MEMORY CAPACITY AND SENTENCE PROCESSING

Edward Gibson
Department of Philosophy, Carnegie Mellon University
Pittsburgh, PA 15213-3890
gibson@cs.cmu.edu

ABSTRACT

The limited capacity of working memory is intrinsic to human sentence processing, and therefore must be addressed by any theory of human sentence processing. This paper gives a theory of garden-path effects and processing overload that is based on simple assumptions about human short term memory capacity.

1 INTRODUCTION

The limited capacity of working memory is intrinsic to human sentence processing, and therefore must be addressed by any theory of human sentence processing. I assume that the amount of short term memory that is necessary at any stage in the parsing process is determined by the syntactic, semantic and pragmatic properties of the structure(s) that have been built up to that point in the parse. A sentence becomes unacceptable for processing reasons if the combination of these properties produces too great a load for the working memory capacity (*cf.* Frazier 1985):

(1)

$$\sum_{i=1}^{n} A_i x_i > K$$

where:

K is the maximum allowable processing load (in processing load units or PLUs),

x_i is the number of PLUs associated with property i,

n is the number of properties,

A_i is the number of times property i appears in the structure in question.

Furthermore, the assumptions described above provide a simple mechanism for the explanation of common psycholinguistic phenomena such as garden-path effects and preferred readings for ambiguous sentences. Following Fodor (1983), I assume that the language processor is an automatic device that uses a greedy algorithm: only the best of the set of all compatible representations for an input string are locally maintained from word to word. One way to make this idea explicit is to assume that restrictions on memory allow at most one representation for an input string at any time (see, for example, Frazier and Fodor 1978; Frazier 1979; Marcus 1980; Berwick and Weinberg 1984; Pritchett 1988). This hypothesis, commonly called the serial hypothesis, is easily compatible with the above view of processing load calculation: given a choice between two different representations for the same input string, simply choose the representation that is associated with the lower processing load.

The serial hypothesis is just one way of placing local memory restrictions on the parsing model, however. In this paper I will present an alternative formulation of local memory restrictions within a parallel framework.

There is a longstanding debate in the psycholinguistic literature as to whether or not more than one representation for an input can be maintained in parallel (see, for example, Kurtzman (1985) or Gorrell (1987) for a history of the debate). It turns out that the parallel view appears to handle some kinds of data more directly than the serial view, keeping in mind that the data are often controversial. For example, it is difficult to explain in a serial model why relative processing load increases as ambiguous input is encountered (see, for example, Fodor *et al.* 1968; Rayner *et al.* 1983; Gorrell 1987). Data that is normally taken to be support for the serial hypothesis includes garden-path effects and the existence of preferred readings of ambiguous input. However, as noted above, limiting the number of allowable representations is only one way of constraining parallelism so that these effects can also be accounted for in a parallel framework.

As a result of the plausibility of a parallel model, I propose to limit the *difference in processing load* that may be present between two structures for the same input, rather than limit the *number* of structures allowed in the processing of an input (*cf.* Gibson 1987; Gibson and Clark 1987; Clark and Gibson 1988). Thus I assume that the human parser prefers one structure over another when the processing load (in PLUs) associated with maintaining the first is markedly lower than the processing load associated with maintaining the second. That is, I assume there exists some arithmetic preference quantity P corresponding to a processing load, such that if the processing loads associated with two representations for the same string differ by load P, then only the representation associated with the smaller of the two loads is pursued.[1] Given the existence of a

---

[1] It is possible that the preference factor is a *geometric* one rather than an *arithmetic* one. Given a geometric preference factor, one structure is preferred over another when the ratio of their processing loads reaches a threshold value. I explore only the arithmetic possibility in this paper; it is possible that the geometric alternative gives results that are as good, although I leave this issue for future research.

preference factor P, it is easy to account for garden-path effects and preferred readings of ambiguous sentences. Both effects occur because of a local ambiguity which is resolved in favor of one reading. In the case of a garden-path effect, the favored reading is not compatible with the whole sentence. Given two representations for the same input string that differ in processing load by at least the factor P, only the less computationally expensive structure will be pursued. If that structure is not compatible with the rest of the sentence and the discarded structure is part of a successful parse of the sentence, a garden-path effect results. If the parse is successful, but the discarded structure is compatible with another reading for the sentence, then only a preferred reading for the sentence has been calculated. Thus if we know where one reading of a (temporarily) ambiguous sentence becomes the strongly preferred reading, we can write an inequality associated with this preference:

(2)

$$\sum_{i=1}^{n} A_i x_i - \sum_{i=1}^{n} B_i x_i > P$$

where:

P is the preference factor (in PLUs),

x_i is the number of PLUs associated with property i,

n is the number of properties,

A_i is the number of times property i appears in the unpreferred structure,

B_i is the number of times property i appears in the preferred structure.

