mary put the cup which Elma poured the milk into on.  
b. # This is the table which Fred broke a plate which Jake saw a spider on near.

These structures are unacceptable in spite of the fact that there are at most two incomplete sentences in the parse of each: at “Elma” in (34a), and at “Jake” in (34b). Similarly, there are at most four incomplete thematic relations in the parse of each. So neither the principle of two sentences nor the thematic-role-based theory predicts the unacceptability of this construction.

In contrast, the unacceptability of these structures is predicted by the SPLT, because the maximal memory complexity of each is  $M(6)+M(3)$  MUs, more than the capacity of linguistic short term memory. This maximal complexity occurs at the point of processing the NP before the two stranded prepositions. For example, in (34a), there are two wh-pronouns which require NP-gap positions at the point of processing “milk”: one for the first instance of “which” and one for the second instance. At this point seven new discourse referents have been processed since the first instance of “which” was processed, and three new discourse referents have been processed since the second instance, leading to a complexity of  $M(7)+M(3)$  MUs. See Gibson (1997) for a discussion of the SPLT’s account of other complex preposition-stranding constructions.

### 6.3. Cross-linguistic support for the SPLT: Japanese linguistic complexity

In addition to explaining complexity phenomena from English, the SPLT also explains complexity phenomena in languages whose structure is very different from that of English. For example, a number of syntactic complexity results from the processing of head-final languages like Japanese are also explained within the SPLT. For example, Babyonyshev & Gibson (1997) provide experimental evidence from acceptability judgement questionnaires demonstrating that the addition of extra nominal arguments to the most embedded clause in nested structures makes the structures harder to comprehend. The structures investigated by Babyonyshev & Gibson have the form in (35):

- (35) a. Obasan-wa [<sub>S'</sub> [<sub>S</sub> bebiisitaa-ga [<sub>S'</sub> [<sub>S</sub> imooto-ga naita ] to ] itta ] to ] omotteiru  
aunt-top babysitter-nom younger-sister-nom cried that said that thinks  
“My aunt thinks that the babysitter said that my younger sister cried”
- b. Obasan-wa [<sub>S'</sub> [<sub>S</sub> bebiisitaa-ga [<sub>S'</sub> [<sub>S</sub> ani-ga imooto-o ijimeta ] to ] itta ] to ] omotteiru  
aunt-top babysitter-nom older-brother-nom younger-sister-acc bullied that said that thinks  
“My aunt thinks that the babysitter said that my older brother bullied my younger sister”

The maximal complexity for (35a) according to the SPLT is  $2M(2)+M(1)$  MUs, which occurs upon processing the verb “naita” (“cried”). At this point, the following category predictions are associated with memory cost:

1. 0 MUs for the prediction of the matrix verb;
2.  $M(2)$  MUs for the prediction of the first embedded verb;
3.  $M(2)$  MUs for the prediction of the outer complementizer;
4.  $M(1)$  MUs for the prediction of the most embedded complementizer.

On the other hand, the maximal complexity for (35b) is  $2M(3)+M(2)$  MUs, which occurs at the point of processing the verb “ijimeta” (“bullied”), and corresponds to the same category predictions as for (35a), but carried for one additional new discourse referent, the embedded direct object “imooto-o” (“sister-acc”) of the transitive verb. The complexity breaks down into the following components:

1. 0 MUs for the prediction of the matrix verb;
2.  $M(3)$  MUs for the prediction of the first embedded verb;

3. M(3) MUs for the prediction of the outer complementizer;
4. M(2) MUs for the prediction of the most embedded complementizer.

The maximal cost is therefore more than that associated with processing a structure like (35a), and the contrast is explained.

Note that the principle of two sentences does not explain the observed acceptability contrast, because each of the structures violates the principle equally. Although the thematic-role-based theory is consistent with these complexity observations, Babyonyshev & Gibson (1997) provide additional processing evidence that contradicts the predictions of the thematic-role-based theory. Babyonyshev & Gibson (1997) also demonstrate that the SPLT explains the contrast between the two embeddings of a relative clause and a sentential complement, the analysis of which is similar to the analysis of the parallel English constructions. See Babyonyshev & Gibson (1997) for more details about how this works out in Japanese.

