Journal of Psycholinguistic Research, Vol. 25, No. 1, 1996

Interference in Short-Term Memory: The Magical Number Two (or Three) in Sentence Processing

Richard L. Lewis[1,2]

Many theories have been proposed to explain difficulty with center embedded constructions, most attributing the problem to some kind of limited-capacity short-term memory. However, these theories have developed for the most part independently of more traditional memory research, which has focused on uncovering general principles such as chunking and interference. This article attempts to gain some unification with this research by suggesting that an interesting range of core sentence processing phenomena can be explained as interference effects in a sharply limited syntactic working memory. These include difficult and acceptable embeddings, as well as certain limitations on ambiguity resolution, length effects in garden path structures, and the requirement for locality in syntactic structure. The theory takes the form of an architecture for parsing that can index no more than two constituents under the same syntactic relation. A limitation of two or three items shows up in a variety of other verbal short-term memory tasks as well.

INTRODUCTION

An important goal in psycholinguistics is understanding the nature of short-term memory (or working memory) and its effects on comprehension. The classic short-term memory (STM) effect in sentence processing is severe

Preparation of this paper was supported by a grant from the James S. McDonnell Foundation to the Human Information Processing Group at Princeton University. Many thanks to Martin Chodorow, Terry Langendoen, Thad Polk, and an anonymous reviewer for helpful comments on the paper and research.

[1] The Ohio State University, Columbus, Ohio.

[2] Address all correspondence to Richard L. Lewis, Department of Computer and Information Science, The Ohio State University, 2015 Neil Avenue, Columbus, Ohio 43210, or rick@cis.ohio-state.edu.

0090-6905/96/0100-0093$09.50/0

difficulty on multiple center embedding (Miller, 1962; Miller & Chomsky, 1963). Over the past three decades, many theories of syntactic processing have been proposed to explain this effect, most attributing the problem to some kind of limited-capacity STM. However, these theories have developed for the most part independently of more traditional memory research, which has focused on uncovering general principles such as chunking (Miller, 1956; Simon, 1974) and interference (McGeoch & McDonald, 1931).

One principle of memory that has been neglected in theories of sentence processing is similarity-based interference. When to-be-remembered items are followed by stimuli that are similar along some dimensions, the original items are more quickly forgotten (Shulman, 1970; Waugh & Norman, 1965). This article attempts to gain some unification with traditional memory research by suggesting that an interesting range of core sentence processing phenomena can be explained as interference effects in a sharply limited syntactic working memory. This research follows the path Newell (1990) urged on psychology—attempting to explain a range of empirical phenomena with a small set of independently motivated architectural mechanisms and principles and their associated constants (Simon, 1974).

The article begins with a brief overview of theories of center embedding and their numeric bounds on processing. Next we shift to traditional accounts of short-term memory, focusing in particular on interference effects and the variety of codes used in short-term memory. The next few sections show how an interference theory of syntactic short-term memory can explain difficult and acceptable embeddings, some limitations on ambiguity resolution, length effects in garden path structures, and the general requirement for locality in syntactic structure. Finally, we speculate on some possible bases for the limitation of two or three, and discuss some empirical problems with the present model and its bound of two.

CENTER EMBEDDING AND BOUNDS ON PROCESSING

The difficulty in comprehending multiple center embedded object relative clauses in English (Miller & Chomsky, 1963) is one of the best known psycholinguistic phenomena.

(1) The salmon that the man that the dog chased smoked fell.

Center embedding refers to what Miller and Chomsky (1963) called nesting of dependencies, which occurs when a constituent *X* is embedded in another constituent *Y*, with material in *Y* to both the left and right of *X*. *Self-embedding* is a special case of center embedding involving constituents of the same category. Self-embedding is the formal property that moves a grammar outside the scope of a finite device, but as Miller and Chomsky pointed out, center embedding in general increases memory load. The classic examples such as (1) exhibit double self-embedding (hence, center embedding) of both noun phrases (NPs) and sentences (Ss):

(2) a. [S The salmon that [S the man that [S the dog chased] smoked] fell.]
b. [NP The salmon that [NP the man that [NP the dog] chased] smoked.]

The difficulty of these constructions has been confirmed in a variety of experimental paradigms (Blauberg & Braine, 1974; Blumenthal, 1966; Hakes & Cairns, 1970; Marks, 1968; Miller & Isard, 1964; Wang, 1970). The basic result is that center embedding is always more difficult than right-branching, which can carry the same amount of content but exhibits no severe limits on depth of embedding. Although this fundamental contrast has been known for nearly 40 years now, it was not until Cowper (1976) and Gibson (1991) that the extent of the relevant cross-linguistic data began to be revealed.

Most theories that account for difficulty with center embeddings fall into three broad classes: linguistic metrics, architectures for parsing, and perceptual strategies [for a review, see Lewis (1995)]. *Metrics* define a measure over syntactic structure which predicts the perceived relative complexity of parsing different structures. The most sophisticated and empirically successful metric is Gibson's (1991; Gibson, Schütze, & Salomon, 1996). *Architectures* specify functional computational mechanisms for parsing, and characterize capacity limitations in terms of some fixed computational resource. PARSIFAL (Marcus, 1980) is one familiar example. The third and smallest class of theories, *perceptual strategies*, posits that difficulty arises from interactions among a set of perceptual mapping strategies that map surface strings to some underlying syntactic structure (Bever, 1970; Kac, 1981).