Given a parsing algorithm together with n properties and their associated processing loads $x_1 \ldots x_n$, we may write inequalities having the form of (1) and (2) corresponding to the processing load at various parse states. An algebraic technique called *linear programming* can then be used to solve this system of linear inequalities, giving an n-dimensional space for the values of x_i as a solution, any point of which satisfies all the inequalities.

In this paper I will concentrate on syntactic properties:[2] in particular, I present two properties based on the θ-Criterion of Government and Binding Theory (Chomsky 1981).[3] It will be shown that these properties, once associated with processing loads, predict a large array of garden-path effects. Furthermore, it is demonstrated that these properties also make desirable predictions with respect to unacceptability due to memory capacity overload.

The organization of this paper is given as follows: first, the structure of the underlying parser is described; second, the two syntactic properties are proposed; third, a number of locally ambiguous sentences, including some garden-paths, are examined with respect to these properties and a solution space for the processing loads of the two properties is calculated; fourth, it is shown that this space seems to make the right predictions with respect to processing overload; conclusions are given in the final section.

2 THE UNDERLYING PARSER

The parser to which the memory limitation constraints apply must construct representations in such a way so that incomplete input will be associated with some structure. Furthermore, the parsing algorithm must, in principle, allow more than one structure for an input string, so that the general constraints described in the previous section may apply to restrict the possibilities. The parsing model that I will assume is an extension of the model described in Clark and Gibson (1988). When a word is input, representations for each of its lexical entries are built and placed in the *buffer*, a one cell data structure that holds a set of tree structures. The parsing model contains a second data structure, the *stack-set*, which contains a set of stacks of buffer cells. The parser builds trees in parallel based on possible attachments made between the buffer and the top of each stack in the stack-set. The buffer and stack-set are formally defined in (3) and (4).

(3) A *buffer cell* is a set of structures $\{ S_1, \ldots, S_n \}$, where each S_i represents the same segment of the input string. The *buffer* contains one buffer cell.

(4) The *stack-set* is a set of stacks of buffer cells, where each stack represents the same segment of the input string:

$$\begin{aligned}
\{ (&\{ S_{1,1,1}, S_{1,1,2}, \ldots, S_{1,1,n_{1,1}} \}, \\
&\{ S_{1,2,1}, S_{1,2,2}, \ldots, S_{1,2,n_{1,2}} \}, \ldots \\
&\{ S_{1,m_1,1}, S_{1,m_1,2}, \ldots, S_{1,m_1,n_{1,m_1}} \}) \\
&\ldots \\
(&\{ S_{p,1,1}, S_{p,1,2}, \ldots, S_{p,1,n_{p,1}} \}, \\
&\{ S_{p,2,1}, S_{p,2,2}, \ldots, S_{p,2,n_{p,2}} \}, \ldots \\
&\{ S_{p,m_p,1}, S_{p,m_p,2}, \ldots, S_{p,m_p,n_{p,m_p}} \}) \}
\end{aligned}$$

where:

p is the number of stacks;

m_i is the number of buffer cells in stack i;

and $n_{i,j}$ is the number of tree structures in the jth buffer cell of stack i.

The motivation for these data structures is given by the desire for a completely unconstrained parsing algorithm upon which constraints may be placed: this algorithm should allow all possible parser operations to occur at each parse state. There are exactly two parser operations: attaching a node to another node and

---

[2] Note that I assume that there also exist semantic and pragmatic properties which are associated with significant processing loads, but which are not discussed here.