#### 6.4. Center-embedding complexity: The length of the most embedded structure

The Japanese evidence due to Babyonyshev & Gibson shows that the number of NP arguments in the most nested clause of a structure has an important effect on its processing complexity. Similar complexity differences should be present in all languages, including English relative clause constructions. In particular, making the most embedded clause in an RC/subject-extraction-RC structure intransitive rather than transitive should make such a structure easier to process:

(36) a. RC/RC, subject-extraction, intransitive:

The administrator who the intern who was sleeping had bothered lost the medical reports.

b. RC/RC, subject-extraction, transitive:

The administrator who the intern who was supervising the nurse had bothered lost the medical reports.

According to the SPLT, the maximal complexity associated with the processing of (36b) is 2M(3) MUs, while the maximal complexity associated with the processing of (36a) is only 2M(2) MUs. Thus the SPLT locality theory predicts a contrast between (36b) and (36a). This predicted difference is due to the fact that the distance between the RC fillers and their role-assigners is shorter in (36a) than in (36b). In contrast, all other theories of syntactic complexity predict no difference between the processing of (36b) and (36a), because each involves the same number of

local violations at the point of maximal complexity, when processing the most embedded subject and following verb.

## 7. Summary and conclusions

The Syntactic Prediction Locality Theory provides a conceptually simple theoretical account of a broad range of linguistic complexity phenomena, including the results from the English questionnaire experiments reported here. According to this theory, the reason that a nested structure is complex has nothing to do with its nested structure *per se*, but rather is derived from the fact that 1) nested structures typically include a parse state at which a number of syntactic categories are predicted, and 2) the distance is substantial between the point at which the categories are predicted and the point that they are resolved. Evidence from other experiments not discussed in depth here has suggested that the best way to quantify distance within the SPLT is in terms of the number of new discourse structures separating a category prediction and its resolution. The results presented here are consistent with this hypothesis.

The most interesting results of the experiments reported here are enumerated as follows:

1. Embedding a relative clause (RC) within a sentential complement (SC) of a noun results in a processable structure, whereas the reverse embedding, consisting of an SC within an RC, results in a much less processable structure.
2. An SC/RC structure (RC within SC) is also much more processable than an RC within another RC (an RC/RC structure).
3. An RC/SC structure (SC within RC) is of similar complexity as an RC/RC structure.
4. There is no interaction between clause type (sentential complement of a noun / sentential subject) and embedding order relative to a relative clause (RC-first, RC-final), suggesting that the unacceptability the embedded sentential subject condition is caused by processing complexity and not a difference in grammaticality.
5. Subject-extracted RC/RC structures are less complex than object-extracted RC/RC structures. This is the same result as has been observed in singly-embedded structures (e.g., Hakes, Evans & Brannon, 1976; Holmes & O'Regan, 1981; King & Just, 1991), but the intuitions are much less obvious in the more complex comparison.



6. The lack of any pre- vs. post-matrix-verb complexity effects for any of the constructions tested here, including RC/RC structures, RC/SC structures and SC/SC structures.
7. The lower complexity of SC/SC structures as compared with RC/RC structures.

The broad empirical coverage together with the conceptual simplicity of the SPLT suggest that locality between the point of predicting a category and its realization plays a fundamental role in language comprehension.

## 8. References

- Anderson, J.R. (1994). *Learning and memory*. Wiley, New York, NY.
- Babyonyshev, M. & Gibson E. (1995). Processing Overload in Japanese. MIT working papers in linguistics, MIT, Cambridge, MA.
- Babyonyshev, M. & Gibson E. (1997). Japanese linguistic complexity. Manuscript, MIT, Cambridge, MA.
- Baddeley, A. (1990). *Human memory*. Allyn and Bacon, Needham Heights, MA.
- Bever, T.G. (1970). The cognitive basis for linguistic structures. in J.R. Hayes (Ed.), *Cognition and the development of language*, John Wiley, New York, NY.
- Bever T.G. & Townsend, D. (1979). Perceptual mechanisms and formal properties of main and subordinate clauses. In W.E. Cooper and E.C.T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Lawrence Erlbaum, Hillsdale, NJ, 159-226.
- Bresnan, J. (1982). *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, MA.
- Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception and Psychophysics*, 12, 73-76.
- Chomsky, N. (1957). *Syntactic Structures*, Mouton, The Hague, The Netherlands.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.
- Chomsky, N. (1986). *Barriers*, Linguistic Inquiry Monograph 13, MIT Press, Cambridge, MA.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In Hale, Kenneth and Samuel Jay Keyser (Eds.) *The view from building 20*, MIT Press, Cambridge, MA, 1-52.
- Chomsky, N. (1995). *The minimalist program*, MIT Press, Cambridge, MA.
- Chomsky, N. & Miller, G.A. (1963). Introduction to the formal analysis of natural languages. In R.D. Luce, R.R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology*, volume 2, 269-321, John Wiley, New York, NY.