Metrics and architectures require specific constant limits if they are to make absolute predictions of acceptability. Table I lists many of the theories of center embedding that have posited some specific bound. Despite the empirical successes represented in this table, only Yngve (1960) used an independently estimated constant [Miller's (1956) 7 +/− 2]. The problem runs much deeper than just stipulating constants, however; there is often no psychological motivation for either the posited structure of short-term memory (stacks, lookahead buffers, etc.), or how the structure should be limited.

There have been attempts to unify linguistic processing with independently motivated psychological constructs. For example, Gathercole and Baddeley (1993) discussed the linguistic functions of the phonological loop,

Table I. Some Posited Processing Bounds in Sentence Processing

Metric or Architecture	Limited resource	Limit
Push-down automaton (Yngve, 1960)	Stack cells	7 +/− 2
Sausage Machine (Frazier & Fodor, 1978)	Lookahead word window	6
Open syntactic requirements (Gibson, 1991)		4
Register-vector finite state machine (Blank, 1989)	State for tracking clause level	3
PARSIFAL (Marcus, 1980)	Lookahead constituent buffer	3
Poker parser (Cowper, 1976)	Clausal processing track, hold cells	3
Connectionist net (Henderson, 1994)	Public stack; tree fragments	2; 4
Open sentence nodes (Kimball, 1973)		2
Open case relations (Stabler, 1994)		1 or 2
ACT (Anderson, Klein, & Lewis, 1977)	Control variables	1
Subroutine architecture (Miller & Isard, 1964)	Return address memory	1

which has been used to account for a variety of immediate memory tasks (Baddeley, 1990; Simon & Zhang, 1985). Just and Carpenter (1992) proposed an activation-based, limited-capacity working memory that has been used to model individual differences in a range of cognitive tasks (Carpenter, Just, & Shell, 1990). Bever (1970) represented one of the earliest efforts to find a general cognitive and perceptual basis for language processing. But in the domain of difficult and acceptable embeddings, none of these theories have come close to achieving the cross-linguistic coverage of the best syntactic metrics and architectures, in particular, those of Cowper (1976) and Gibson (1991).

INTERFERENCE IN A MULTIFACETED WORKING MEMORY

In this section, I will briefly attempt to establish two things. First, working memory uses a multiplicity of different codes, and exhibits interference effects when items are stored using similar codes. Second, the standard verbal STM tasks do not use the same codes that are functionally required by syntactic parsing. These two claims lead naturally to positing a portion of working memory that is coded specifically in terms of syntactic features, but which still exhibits interference effects.

Retroactive Interference and Within-Category Similarity Effects

Table II lists a set of content categories for which there is some evidence of interference effects in short-term memory, independent of decay. The table refers primarily to studies demonstrating *retroactive interference*, which occurs when a to-be-remembered stimulus is followed by a set of distractor items that the subject must attend to in some fashion.[3] The general finding is that distractor items that are similar in kind to the original stimulus cause forgetting, while dissimilar items are far less disruptive. For example, immediate memory for odor is disrupted by interpolated tasks involving odors, but is unaffected by a distracting verbal task such as counting backward (Walk & Johns, 1984).

Most of this evidence is consistent with a model in which similarity has an effect only at a fairly gross level (visual vs. phonological, odor vs. verbal), and in which interference can be construed primarily as a matter of displacement (Waugh & Norman, 1965). However, there are *within-category* similarity effects which provide stronger support for the concept of similarity-based interference and demonstrate that the mechanisms of forgetting in working memory are probably more subtle and complex than displacement at a gross level.

The most familiar within-category effect is the phonological similarity effect. Ordered recall of phonologically similar lists of words, consonants, or nonsense trigrams is worse than recall of dissimilar lists (Baddeley, 1966; Conrad, 1963; Wickelgren, 1965). A similar effect was found with immediate memory for signs in American Sign Language, where the relevant

[3] Although proactive interference (PI) may occur in short-term memory as well, PI as it has been studied may more properly be considered a phenomenon of long-term memory (Baddeley, 1990; Shulman, 1970).

Table II. Kinds of Immediate Memory Subject to Interference

Conceptual/semantic	(Potter, 1976)
Kinesthetic	(Williams, Beaver, Spencer, & Rundell, 1969)
Odor	(Walk & Johns, 1984)
Sign language	(Poizner, Bellugi, & Tweney, 1981)
Tone	(Deutsch, 1970)
Verbal/phonological	(Schiffrin, 1973; Waugh & Norman, 1965)
Visual	(Logie, Zucco, & Baddeley, 1990)

measure of similarity was formational (Poizner, Bellugi, & Tweney, 1981). Similarity effects have also been demonstrated in visual memory tasks. For example, distractor gratings with similar orientations to the target stimuli produced more interference than those with dissimilar orientations (Magnussen, Greenslee, Asplund, & Dyrnes, 1991).

Traditional Verbal STM in Syntactic Processing

The suggestion that theories of center embedding have largely ignored traditional work on STM raises the important question: Does verbal short-term memory as studied in these classic experiments actually have anything to do with syntactic processing?

The answer may be: yes, but very little. The reason is fairly straightforward when one considers the functional requirements of parsing and remembering lists of unrelated stimuli. There is good evidence that traditional verbal memory tasks rely on phonological codes in working memory (Baddeley, 1990). Parsing, on the other hand, requires the temporary storage of partial syntactic structures. With the possible exception of linear precedence information, none of the functional models listed in Table I incorporate representations that resemble the content of phonological STM.