[3] In another syntactic theory, similar properties may be obtained from the principles that correspond to the θ-Criterion in that theory. For example, the completeness and coherence conditions of Lexical Functional Grammar (Bresnan 1982) would derive properties similar to those derived from the θ-Criterion. The same empirical effects should result from these two sets of properties.

pushing a buffer cell onto a stack. In order to allow both of these operations to be performed in parallel, it is necessary to have the given data structures: the buffer and the stack-set. For example, consider a parser state in which the buffer is non-empty and the stack-set contains only a single cell stack:

(5)
Stack-set: $\{ \{ \{ S_1, ..., S_n \} \} \}$
Buffer: $\{ B_1, ..., B_m \}$

Suppose that attachments are possible between the buffer and the single stack cell. The structures that result from these attachments will take up a single stack cell. Let us call these resultant structures $A_1, A_2, ..., A_k$. If all possible operations are to take place at this parser state, then the contents of the current buffer must also be pushed on top of the current stack. Thus two stacks, both representing the same segment of the input string will result:

(6)
Stack 1: $\{ \{ \{ A_1, ..., A_k \} \} \}$
Stack 2: $\{ \{ \{ B_1, ..., B_m \} \{ S_1, ..., S_n \} \} \}$

Since these two stacks break up the same segment of the input string in different ways, the stack-set data structure is necessary.

3 TWO SYNTACTIC PROPERTIES DERIVABLE FROM THE θ-CRITERION

Following early work in linguistic theory, I distinguish two kinds of categories: *functional* categories and *thematic* or *content* categories (see, for example, Fukui and Speas (1986) and Abney (1987) and the references cited in each). Thematic categories include nouns, verbs, adjectives and prepositions; functional categories include determiners, complementizers, and inflection markers. There are a number of properties that distinguish functional elements from thematic elements, the most crucial being that functional elements mark grammatical or relational features while thematic elements pick out a class of objects or events. I will assume as a working hypothesis that only those syntactic properties that have to do with the thematic elements of an utterance are relevant to preferences and overload in processing. One principle of syntax that is directly involved with the thematic content of an utterance in a Government-Binding theory is the θ-Criterion:

(7) Each argument bears one and only one θ-role (thematic role) and each θ-role is assigned to one and only one argument (Chomsky 1981:36).

I hypothesize that the human parser attempts to locally satisfy the θ-Criterion whenever possible. Thus given a thematic role, the parser prefers to assign that role, and given a thematic element, the parser prefers to assign a role to that element. These assumptions are made explicit as the following properties:

(8) The Property of Thematic Reception (PTR):
Associate a load of x_{TR} PLUs of short term memory to each thematic element that is in a position that can receive a thematic role in some co-existing structure, but whose θ-assigner is not unambiguously identifiable in the structure in question.

(9) The Property of Thematic Assignment (PTA):
Associate a load of x_{TA} PLUs of short term memory to each thematic role that is not assigned to a node containing a thematic element.

Note that the Properties of Thematic Assignment and Reception are stated in terms of *thematic* elements. Thus the Property of Thematic Reception doesn't apply to functional categories, whether or not they are in positions that receive thematic roles. Similarly, if a thematic role is assigned to a functional category, the Property of Thematic Assignment does not notice until there is a thematic element inside this constituent.

4 AMBIGUITY AND THE PROPERTIES OF THEMATIC ASSIGNMENT AND RECEPTION

Consider sentence (10) with respect to the Properties of Thematic Assignment and Reception:

(10) John expected Mary to like Fred.

The verb *expect* is ambiguous: either it takes an NP complement as in the sentence *John expected Mary* or it takes an IP complement as in (10).[4] Consider the state of the parse of (10) after the word *Mary* has been processed:

(11) a. $[_{IP}$ $[_{NP}$ John $]$ $[_{VP}$ expected $[_{NP}$ Mary $]]]$
b. $[_{IP}$ $[_{NP}$ John $]$ $[_{VP}$ expected $[_{IP}$ $[_{NP}$ Mary $]$ $]]]$

In (11a), the NP *Mary* is attached as the NP complement of *expected*. In this representation there is no load associated with either of the Properties of Thematic Assignment or Reception since no thematic elements need thematic roles and no thematic roles are left unassigned. In (11b), the NP *Mary* is the specifier of a hypothesized IP node which is attached as the complement of the other reading of *expected*.[5] This representation is associated with at least x_{TR} PLUs since the NP *Mary* is in a position that can be associated with a thematic role, the subject position, but whose θ-assigner is not yet identifiable. No load is associated with the Property of Thematic Assignment, however, since both thematic roles of the verb *expected* are assigned to nodes that contain thematic elements. Since

---

[4] Following current notation in GB Theory, IP (Inflection Phrase) = S and CP (Complementizer Phrase) = S′ (Chomsky 1986).