- Cowper, E.A. (1976). *Constraints on sentence complexity: A model for syntactic processing*, Ph.D. dissertation, Brown University, Providence, RI.
- De Roeck, A. Johnson, R., King, M., Rosner, M., Sampson, G. & Varile, N. (1982). A myth about center-embedding. *Lingua*, 58, 327-340.
- Forster, K.I. & Olbrei, I. (1973). Semantic heuristics and syntactic analysis. *Cognition*, 2, 319-347.
- Frank, R. (1992). *Syntactic locality and tree-adjoining grammar: Grammatical, acquisition and processing perspectives*, Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA.
- Frazier, L. (1985). Syntactic complexity. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language processing: Psychological, computational and theoretical perspectives*, Cambridge
- Frazier, L., & Fodor, J.D. (1978). The Sausage Machine: A new two-stage parsing model. *Cognition* 6, 291-325.
- Frazier, L. & Rayner, K. (1988). Parameterizing the language-processing system: Left- vs. right-branching within and across languages. In J.A. Hawkins (ed.), *Explaining language universals*, pp. 247-279. Basil Blackwell, Oxford, UK,
- Garrod, S. & Sanford, A. J. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, 16, 77-90.
- Garrod, S.C. & Sanford, A.J. (1994). Resolving Sentences in a Discourse Context: How Discourse Representation Affects Language Understanding. In Gernsbacher, M. A. (Ed.), *Handbook of Psycholinguistics*. Academic Press, San Diego, 675-698.
- Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*, Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.
- Gibson, E. (1997). Linguistic complexity: Locality of syntactic dependencies. MIT manuscript.
- Gibson, E., Grodner, D. & Tunstall, S. (1997). Evidence for a syntactic complexity metric in ambiguity resolution independent of lexical effects. Poster presented at the 10th CUNY sentence processing conference, Santa Monica, CA.
- Gibson, E., Hickok, G. & Schütze, C. (1994). Processing empty categories: A parallel approach. *Journal of Psycholinguistic Research*, 23, 381-406.
- Gibson, E., Pearlmutter, N., Canseco-Gonzales, E. & Hickok, G. (1996). Recency Preference in the Human Sentence Processing Mechanism. *Cognition*, 59, 23-59.
- Gibson, E. & Thomas, J. (1996). The processing complexity of English center-embedded and self-embedded structures. In C. Schütze (Ed.) *Proceedings of the NELS 26 workshop on language processing*, MIT occasional papers in linguistics 9, Cambridge, MA.
- Gibson, E. & Thomas, J. (1997). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. MIT manuscript.
- Grimshaw, J. (1991). Extended projection. Brandeis University manuscript.
- Hakes, B., Evans, J. & Brannon, L. (1976). Understanding sentences with relative clauses. *Memory and Cognition*, 4, 283-296.

- Hakuta, K. (1981). Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition* 9, 197-236.
- Haliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. Longman, London.
- Haviland, S.E. & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.
- Hawkins, J.A. (1990). A parsing theory of word order universals. *Linguistic Inquiry*, 21, 223-262.
- Hawkins, J.A. (1994). *A performance theory of order and constituency*, Cambridge University Press, Cambridge, UK.
- Heim, I. (1982). *The semantics of definite and indefinite noun phrases*, Ph.D. thesis, University of Massachusetts, Amherst, MA.
- Holmes, V.M., O'Regan, J.K. (1981). Eye fixation patterns during the reading of relative clause sentences. *Journal of Verbal Learning and Verbal Behavior*, 20, 417-430.
- Jarvella, R.J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 10, 409-416.
- Jarvella, R.J. (1979). Immediate memory and discourse processing. In G.H. Bower (Ed.) *The psychology of learning and motivation*, vol. 13. Academic Press, New York, pp. 379-421.
- Just, M.A. and Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Kac, M.B. (1981). Center-embedding revisited. In *Proceedings of the third annual conference of the Cognitive Science Society*, Lawrence Erlbaum, Hillsdale, NJ, pp. 123-124.
- Kamp, H. (1981). A theory of truth and semantic representations. In J. Groenendijk (Ed.) *Formal methods in the study of language*, Mathematisch Centrum, Amsterdam, The Netherlands.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition* 2, 15-47.
- King, J. & Just, M.A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.
- Koster, J. (1978) "Why Subject Sentences Don't Exist," in S.J. Keyser (ed.), *Recent Transformational Studies in European Languages*, MIT Press, Cambridge, MA.
- Lewis, R. (1993) *An architecturally-based theory of human sentence comprehension*, Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.
- Lewis, R. (1996). *A theory of grammatical but unacceptable embeddings*, *Journal of Psycholinguistic Research*, 25, 93-116.
- MacWhinney, B. (1987). The Competition Model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*, Lawrence Erlbaum, Hillsdale, NJ, 249-308.
- Marslen-Wilson, W., Tyler, L.K. & Seidenberg, M. (1978). Sentence processing and the clause boundary. In W.J.M. Levelt and G.B. Flores d'Arcais (Eds.), *Studies in the perception of language*, John Wiley, New York, NY, 219-246.