There is also empirical evidence for the independence of parsing from phonological STM. Early studies with center embedding clearly demonstrated conditions where phonological STM is intact, but parsing is impaired. For example, Larkin and Burns (1977) showed that subjects could repeat verbatim the words in a difficult center embedded sentence without being able to correctly parse the structure. Conversely, some brain-damaged subjects with severely impaired phonological short-term memories were still able to parse even complex constructions (e.g., Martin, 1993). Potter's (1982) RSVP (rapid serial visual presentation) experiments provided a condition where parsing is intact but phonological STM is impaired in normals.

The functional considerations, the evidence of a double dissociation between parsing and phonological STM, and the empirical success of purely syntactic theories suggest that most of the action in explaining difficult and acceptable embeddings is to be found in that part of working memory based on syntactic features, not phonological features.

AN INTERFERENCE THEORY OF SYNTACTIC WORKING MEMORY

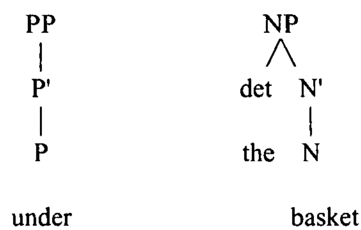
The STM theory presented here is part of a more comprehensive computational model of real-time sentence comprehension, NL-Soar (Lewis, 1993; Lewis & Lehman, 1994). NL-Soar is built within Soar, a theory of the human cognitive architecture (Laird, Newell, & Rosenbloom, 1987; Newell, 1990; Rosenbloom, Lehman, & Laird, 1993). In NL-Soar, Soar's working memory is used to store the partial products of incremental comprehension, including syntactic structure. The model of syntactic working memory described below grew out of an attempt to ensure that the most primitive cognitive processes in Soar operate efficiently (matching associations against working memory) (Lewis, 1993; Tambe, Newell, & Rosenbloom, 1990).

The part of NL-Soar's working memory that buffers partial constituents is called the H/D set, for Heads and Dependents. Assuming that the function of syntactic parsing is to build surface structure, X-bar positions (specificier of IP (spec-IP), complement of V' (comp-V'), adjunct of N' (adjoin-N'), etc.) (Chomsky, 1986) comprise one possible useful set of indices for partial structures. The Heads part of the H/D set indexes nodes by the structural positions they may head. The Dependents part of the H/D set indexes nodes by the structural relations they may fill. For example, (3a) gives the partial H/D set corresponding to the situation in (3b):

(3) a.

Heads	comp-P':	[P' under]
Dependents	comp-P':	[NP the basket]
	comp-V':	[NP the basket]
	adjoin-N':	[PP under]

b.



Parsing is now a process of projecting X0 nodes (Pritchett, 1991), and matching potential heads with potential dependents. In (3), [P' under] and [NP the basket] may be joined by the structural relation comp-P', since [NP the basket] is indexed under comp-P' in the Dependents set, and [P' under] is indexed under comp-P' in the Heads set. Furthermore, [PP under] may potentially be adjoined to an N' as a postnominal modifier, so it is indexed under adjoin-N' in the Dependents set, and so on.

Once the PP *under the basket* has been formed, the H/D set contains the following:

(4)	Heads	comp-P':	[P' under [NP the basket]]
	Dependents	adjoin-N':	[PP under [NP the basket]]

When a node is assigned a structural position, it is removed from the Dependents set. However, because the H/D set provides all access to syntactic structure (i.e., it is not shunted off to another buffer for further processing), *under* must remain in the Heads set to support semantic interpretation.

In the example above, each relation indexes just one node, but in principle multiple constituents may be indexed by the same relation. Consider the case of a single object relative clause:

(5) a. The cat that the bird found jumped.

b.

Dependents	spec-IP:	[NP the cat], [NP the bird]

In (5), both NPs must be momentarily indexed under spec-IP in the Dependents set, since both eventually occupy a subject position.

We can make this model an interference theory by simply placing a limit on the number of nodes that each relation may index. In this way, structures are similar and interfere with each other when they must be indexed by the same relations. How much should each relation index? If we assume a uniform limit across all relations in the H/D set, that limit must be at least two in order to achieve the minimal functionality required to parse natural language (Lewis, 1995). The reason is simple: To compose propositions (as in a relative or complement clause), two clauses must often be held simultaneously in working memory. Furthermore, clauses by their nature have similar structures and will often interfere with each other. We will return to a discussion of this memory bound later in the paper, but for now simply assume it to test its empirical adequacy.

ACCOUNTING FOR SOME LIMITATIONS IN HUMAN PARSING

This section describes the model's explanation of a variety of cross-linguistic phenomena in parsing. The first few subsections deal with difficult and acceptable embeddings, presenting a subset of the 50+ constructions analyzed in Lewis (1995). The remaining subsections consider ambiguity resolution, garden path effects, and grammatical locality constraints.

Embedded Relative Clauses

Consider again the difficult center embedding:

(6) The boy that the man that the woman hired hated cried.

The Dependents set must index the three initial NPs under spec-IP, since all three NPs will eventually occupy subject position:

(7)	Dependents	spec-IP:	[NP the boy], [NP the man], [NP the woman]

Failure will occur at one of the final verbs (which verb depends on which NP is dropped from the H/D set). It is important to note that failure does not occur just because nodes are removed from the set. Just as in standard short-term memory tasks, interference is only a problem if the item that is interfered with must be accessed later.