[5] I assume some form of hypothesis-driven node projection so that noun phrases are projected to those categories that they may specify. Motivation for this kind of projection algorithm is given by the processing of Dutch (Frazier 1987) and the processing of certain English noun phrase constructions (Gibson 1989).

there is no difficulty in processing sentence (10), the load difference between these two structures cannot be greater than P PLUs, the preference factor in inequality (2). Thus the inequality in (12) is obtained:

(12) $x_{TR} \leq P$

Since the load difference between the two structures is not sufficient to cause a strong preference, both structures are maintained. Note that this is an important difference between the theories presented here and the theory presented in Frazier and Fodor (1978), Frazier (1979) and Pritchett (1988). In each of these theories, only one representation can be maintained, so that either (11a) or (11b) would be preferred. In order to account for the lack of difficulty in parsing (10), Frazier and Pritchett both assume that reanalysis in certain situations is not expensive. No such stipulation is necessary in the framework given here: it is simply assumed that *all* reanalysis is expensive.[6]

Consider now sentence (13) with respect to the Properties of Thematic Assignment and Reception:

(13) John expected her mother to like Fred.

Consider the state of the parse of (13) after the word *her* has been processed. In one representation the NP *her* will be attached as the NP complement of *expected*:

(14) $[_{IP}$ $[_{NP}$ John] $[_{VP}$ expected $[_{NP}$ her]]]

In this representation there is no load associated with either of the Properties of Thematic Assignment or Reception since no thematic objects need thematic roles and no thematic roles are left unassigned. In another representation the NP *her* is the specifier of a hypothesized NP which is pushed onto a substack containing the other reading of the verb *expected*:

(15) { { $[_{IP}$ $[_{NP}$ John] $[_{VP}$ expected $[_{IP}$ e]]] }
{ $[_{NP}$ $[_{NP}$ her]] } }

This representation is associated with at least x_{TA} PLUs since the verb *expected* has a thematic role to assign. However, no load is associated with the genitive NP specifier *her* since its θ-assigner, although not yet present, is unambiguously identified as the head of the NP to follow (Chomsky (1986a)).[7] Thus the total load associated with (15) is x_{TA} PLUs. Since there is no difficulty in processing sentence (10), the load difference between these two structures cannot be greater than P PLUs. Thus the second inequality, (18), is obtained:

(18) $x_{TA} \leq P$

Now consider (19):[8]

(19) # I put the candy on the table in my mouth.

This sentence becomes ambiguous when the preposition *on* is read. This preposition may attach as an argument of the verb *put* or as a modifier of the NP *the candy*:

(20) a. I $[_{VP}$ $[_{V'}$ $[_{V}$ put] $[_{NP}$ the candy] $[_{PP}$ on]]]
b. I $[_{VP}$ $[_{V'}$ $[_{V}$ put] $[_{NP}$ the candy $[_{PP}$ on]]]]

At this point the argument attachment is strongly preferred. However, this attachment turns out to be incompatible with the rest of the sentence. When the word *mouth* is encountered, no pragmatically coherent structure can be built, since tables are not normally found in mouths. Thus a garden-path effect results. Consider the parse state depicted in (20) with respect to the Properties of Thematic Assignment and Reception. The load associated with the structure resulting from argument attachment is x_{TA} PLUs since, although the θ-grid belonging to the verb *put* is filled, the thematic role assigned by the preposition *on* remains unassigned. On the other hand, the load associated with the modifier attachment is $2 * x_{TA} + x_{TR}$ PLUs since 1) both the verb *put* and the preposition *on* have thematic roles that need to be assigned and 2) the PP headed by *on* receives a thematic role in the argument attachment structure, while it receives no such role in the structure under consideration. Thus the difference between the loads associated with the two structures is $x_{TA} + x_{TR}$ PLUs. Since the argument attachment structure is strongly preferred over the other structure, I hypothesize that this load is greater than P PLUs:

(21) $x_{TA} + x_{TR} > P$

Now consider the the well-known garden-path sentence in (22):

(22) # The horse raced past the barn fell.