- Miller, G.A. & Chomsky, N. (1963). Finitary models of language users. In R.D. Luce, R.R. Bush and E. Galanter (Eds.) *Handbook of Mathematical Psychology*, volume 2, John Wiley, New York, NY, pp. 419-491.
- Miller, G.A. & Isard, S. (1964). Free recall of self-embedded English sentences. *Information and Control* 7, 292-303.
- Milsark, G.L. (1983). On Length and Structure in Sentence Parsing. *Cognition*, 13, 129-134.
- Pickering, M. & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, 6, 229-259.
- Pollard, C. & Sag, I.A. (1994). *Head-driven phrase structure grammar*, University of Chicago Press, Chicago, IL.
- Roberts, R., Hooker, C. & Gibson, E. (1997). Sentence memory capacity and working memory. Poster presented at the 10th CUNY sentence processing conference, Santa Monica, CA.
- Schlesinger, I.M. (1968). *Sentence structure and the reading process*, Mouton, The Hague, The Netherlands.
- Schütze, C. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. University of Chicago Press, Chicago, IL.
- Stabler, E.P. (1994). The finite connectivity of linguistic structures. In Clifton, C., L. Frazier, and K. Rayner (Eds.), *Perspectives in Sentence Processing*, Lawrence Erlbaum, 303-336.
- Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, 23, 295-322.
- Stolz, W.S. (1967) "A Study of the Ability to Decode Grammatically Novel Sentences," *Journal of Verbal Learning and Verbal Behavior*, 6, 867-873.
- Thomas, J. (1995). Center-embedding and self-embedding in human language processing. Masters thesis, MIT.
- Waugh N.C. & Norman, D.A. (1965). Primary memory. *Psychological Review*, 72, 89-104.
- Yngve, V.H. (1960). A Model and an Hypothesis for Language Structure. *Proceedings of the American Philosophical Society*, 104, 444-466.

## A Experimental Items for Experiment 1

The items used in Experiment 1 are listed below. All seven conditions are given for the first item, with only the SC/RC and RC/RC conditions for the remainder. Most of the components of the other versions can be computed from these as described in Section 3. The complete set of items is available from the first author. The first seventeen used “whether” as the head of the sentential subject in the SS items, while the remaining eighteen used “that”.