The uniform limit of two nodes per relation in the H/D set theory does not entail a prohibition against all double center embedding. If that was the case, it would essentially be equivalent to Kimball's (1973) principle of two sentences, and fall prey to the same empirical problems. Consider the pseudocleft construction, which has something like a *wh*-clause in subject position:

(8) a. What the man saw was a dog.
b. [IP [NP [CP What the man saw]] was a dog.]

The initial *wh*-clause in pseudo-clefts is an indirect question (Baker, 1989; McCawley, 1988). The interaction of this structure with the H/D set leads to some interesting predictions. Because the initial *wh*-word does not occupy spec-IP, it should be possible to embed an additional relative clause within the indirect question without causing difficulty. This prediction turns out to be true (Gibson, 1991)[4]:

[4] The fact that *wh*-phrases do not occupy spec-IP means that the following construction is incorrectly predicted to be acceptable (as pointed out by an anonymous reviewer):

(33) Which bird that the mouse that the cat chased scared flew away?

In (33), *which bird* need not be indexed under spec-IP in the Dependents set for the parse to go through. Furthermore, under the assumption that the matrix CP is not projected until *flew*, there will not be interference in the Heads set on spec-CP. However, this strict bottom-up assumption should perhaps be overturned in favor of a parsing strategy in which specifiers trigger creation of their heads (e.g, Gibson, 1991; Inoue & Fodor, 1995). Adopting such a parsing strategy has the following implications. First, structures such as (33) are correctly predicted to be difficult, due to interference on spec-CP in the Heads set:

(34)	Heads	spec-CP:	[CP which bird], [CP that...], [CP that...]

Furthermore, center embedded subject relatives such as (10) and (11) are now predicted to be difficult. Finally, this parsing strategy should make NL-Soar more consistent with phenomena in head-final languages which may indicate that humans do not wait for the end of the clause to begin projecting clausal structure (Inoue & Fodor, 1995). Fortunately, although the details of some analyses may change, the other predictions of the theory remain the same.

(9) a. What the woman that John married likes is smoked salmon.

[s What [s the woman that [s John married] likes] is smoked salmon.]

b.	Dependents	spec-IP:	[NP woman], [NP John]
		spec-CP:	[NP what]

One robust finding in psycholinguistics is that English object relatives are more difficult to process than subject relatives (Ford, 1983; Holmes & O'Regan, 1981; Hudgins & Cullinan, 1978). Cowper (1976) and Kac (1981) noted that this is true of double embeddings as well: Double-embedded subject relatives such as (10) below are easier to process than object relatives:

(10) a. The book that the man who hired me wrote deals with politics. (Cowper, 1976)
b. [s The book that [s the man who [s hired me] wrote] deals with politics.]

In (10), only two NPs (*the book* and *the man*) must be indexed simultaneously. Although there is a trace NP in subject position of the IP headed by *hired*, the generation and attachment of traces is triggered by their assigners, and therefore they do not need to be temporarily buffered in the Dependents set.

Unlike (6), however, (10) does not exhibit double self-embedding of complex NPs. Subject relatives which do exhibit double NP self-embedding seem worse:

(11) a. The boy that the man that hired the woman yesterday hated cried.
b. [NP The boy that [NP the man that hired [NP the woman] yesterday] hated.]

In fact, Gibson and Thomas (1995) presented evidence that center embedded subject relatives can be unacceptably difficult. Since NL-Soar predicts that these should be acceptable (but see footnote 4), there remains a problem in capturing the contrast between easy and difficult subject relatives. No model currently accounts for this contrast, and it is not clear if self-embedding of NPs is the critical difference.

There are also some triple-subject constructions that the model predicts should be difficult, yet seem easier to process than the classic (6):

(12) a. The reporter everyone I met trusts had predicted that Thieu would resign in the spring. (T. Bever, personal communication, August 4, 1995)

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

(13) The claim that the man that John hired is incompetent upset me. (Gibson, 1991)

Sentences like those in (12) are counterexamples to the claim that double center embedded relative clauses are always difficult in English. Example (13) also involves three subjects, but mixes a complement clause with a relative clause. Interestingly, the constructions in (12) and (13) fit the general prediction of the interference hypothesis that making the constituents more distinct or dissimilar in some way should help processing, as first noted by Bever (1970). The fact that they are nevertheless predicted to be difficult by the H/D set suggests that something may be missing in the way that the H/D set has formalized similarity.

Is the Problem Center Embedding or Gap-filling?

Difficulty can arise with center embedding independently of problems computing filler-gap relations, consistent with Miller and Chomsky's (1963) initial analysis. For example, Mazuka et al. (1989) found that double center embedding sentential complements and adjuncts in Japanese can cause difficulty, even though no filler–gap relations are involved:

(14) Akira-ga Tosiko-ga Hazime-ga nakidasita toki okidasita no ni kizuita.
Akira NOM Tosiko NOM Hazime NOM started crying when got-up that noticed.
(Akira noticed that Toshiko got up when Hajime started crying.)

In (14), each of the NPs occupies subject position, requiring the H/D set to index three NPs on the spec-IP relation:

(15)	Dependents	spec-IP:	[NP Akira], [NP Tosiko], [NP Hazime]

Interestingly, the H/D set does predict that computing filler–gap relations can cause problems as well. Consider *tough-movement* constructions:

(16) Fred$_i$ is tough [CP e$_i$ [C' for the man to please e$_i$]].