The structure for the input *the horse raced* is ambiguous between at least the two structures in (23):

(23) a. $[_{IP}$ $[_{NP}$ the horse] $[_{VP}$ raced]]
b. $[_{IP}$ $[_{NP}$ the $[_{N'}$ $[_{N'}$ horse$_i$] $[_{CP}$ O_i raced]]]]

Structure (23a) has no load associated with it due to either the PTA or the PTR. Crucially note that the verb *raced* has an intransitive reading so that no load is required via the Property of Thematic Assignment. On the other hand, structure (23b) requires a load of $2 * x_{TR}$ PLUs since 1) the noun phrase *the horse* is in a position that can receive a thematic role, but currently does not and 2) the operator O_i is in a position that may be associated with a thematic role, but is not yet

---

[6] See Section 4.1 for a brief comparison between the model proposed here and serial models such as those proposed by Frazier and Fodor (1978) and Pritchett (1988).

[7] Note that specifiers do not always receive their thematic roles from the categories which they specify. For example, a non-genitive noun phrase may specify any major category. In particular, it may specify an IP or a CP. But the specifier of these categories may receive its thematic role through chain formation from a distant θ-assigner, as in (16):

(16) John appears to like beans.

Note that there is no NP that corresponds to (16) (Chomsky (1970)):

(17) * John's appearance to like beans.

[8] I will prefix sentences that are difficult to parse because of memory limitations with the symbol "#". Hence sentences that are unacceptable due to processing overload will be prefixed with "#", as will be garden-path sentences.

associated with one.[9] Thus the difference between the processing loads of structures (23a) and (23b) is $2 * x_{TR}$ PLUs. Since this sentence is a strong garden-path sentence, it is hypothesized that a load difference of $2 * x_{TR}$ PLUs is greater than the allowable limit, P PLUs:

(24) $2 * x_{TR} > P$

A surprising effect occurs when a verb which *optionally* subcategorizes for a direct object, like *race*, is replaced by a verb which *obligatorily* subcategorizes for a direct object, like *find*:

(25) The bird found in the room was dead.

Although the structures and local ambiguities in (25) and (22) are similar, (22) causes a garden-path effect while, surprisingly, (25) does not. To determine why (25) is not a garden-path sentence we need to examine the local ambiguity when the word *found* is read:

(26) a. $[_{IP}\ [_{NP}$ the bird $]\ [_{VP}\ [_{V'}\ [_{V}$ found $]\ [_{NP}\ ]\ ]]]$
b. $[_{IP}\ [_{NP}$ the $[_{N'}\ [_{N'}$ bird$_i$ $]\ [_{CP}\ O_i$ found $]\ ]]\ ]$

The crucial difference between the verb *found* and the verb *raced* is that *found* requires a direct object, while *raced* does not. Since the θ-grid of the verb *found* is not filled in structure (26a), this representation is associated with x_{TA} PLUs of memory load. Like structure (23b), structure (26b) requires $2 * x_{TR}$ PLUs. Thus the difference between the processing loads of structures (26a) and (26b) is $2 * x_{TR} - x_{TA}$ PLUs. Since no garden-path effect results in (25), I hypothesize that this load is less than or equal to P PLUs:

(27) $2 * x_{TR} - x_{TA} \leq P$

Furthermore, these results correctly predict that sentence (28) is not a garden-path sentence either:

(28) The bird found in the room enough debris to build a nest.

Hence we have the following system of inequalities:

(29) a. $x_{TR} \leq P$
b. $x_{TA} \leq P$
c. $x_{TA} + x_{TR} > P$
d. $2 * x_{TR} > P$
e. $2 * x_{TR} - x_{TA} \leq P$

This system of inequalities is consistent. Thus it identifies a particular solution space. This solution space is depicted by the shaded region in Figure 1.

Note that, pretheoretically, there is no reason for this system of inequalities to be consistent. It could have been that the parser state of one of the example sentences forced an inequality that contradicted some previously obtained inequality. This situation would have had one of three implications: the properties being considered might be incorrect; the properties being considered might be incomplete; or the whole approach might be incorrect. Since this situation has not yet been observed, the results mutually support one another.