1. SC/RC: The chance that the administrator who the nurse supervised had lost the medical reports bothered the intern a great deal.  
RC/SC: The intern who the chance that the administrator had lost the medical reports had bothered a great deal supervised the nurse.  
SS/RC: Whether the administrator who the nurse supervised had lost the medical reports bothered the intern a great deal.  
RC/SS: The intern who whether the administrator had lost the medical reports had bothered a great deal supervised the nurse.  
RC/RC: The administrator who the intern who the nurse supervised had bothered a great deal lost the medical reports.  
RC/RC-subject-extraction: The administrator who the intern who supervised the nurse had bothered a great deal lost the medical reports.  
RC/RC-post-verbal: The medical reports were lost by the administrator who the intern who the nurse supervised had bothered a great deal.
2. SC/RC: The discovery that the babysitter who the parents were fond of had spanked the child disturbed the neighbors to some degree.  
RC/RC: The babysitter who the neighbors who the parents were fond of had disturbed to some degree spanked the child.
3. SC/RC: The news that the lawyer who the district attorney influenced would defend the murderer affected the celebrity case quite strongly.  
RC/RC: The lawyer who the celebrity case which the district attorney influenced had affected quite strongly would defend the murderer.
4. SC/RC: The suggestion that the preacher who the voters adored will give a moving speech might unnerve the politician a little bit.  
RC/RC: The preacher who the politician who the voters adored might unnerve a little bit will give a moving speech.
5. SC/RC: The likelihood that the graduate student who the professor respected had plagiarized the article upset the secretary quite a lot.  
RC/RC: The graduate student who the secretary who the professor respected had upset quite a lot plagiarized the article.
6. SC/RC: The disclosure that the union members who the news stories criticized had agreed upon a contract encouraged the executive tremendously.  
RC/RC: The union members who the executive who the news stories criticized had encouraged tremendously agreed upon a contract.
7. SC/RC: The possibility that the actor who the director yelled at would try to reorganize the scene maddened the producer beyond words.  
RC/RC: The actor who the producer who the director yelled at had maddened beyond words may try to reorganize the scene.
8. SC/RC: The message that the janitor who the doorman chatted with forgot to empty the trash infuriated the supervisor all night long.  
RC/RC: The janitor who the supervisor who the doorman chatted with had infuriated all night long forgot to empty the trash.
9. SC/RC: The notion that the bus driver who the policeman screamed at might hit the pedestrian scared the passenger out of his wits.  
RC/RC: The bus driver who the passenger who the policeman screamed at had scared out of his wits might hit the pedestrian.

10. SC/RC: The knowledge that the athlete who the doctor helped had scored the winning goal inspired the little boy to great achievements.  
RC/RC: The athlete who the little boy who the doctor helped had inspired to great achievements scored the winning goal.
11. SC/RC: The report that the fans who the hooligans ridiculed cheered for the home team excited the star player a huge amount.  
RC/RC: The fans who the star player who the hooligans ridiculed had excited a huge amount cheered for the home team.
12. SC/RC: The myth that the guard who the fair maiden married had killed the dragon impressed the knight immensely.  
RC/RC: The guard who the knight who the fair maiden married had impressed immensely killed the dragon.
13. SC/RC: The inference that the mindless lackey who the hero had fought with would bungle the job exasperated the evil genius tremendously.  
RC/RC: The mindless lackey who the evil genius who the hero had fought with had exasperated tremendously bungled the job.
14. SC/RC: The accusation that the cheerleader who the quarterback dated had snubbed the nerds aggravated the class president a great deal.  
RC/RC: The cheerleader who the class president who the quarterback dated had aggravated a great deal snubbed the nerds.
15. SC/RC: The implication that the dictator who the diplomat had insulted closed the embassy embarrassed the prime minister in parliament.  
RC/RC: The dictator who the prime minister who the diplomat insulted had embarrassed in parliament closed the embassy.
16. SC/RC: The threat that the beggar boy who the genie served would speak the magic words terrified the evil king immensely.  
RC/RC: The beggar boy who the evil king who the genie served had terrified immensely spoke the magic words.
17. SC/RC: The hunch that the serial killer who the waitress had trusted might hide the body frightened the FBI agent into action.  
RC/RC: The serial killer who the FBI agent who the waitress trusted had frightened into action hid the body.
18. SC/RC: The confession that the housewife who the whole town gossiped about bought no cosmetics distracted the Avon lady from her work.  
RC/RC: The housewife who the Avon lady who the whole town gossiped about had distracted from her work bought no cosmetics.
19. SC/RC: The fact that the bike messenger who the car just missed blocked traffic angered the motorcyclist just a little.  
RC/RC: The bike messenger who the motorcyclist who the car just missed had angered just a little blocked traffic.
20. SC/RC: The suspicion that the consultant who the shareholders mistrusted wrote a bad report alienated the management quite a lot.  
RC/RC: The consultant who the management who the shareholders mistrusted had alienated quite a lot wrote a bad report.
21. SC/RC: The awareness that the pilot who the gunner had warned broke radio silence alerted the navigator to the danger.  
RC/RC: The pilot who the navigator who the gunner had warned had alerted to the danger broke radio silence.
22. SC/RC: The revelation that the beatnik who everyone seemed to have met wore women's clothing fascinated the writer a huge amount.  
RC/RC: The beatnik who the writer who everyone seemed to have met had fascinated a huge amount wore women's clothing.
23. SC/RC: The observation that the tourist who the local children teased took dozens of pictures mystified the rug merchant to some degree.  
RC/RC: The tourist who the rug merchant who the local children teased had mystified to some degree took dozens of pictures.
24. SC/RC: The danger that the terrorist who the assassin shot at planted a bomb provoked the spy very strongly.  
RC/RC: The terrorist who the spy who the assassin shot at had provoked very strongly planted a bomb.