In (16), the matrix subject is related to the object of the embedded clause via movement through spec-CP. Embedding two relatives in such a construction is unacceptably difficult (Cowper, 1976):

(17) Fred is easy for the woman who the man who hired me married to please.

Stacking three subject NPs cannot be the explanation of difficulty in (17), because at most two NPs must be buffered simultaneously under spec-IP (*woman* and *man*). Nevertheless, the H/D set cannot parse this construction because three CPs must be indexed simultaneously in the Heads set under spec-CP so that the traces can be properly generated:

(18)	Heads	spec-CP:	[CP for], [CP who], [CP who]

Is the Problem Stacking or Embedding?

The H/D set explanation for difficulty with center embedding places the blame on stacking (consecutive occurrences) of NPs, rather than center embedding per se. An interesting study by Hakuta (1981) was able to tease apart these two factors. Controlling for depth of embedding, he found that Japanese children had much less trouble with center embedded structures which were transformed so that the amount of stacking was reduced. Mazuka et al. (1989) also pointed out that difficult double center embeddings such as (14) become easy to process if one or two of the initial overt NPs are dropped. The H/D set predicts this because then only one or two NPs must be indexed under spec-IP.

Although the Hakuta (1981) study showed that stacking can be a source of difficulty in Japanese, it need not be. In fact, stacking NPs appears to be more acceptable in head-final languages such as Japanese. Remarkably, Japanese sentences that stack up to five initial NPs are acceptable for some speakers:

(19) John-wa Bill-ni Mary ga Sue-ni Bob-o syookai sita to it-ta.
John TOP Bill DAT Mary NOM Sue DAT Bob ACC introduced COMP say PERF.
(John said to Bill that Mary introduced Bob to Sue.)

Although such sentences are surely complex, they do not cause the failure associated with (6) or (14). The H/D set can handle structures such as (19) because no single structural relation must buffer more than two NPs[5]:

(20)	Dependents	spec IP:	[NP John], [NP Mary]
		comp-V':	[NP Bob]
		comp2-V':	[NP Bill], [NP Sue]

[5] The overt case marking does not in and of itself explain the contrast between Japanese and English stacked NPs. Note that the explicit case marking in (14)—with only three stacked NPs—does not seem to help.

Limitations on Ambiguity Resolution

The H/D set has implications for ambiguity resolution as well. In particular, it makes the strong prediction that at most two nodes are available to assign the same structural relation at any given time. For example, consider the right branching structure in (21):

(21) a. Amparo saw the dog under the box on the table in the room next to the library.

b.

Heads	adjoin-N':	[N' dog], [N' library]

At any given point, at most two noun phrases are available for postnominal modification, because only two NPs may be indexed under the adjoin-N' relation in the Heads set. [In fact, (21) can be successfully parsed by just keeping the most recent NP open for attachment.]
In this way, the H/D set naturally serves a function similar to *closure principles* (Church, 1980; Frazier & Rayner, 1982; Kimball, 1973), which keep a small set of syntactic nodes open for further attachment. The H/D set bears the closest resemblance to Church's (1980) closure principle, which posits that only the two most recent nodes of the same category may be kept open at any time. Church's principle and those derived from it have the best empirical coverage (Church, 1980; Gibson, 1991). The main differences between NL-Soar and Church's principle is that closure in NL-Soar derives from the structure of the parsing architecture, and the limit on open nodes is on the basis of similar structural relations, not categories.

One way of testing the theory is to construct material with three potential sites and syntactically *force* attachment to each of the three sites as the experimental manipulation. The H/D set predicts that, at least some of the time, one of the sites should cause difficulty. Gibson, Pearlmutter, Canseco-Gonzales, and Hickock (1993) conducted a study in English and Spanish using material with three potential NP attachment sites, and found that forcing PP attachment to one of the sites (the intermediate site) often did cause difficulty. Gordon (1982) uncovered a similar pattern in adverb attachment. The H/D set does not predict which of the three sites will be difficult, only that one must be. Predicting the site of difficulty requires further specifying exactly how conflicts are resolved when more than two nodes are vying for the same relation.

A general recency preference has played an important role in several sentence processing models (e.g., Gibson, 1991; Kimball, 1973). In fact, by abstracting away from the effects of any particular strategies for handling conflicts in the H/D set, we can see that the basic structure of the H/D set does give rise to a kind of recency preference. More precisely:

(22) *H/D set recency preference*: Given a sequence of syntactic nodes $x_1, x_2, \ldots, x_n$, $n > 2$, that potentially assign some structural relation ρ, attachment to more recent nodes via ρ is more likely than attachment to less recent nodes, all things being equal.

"All things being equal" means that when three nodes are vying for the same position, all three nodes are equally likely candidates for dropping from working memory. Preference (22) can be derived with a simple probabilistic analysis [see Lewis (1993) for details]. The basic insight is that earlier nodes will have had more opportunities to be displaced by incoming material. This is essentially the interference or displacement account of recency, as opposed to a decay-based account (Kempen & Vosse, 1989; Stevenson, 1994).