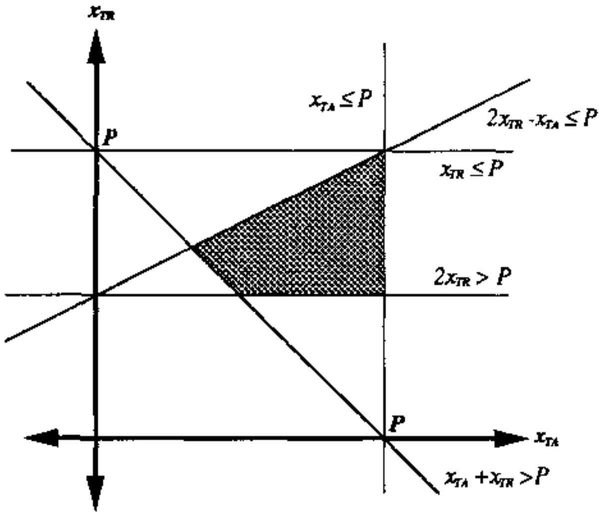

Figure 1: The Solution Space for the Inequalities in (29)

4.1 A COMPARISON WITH SERIAL MODELS

Because serial models of parsing can maintain at most one representation for any input string, they have difficulty explaining the lack of garden-path effects in sentences like (10) and (25):

(10) John expected Mary to like Fred.

(25) The bird found in the room was dead.

As a result of this difficulty Pritchett (1988) proposes the Theta Reanalysis Constraint:[10]

(30) Theta Reanalysis Constraint (TRC): Syntactic reanalysis which interprets a θ-marked constituent as outside its current θ-Domain and as within an existing θ-Domain of which it is not a member is costly.

(31) θ-Domain: α is in the γ θ-Domain of β iff α receives the γ θ-role from β or α is dominated by a constituent that receives the γ θ-role from β.

As a result of the Theta Reanalysis Constraint, the necessary reanalysis in each of (10) and (25) is not expensive, so that no garden-path effect is predicted. Furthermore, the reanalysis in sentences like (22) and (19) violates the TRC, so that the garden-path effects are predicted.

However, there are a number of empirical problems with Pritchett's theory. First of all, it turns out that the

---

[9] In fact, this operator will be associated with a thematic role as soon as a gap-positing algorithm links it with the object of the passive participle *raced*. However, when the attachment is initially made, no such link yet exists: the operator will initially be unassociated with a thematic role.

[10] Frazier and Rayner (1982) make a similar stipulation to account for problems with the theory of Frazier and Fodor (1978). However, their account fails to explain the lack of garden-path effect in (25). See Pritchett (1988) for a description of further problems with their analysis.

Theta Reanalysis Constraint as defined in (30) incorrectly predicts that the sentences in (32) do not induce garden-path effects:

(32) a. # The horse raced past the barn was falling.
b. # The dog walked to the park seemed small.
c. # The boat floated down the river was a canoe.

For example, consider (32a). When the auxiliary verb *was* is encountered, reanalysis is forced. However, the auxiliary verb *was* does not have a thematic role to assign to its subject, *the dog*, so the TRC is not violated. Thus Pritchett's theory incorrectly predicts that these sentences do not cause garden-path effects.

Other kinds of local ambiguity that do not give the human parser difficulty also pose a challenge to serial parsers. Marcus (1980) gives the sentences in (33) as evidence that any deterministic parser must be able to look ahead in the input string:[11]

(33) a. Have the boys taken the exam today?
b. Have the boys take the exam today.

Any serial parser must be able to account for the lack of difficulty with either of the sentences in (33). It turns out that the Theta Reanalysis Constraint does not help in cases like these: no matter which analysis is pursued first, reanalysis will violate the TRC.

4.2 EMPIRICAL SUPPORT: FURTHER GARDEN-PATH EFFECTS

Given the Properties of Thematic Assignment and Reception and their associated loads, we may now explain many more garden-path effects. Consider (34):

(34) # The Russian women loved died.