25. SC/RC: The story that the small town girl who strangers never talked to would eat live frogs intrigued the mysterious visitor intently.  
RC/RC: The small town girl who the mysterious visitor who strangers never talked to had intrigued intently ate live frogs.
26. SC/RC: The idea that the drummer who groupies loved spat on the crowd had disgusted the singer beyond words.  
RC/RC: The drummer who the singer who groupies loved had disgusted beyond words spat on the crowd.
27. SC/RC: The disclosure that the medic who the doctor yelled to dropped the bandage startled the victim quite a lot.  
RC/RC: The medic who the victim who the doctor yelled to had startled quite a lot dropped the bandage.
28. SC/RC: The information that the frightened child who the rescue worker looked for survived the crash comforted the old woman greatly.  
RC/RC: The frightened child who the old woman who the rescue worker looked for had comforted greatly survived the crash.
29. SC/RC: The thought that the leading man who the starlet adored forgot the crucial line entertained the comedian immensely.  
RC/RC: The leading man who the comedian who the starlet adored had entertained immensely forgot the crucial line.
30. SC/RC: The notification that the violinist who the corporate sponsors flattered skipped the rehearsal disappointed the conductor a great deal.  
RC/RC: The violinist who the conductor who the corporate sponsors flattered had disappointed a great deal skipped the rehearsal.
31. SC/RC: The reminder that the student who the professor met with disrupted the class tormented the teaching assistant a large amount.  
RC/RC: The student who the teaching assistant who the professor met with had tormented a large amount disrupted the class.
32. SC/RC: The rumor that the suspected mobster who the media attacked investigated the campaign contributions intimidated the ambitious attorney a little bit.  
RC/RC: The suspected mobster who the ambitious attorney who the media attacked had intimidated a little bit investigated the campaign contributions.
33. SC/RC: The realization that the burglar who the returning couple startled dropped the TV alarmed the paranoid neighbor to some degree.  
RC/RC: The burglar who the paranoid neighbor who the returning couple startled had alarmed to some degree dropped the TV.
34. SC/RC: The probability that the teenager who the architect hired had ruined the lawn irritated the bossy contractor at the end of the day.  
RC/RC: The teenager who the bossy contractor who the architect hired had irritated at the end of the day ruined the lawn.
35. SC/RC: The testimony that the child who the psychologist talked to had lost a lot of sleep worried the divorcing parents tremendously.  
RC/RC: The child who the divorcing parents who the psychologist talked to had worried tremendously had lost a lot of sleep.

## B Experimental Items for Experiment 2

The items used in Experiment 2 are listed below. All ten conditions are given for the first item, with only the SC/SC, pre-verbal condition given for the remainder, from which the other conditions can be computed as discussed in Section 5.

1. RC/RC-object-extraction, pre-verbal: The colonial settlers who the tribe which the sun god was protecting had frightened were not taken seriously by the government.  
RC/RC-object-extraction, post-verbal: The government did not take seriously the colonial settlers who the tribe which the sun god was protecting had frightened.  
RC/RC-subject-extraction, pre-verbal: The colonial settlers who the sun god which was protecting the tribe had

frightened were not taken seriously by the government.

RC/RC-subject-extraction, post-verbal: The government did not take seriously the colonial settlers who the sun god which was protecting the tribe had frightened.

RC/VP-gerund, pre-verbal: The colonial settlers who the sun god protecting the tribe had frightened were not taken seriously by the government.

RC/SC, pre-verbal: The colonial settlers who the myth that the sun god was protecting the tribe had frightened were not taken seriously by the government.

RC/SC, post-verbal: The government did not take seriously the colonial settlers who the myth that the sun god was protecting the tribe had frightened.

SC/SC, pre-verbal: The notion that the myth that the sun god was protecting the tribe might frighten the colonial settlers was not taken seriously by the government.

SC/SC, post-verbal: The government did not take seriously the notion that the myth that the sun god was protecting the tribe might frighten the colonial settlers.

SS/SC: That the myth that the sun god was protecting the tribe might frighten the colonial settlers was not taken seriously by the government.