This result only holds for nodes that are competing for the same structural relation. For example, verbs do not compete with nouns for PP attachment, nor do complement attachments compete with adjuncts. This is consistent with the fact that right association is not a good predictor across syntactic categories, or between argument/adjunct ambiguities (Abney, 1989; Whittemore & Ferrara, 1990).[6] Note that (22) is not at odds with the Gibson et al. (1993) and Gordon (1982) results cited above, which clearly indicate that something in addition to recency is at work. Preference (22) only reveals the central tendency of the model. There may be specialized strategies which violate strict recency in many cases.

Length Effects in Garden Path Structures

Consider the following easy and difficult structural ambiguities:

(23) John believed the professor was honest.
(24) The horse raced past the barn fell. (Bever, 1970)

One striking difference between many unproblematic ambiguities like (23) and the classic (24) is the distance between the ambiguity and the disambiguating material. In (23), the disambiguating information follows immediately on the heels of the ambiguity, while in the classic garden path, there is an intervening phrase (*past the barn*) between the ambiguous raced and the disambiguating *fell*. Such a difference leads naturally to the hypothesis that distance-to-disambiguation is a crucial factor.

---

[6] Of course, there are right association effects that are not accounted for by the H/D set, because the limitations do not show up until more than two competing nodes have been encountered. It is still unclear whether this indicates an independent recency principle is at work which should be incorporated into the model, as in Gibson (1991).

However, the distance-to-disambiguation can be extended in structures like (23) without causing severe garden path effects (Pritchett, 1992):

(25) a. Ron believed the ugly little linguistics professor.
b. Ron believed the ugly little linguistics professor he had met the week before in Prague disliked him.

NL-Soar includes a limited repair mechanism for handling unproblematic ambiguities (Lewis, 1993). As long as the complement relation assigned by the relevant verb (in this case *believe*) is still accessible in working memory, then the structure can be repaired regardless of the length of the object noun phrase.

Nevertheless, Warner and Glass (1987) *did* manage to produce an apparently length-induced garden path effect, using exactly the same object/subject ambiguity:

(26) The girls believe the man who believes the very strong ugly boys struck the dog killed the cats.

Surprisingly, NL-Soar accounts for such garden path effects. The intervening material is such that all the structural relations required for a successful repair cannot be held in the working memory. For the repair to be successful, the H/D set must support the following structure upon reading *struck*[7]:

(27)	Heads	comp-V':	[v' believes], [v' believe], [v' struck]

This exceeds the capacity of two nodes per relation. Thus, the limited working memory produces a garden path effect because the relevant structural relations are not available to repair. The important factor is not simply the length of the intervening material, but the interaction of the structure of the intervening material with the structure to be repaired.[8]

[7] In (26), *the man* and *the boys* are taken as complements of *believe* and *believes*, respectively. When *struck* arrives, it must be indexed under the Comp-V' relation so that it may take its complement. *Believes* must also still be available on the Comp-V' relation, because the complement of *believes* must change from the *boys* to *struck*. However, *believe* must also be on the Comp-V' relation, because the complement of *believe* must be changed from *the man* to *the man killed.*

[8] Contrary to the theory presented here, Ferreira and Henderson (1991) presented evidence that the crucial factor is the length of the ambiguous region after the head of the object phrase, not the structural complexity. However, their material consisted of subject/object ambiguities that can cause difficulty even in short versions, in contrast to (23) above:

(35) When the men hunt the birds typically scatter.

Thus, Ferreira and Henderson were investigating the factors that cause a mild but noticeable garden path effect to become a severe one, while the theory above is concerned with the structural factors that can cause a perfectly acceptable ambiguity to become a noticeable garden path. In short, both length and structure may play a role.

Working Memory and Syntactic Locality

The H/D set suggests that some, but not all, extraction violations may be accounted for by interference effects in short-term memory. Consider the severe violation in (28):

(28) *Who$_i$ does Phineas know a boy who hates the man who saw t$_i$?

The H/D set cannot maintain the structures in working memory necessary to establish the long distance relation between the object of *saw* and the initial *who*. The structure required to establish this relation involves three CPs (complement phrases):

(29) [CP who does . . . [CP who hates . . . [CP who saw . . .]]]

Binding the object trace with the wh-word antecedent requires accessing the antecedent in spec-CP (specifier of CP) and following three comp-C' (complement of C') links (among others) to reach the object position in the final embedded clause. This exceeds the two-valued capacity of the H/D set:

(30)	Heads	spec-CP:	[CP who does]
		comp-C':	[C' who does], [C' who hates], [C' who saw]

In general, the H/D set will rule out extractions of the form

(31) [α . . . X . . . [α . . . [α . . . Y . . .]]]

which essentially corresponds to Ross's (1967) *wh*-island constraint when α is S. This is also enticingly close to subjacency (Chomsky, 1973), where the prohibition against crossing two bounding nodes corresponds to the limit of two in the H/D set. However, the H/D set clearly does not capture all the violations that subjacency captures, because α must be the same category in the H/D set, while subjacency predicts that violations occur when two bounding nodes are crossed, which may or may not be the same category (complex NP violations are an example, crossing S and NP bounding nodes.)

What the H/D set does predict is that eventually all long-distance dependencies will become unprocessable due to syntactic interference in short-term memory. Thus, there is clear motivation for some kind of locality constraint on grammatical representation. However, within the small range of processable extractions, there may still be other grammatical principles at work that rule out certain extractions.