2. The impression that the comment that the star athlete was receiving bribes would upset the coach was supported by the university official.
3. The danger that the pronouncement that locusts were infesting the farmlands might irritate the king was feared by the peasants.
4. The prediction that the story that the preacher was sleeping with the actress would scandalize the parishioners was overlooked by the newsman.
5. The information that the evidence that the aliens were kidnapping the children would expose the invasion plan was analyzed by the conspiracy buff.
6. The suggestion that the announcement that the corporation was firing the workers might infuriate the labor union was supported by the news broadcast.
7. The possibility that the information that the actress was seducing the director might aggravate the lead actor was ignored by the producer.
8. The implication that the belief that the artist was creating a new art form might influence the review was rejected by the scholarly journal.
9. The speculation that the feeling that the diplomat was mishandling the negotiations might irritate the foreign minister was denounced by the president.
10. The allegation that the proof that the stock trader was defrauding the investors might implicate the overseas bank was investigated by the regulators.
11. The report that the revelation that the landlord was faking the receipts would provoke the attorney was supported by the tenant's union.
12. The threat that the news that the terrorists were kidnapping the delegate might trick the bumbling law enforcement authorities was ignored by the negotiators.
13. The prospect that the fact that the sultry singer was singing the passionate love song might impress the mysterious stranger was feared by the jealous club owner.
14. The chance that the speculation that the author was writing the disturbing children's story might frighten off the veteran illustrator was discounted by the nervous publisher.
15. The implication that the story that the surrealist painter was offending the conservative politician might disturb the rich patrons was rejected by the gallery owner.
16. The rumor that the disclosure that the marketing whiz was masterminding the new ad campaign might boost the profits was analyzed by the new CEO.
17. The likelihood that the announcement that the sleazy pop-singer was endorsing the lipstick might offend the right-wing consumer group was discounted by the cosmetics mogul.
18. The testimony that the threat that the mobster was intimidating the reporters would silence the witness was questioned by the district attorney.
19. The likelihood that the evidence that the police officer was beating the suspect would outrage the journalists was disregarded by the lawyer.
20. The prediction that the certainty that the killer virus was decimating the small village might panic the capital city was ignored by the health officials.



21. The suggestion that the allegations that the professor was favoring the graduate student might disappoint the review committee irritated the department head.
22. The realization that the announcement that the flight attendants were serving drinks might calm the passengers was reassuring to the pilots of the damaged airliner.
23. The expectation that the reports that the teachers were neglecting the students might anger the school board troubled the superintendent.
24. The acknowledgement that the news that the waitress was working the graveyard shift might impress the assistant manager annoyed the cook.
25. The possibility that the belief that the hometown team was playing in the championship might excite the sick child uplifted the city residents.
26. The disclosure that the prospect that the Nazis might be seeking the lost temple had scared the archeologist embarrassed the university.
27. The prospect that the report that the bear was bothering the tourists would scare away the rich Americans annoyed the park ranger.
28. The chance that the rumor that the strange hermit was worshiping the nature spirit might intrigue the local kids troubled the villagers.
29. The misconception that the chance that the social worker was reporting the gun might scare away the emotionally disturbed teenager troubled the foster parents.
30. The confession that the fact that the noble huntsman was kissing the sleeping princess might lift the curse upset the evil witch-queen.
31. The implication that the expectation that the beautiful assistant was releasing the deadly cobra would distract the exotic conjuror fascinated the enraptured audience.
32. The feeling that the threat that the disgruntled client was stalking the lawyer might frighten the staff bothered the policeman.
33. The acknowledgement that the warning that the white supremacists were training an army might motivate the governor into action reassured the legislators.
34. The perception that the allegation that the lobbyist was bribing the politicians would offend the ethics committee reassured the constituents.
35. The suggestion that the impression that the animal trainer was abusing the elephants might unnerve the visitors alarmed the circus owner.
36. The likelihood that the idea that the dissident was publishing a newsletter would anger the dictator terrified the rebels.
37. The realization that the perception that the child prodigy was solving the difficult problem might embarrass the older boy frustrated the guidance counselor.
38. The realization that the promise that the money was helping poor people might inspire the priest consoled the grieving widow.
39. The hunch that the news that the tornado was destroying the small town might panic the touchy investors bothered the mayor.
40. The possibility that the suspicion that the secret agent was locking the cabin door might hinder the hired assassin infuriated the criminal mastermind.