THE BOUND OF TWO OR THREE IN OTHER STM PHENOMENA

None of the posited processing bounds in sentence processing theories (Table I) have had quite the magic of Miller's (1956) original 7 +/− 2. The limitations were only motivated by one kind of data, and until Gibson (1991), even the coverage of that one kind of data was not compelling. I hope now to make good on an implicit promise in the title by showing that the constant two or three shows up often enough in other guises that it should at least raise an eyebrow.

We have already seen that a bound of two or three may be behind a variety of syntactic processing effects. But in addition to this, this constant seems to arise in immediate verbal memory tasks not involving phonological short-term memory. For example, when presentation of words is so rapid that subjects do not have time to encode them phonologically, average word span drops to 2.6 (Potter, 1982). Rapid presentation is not necessary to reduce memory span. Simon and Zhang (1985) found that the average span for Chinese homophones (all pronounced /gong/, so a purely phonological memory would not be very useful) was 2.7. Table III summarizes these and other phenomena.

Table III may represent nothing more than some coincidences. However, it is possible to give these data a plausible interference-based interpretation. For any, given task, subjects must adopt some means of coding stimuli in working memory. Furthermore, the subject must essentially come to the task prepared with this set of codes; there is not time in a single task to learn a new encoding scheme. For the vast majority of verbal short-term memory tasks, subjects adopt a phonological encoding, which provides a ready set of over-learned features to discriminate the contents of memory. However, in tasks where a phonological code is not available or where phonological interference is too great, there may essentially be no means at hand to discriminate the to-be-remembered stimuli. This is the case with the tasks in Table III, by design. In short, what we could be seeing in the data

Table III. The Magical Number Two (or Three) in Verbal Short-Term Memory

Maximum subject NPs stacked in center-embedding (Lewis, 1993)	2 or 3
Phrases of a single category open for attachment (Church, 1980; Gibson, 1991; Gibson & Pearlmutter, 1994)	2 or 3
Bounding nodes crossed in Wh-movement for violation (Chomsky, 1973)	2
Average span for RSVP words (Potter, 1982)	2.6
Average span for unpronounceable Chinese radicals (Simon & Zhang, 1985)	3.0
Average span for Chinese homophones (Simon & Zhang, 1985)	2.7
Average reading span (Daneman & Carpenter, 1980)	3.15

in Table III is the limited capacity of human memory to keep in mind only two or three similar chunks at the same time. It is of particular interest that the average reading span of subjects in the Daneman and Carpenter (1980) task falls into this range, because this leaves open the exciting possibility of unifying the interference theory presented here with the Just and Carpenter (1992) account of individual differences in working memory capacity.

IS TWO ENOUGH?

The predictions of the H/D set depend on the bound of two and adopting the full set of X-bar positions. Although the empirical coverage is promising [see Lewis (1995) for the full data set], there are still problems.

The most significant problem, noted earlier, is that there are constructions that require buffering three subjects, but that are nonetheless acceptable. As one additional example, Babyonyshev and Gibson (1995) point out that triple-embedded complement clauses in Japanese are acceptable, despite the fact that the first three NPs occupy spec-IP (and therefore should overload the H/D set):

(32) Taroo-ga Akira-ga Hanako-ga nakidasita to itta to omotteiru.
Taroo Akira Hanako started-crying that said that knows.
(Taroo thinks that Akira said that Hanako started crying.)

This suggests that the bound should be three, and not two (as does some of the data in Table III, if we take it seriously). Making this move, however, would require finding additional sources of difficulty in the standard double center embedding cases. The Gibson et al. metric does this because it assigns additional cost to the empty operator positions associated with the relative clauses. However, this may not be the correct generalization, for we are still left with the fact that the double center embedded (14) is very difficult, but contains no relative clauses. At present, neither theory can account for the contrast between (32) and (14). It appears that there may be additional cost associated with adjuncts in general, regardless of whether they involve relativization.

CONCLUSION

This paper suggests that sentence processing theories can benefit from incorporating ideas from traditional work in short-term memory. NL-Soar demonstrates that it is possible to combine general principles of memory (in particular similarity-based interference) with linguistic analyses to produce a detailed model of syntactic working memory that is empirically powerful, and psychologically and functionally grounded.

NL-Soar's working memory captures two important theoretical convergences in recent accounts of syntactic short-term memory:

- The source of memory load is open or unsatisfied syntactic relations (Abney & Johnson, 1991; Gibson, 1991; Stabler, 1994). This leads naturally to a focus on stacking, rather than embedding per se, as the source of difficulty (Gibson, 1991; Hakuta, 1981; Mazuka et al., 1989).
- Increasing similarity makes things more difficult. This shows up clearly in the "broad but shallow" memory models of Reich (1969) and Stabler (1994), and the original self-embedding metric of Miller and Chomsky (1963).

Although the accumulated phenomena of short-term memory are vast, there is a great potential payoff for continued attempts to unify traditional STM phenomena with functional tasks such as language comprehension.

REFERENCES

Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research, 18*, 129–144.

Abney, S. P., & Johnson, M. (1991). Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research, 20*, 233–250.

Anderson, J. R., Kline, P. J., & Lewis, C. H. (1977). A production system model of language processing. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18*, 362–365.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Boston: Allyn and Bacon.

Baker, C. L. (1989). *English syntax*. Cambridge, MA: MIT Press.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

Blank, G. D. (1989). A finite and real-time processor for natural language. *Communications of the ACM, 32*, 1174–1189.

Blauberg, M. S., & Braine, M. D. S. (1974). Short-term memory limitations on decoding self-embedded sentences. *Journal of Experimental Psychology, 102*, 745–748.

Blumenthal, A. L. (1966). Observations with self-embedded sentences. *Psychonomic Science, 6*, 453–454.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*, 404–431.

Chomsky, N. (1973). Conditions on transformations. In S. R. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston.

Chomsky, N. (1986). *Barriers*. Cambridge, MA: MIT Press.
Church, K. W. (1980). *On memory limitations in natural language processing* (Tech. Rep. MIT/LCL/TR-245). Cambridge, MA: Laboratory for Computer Science, Massachusetts Institute of Technology.
Conrad, R. (1963). Acoustic confusions and memory span for words. *Nature, 197*, 1029–1030.
Cowper, E. A. (1976). *Constraints on sentence complexity: A model for syntactic processing*. Unpublished doctoral dissertation, Brown University.
Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.
De Roeck, A., Johnson, R., King, M., Rosner, M., Sampson, G., & Varile, N. (1982). A myth about center-embedding. *Lingua, 58*, 327–340.
Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science, 168*, 1604–1605.
Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 30*, 725–745.
Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 203–218.
Frank, R. E. (1992). *Syntactic locality and tree adjoining grammar: Grammatical, acquisition and processing perspectives*. Unpublished doctoral dissertation, University of Pennsylvania.
Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6*, 291–325.
Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.
Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove, U.K.: Erlbaum.
Gibson, E. A. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. Doctoral dissertation, Carnegie Mellon. (Available Tech. Rep. CMU-CMT-91-125, from Center for Machine Translation, Carnegie Mellon University, Pittsburgh, PA 15213.)
Gibson, E. A., & Pearlmutter, N. (1994). A corpus-based analysis of psycholinguistic constraints on PP attachment. In C. Clifton, L. Frazier, & K. Rayner, (Eds.), *Perspectives in sentence processing*. Hillsdale, NJ: Erlbaum.
Gibson, E. A., Pearlmutter, N., Canseco-Gonzales, E., & Hickok, G. (1993). *Cross-linguistic attachment preferences: Evidence from English and Spanish*. Manuscript submitted for publication.
Gordon, R. (1982). *The listener resolves an ambiguity*. Unpublished paper for the Westinghouse science talent search.
Hakes, D. T., & Cairns, H. S. (1970). Sentence comprehension and relative pronouns. *Perception & Psychophysics, 8*, 5–8.
Hakuta, K. (1981). Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition, 9*, 197–236.
Henderson, J. (1994). Description based parsing in a connectionist network. Doctoral dissertation, University of Pennsylvania. (Available as Tech. Rep. IRCS-94-12 from The Institute for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut St., Suite 400C, Philadelphia, PA 19104.)
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.

Hudgins, J. C., & Cullinan, W. L. (1978). Effects of sentence structure on sentence elicited imitation responses. *Journal of Speech and Hearing Research, 21*, 809–819.

Inoue, A., & Fodor, J. D. (1995). Information-paced parsing of Japanese. In R. Mazuka & N. Nagai, (Eds.), *Japanese sentence processing* (pp. 9–63). Hillsdale, NJ: Erlbaum.

Just, M. A., & 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* (pp. 123–124). Hillsdale, NJ: Erlbaum.

Kempen, G., & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. *Connection Science, 1*, 273–290.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition, 2*, 15–47.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence, 33*, 1–64.

Larkin, W., & Burns, D. (1977). Sentence comprehension and memory for embedded structure. *Memory & Cognition, 5*, 17–22.

Lewis, R. L. (1993). *An architecturally-based theory of human sentence comprehension.* Doctoral dissertation, Carnegie Mellon University. (Available as Tech. Rep. CMU-CS-93-226 from Computer Science Dept., Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, or reports @ cs.cmu.edu.)

Lewis, R. L. (1995). *A theory of grammatical but unacceptable embeddings.* Submitted manuscript, Cognitive Science Laboratory, Princeton University.

Lewis, R. L., & Lehman, J. F. (1994). *A theory of the computational architecture of sentence comprehension.* Submitted manuscript, Cognitive Science Laboratory, Princeton University.

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica, 75*, 55–74.

Magnussen, S., Greenlee, M. W., Asplund, R., & Dyrnes, S. (1991). Stimulus-specific mechanisms of visual short-term memory. *Vision Research, 31*, 1213–1219.

Marcus, M. P. (1980). *A theory of syntactic recognition for natural language.* Cambridge, MA: MIT Press.

Marks, L. E. (1968). Scaling of grammaticalness of self-embedded English sentences. *Journal of Verbal Learning and Verbal Behavior, 7*, 965–967.

Martin, R. C. (1993). Short-term memory and sentence processing: Evidence from neuropsychology. *Memory & Cognition, 21*, 176–183.

Mazuka, R., Itoh, K., Kiritani, S., Niwa, S., Ikejiru, K., & Naitoh, K. (1989). Processing of Japanese garden path, center-embedded, and multiply left-embedded sentences. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics (Tokyo), 23*, 187–212.

McCawley, J. D. (1988). *The syntactic phenomena of English, volume 2.* Chicago: The University of Chicago Press.

McGeoch, J. A., & McDonald, W. T. (1931). Meaningful relation and retroactive inhibition. *American Journal of Psychology, 43*, 579–588.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *The Psychological Review, 63*, 81–97.

Miller, G. A. (1962). Some psychological studies of grammar. *American Psychologist, 17*, 748–762.