Up until the last word, this sentence is ambiguous between two readings: one where *loved* is the matrix verb; and the other where *loved* heads a relative clause modifier of the noun *Russian*. The strong preference for the matrix verb interpretation of the word *loved* can be easily explained if we examine the possible structures upon reading the word *women*:

(35) a. [$_{IP}$ [$_{NP}$ the Russian women]]
b. [$_{IP}$ [$_{NP}$ the [$_{N'}$ [$_{N'}$ Russian$_i$] [$_{CP}$ [$_{NP}$ O_i] [$_{IP}$ [$_{NP}$ women]]]]]]

Structure (35a) requires x_{TR} PLUs since the NP *the Russian women* needs but currently lacks a thematic role. Structure (35b), on the other hand, requires at least $3 * x_{TR}$ PLUs since 1) two noun phrases, *the Russian* and *women*, need but currently lack thematic roles; and 2) the operator in the specifier position of the modifying Comp phrase can be associated with a thematic role, but currently is not linked to one. Since the difference between these loads is $2 * x_{TR}$, a garden-path effect results.

Consider now (36):

(36) # John told the man that Mary kissed that Bill saw Phil.

When parsing sentence (36), people will initially analyze the CP *that Mary kissed* unambiguously as an argument of the verb *told*. It turns out that this hypothesis is incompatible with the rest of the sentence, so that a garden-path effect results. In order to see how the garden-path effect comes about, consider the parse state which occurs after the word *Mary* is read:

(37) a. [$_{IP}$ [$_{NP}$ John] [$_{VP}$ [$_{V'}$ [$_{V}$ told] [$_{NP}$ the man] [$_{CP}$ that] [$_{IP}$ [$_{NP}$ Mary]]]]]]
b. [$_{IP}$ [$_{NP}$ John] [$_{VP}$ [$_{V'}$ [$_{V}$ told] [$_{NP}$ the [$_{N'}$ [$_{N'}$ man$_i$] [$_{CP}$ [$_{NP}$ O_i] that [$_{IP}$ [$_{NP}$ Mary]]]]]]

Structure (37a) requires no load by the PTA since the θ-grid of the only θ-assigner is filled with structures that each contain thematic elements. However, the noun phrase *Mary* requires x_{TR} PLUs by the Property of Thematic Reception since this NP is in a thematic position but does not yet receive a thematic role. Thus the total load associated with structure (37a) is x_{TR} PLUs. Structure (37b), on the other hand, requires a load of $x_{TA}+2*x_{TR}$ since 1) the thematic role PROPOSITION is not yet assigned by the verb *told*; 2) the operator in the specifier position of the CP headed by *that* is not linked to a thematic role; and 3) the NP *Mary* is in thematic position but does not receive a thematic role yet. Thus the load difference is $x_{TA}+x_{TR}$ PLUs, enough for the more expensive one to be dropped. Thus only structure (37a) is maintained and a garden-path effect eventually results, since this structure is not compatible with the entire sentence. Hence the Properties of Thematic Assignment and Reception make the correct predictions with respect to (36).

Consider the garden-path sentence in (38):

(38) # John gave the boy the dog bit a dollar.

This sentence causes a garden-path effect since the noun phrase *the dog* is initially analyzed as the direct object of the verb *gave* rather than as the subject of a relative clause modifier of the NP *the boy*. This garden-path can be explained in the same way as previous examples. Consider the state of the parse after the NP *the dog* has been processed:

(39) a. [$_{IP}$ [$_{NP}$ John] [$_{VP}$ [$_{V'}$ [$_{V}$ gave] [$_{NP}$ the boy] [$_{NP}$ the dog]]]]
b. [$_{IP}$ [$_{NP}$ John] [$_{VP}$ [$_{V'}$ [$_{V}$ gave] [$_{NP}$ the [$_{N'}$ [$_{N'}$ boy$_i$] [$_{CP}$ [$_{NP}$ O_i] [$_{IP}$ [$_{NP}$ the dog]]] [$_{NP}$]]]]]

While structure (39a) requires no load at this stage, structure (39b) requires $2 * x_{TR} + x_{TA}$ PLUs since 1) one thematic role has not yet been assigned by the verb *gave*; 2) the operator in the specifier position of the CP modifying *boy* is not linked to a thematic role; and 3) the NP *the dog* is in a thematic position but does not yet receive a thematic role. Thus structure (39a) is strongly preferred and a garden-path effect results.

The garden-path effect in (40) can also be easily explained in this framework:

(40) # The editor authors the newspaper hired liked laughed.

---

[11] Note that model that I am proposing here is a parallel model, and therefore is nondeterministic.

Consider the state of the parse of (40) after the word *authors* has been read:

(41) a. $[_{IP}\ [_{NP}$ the editor $]\ [_{VP}\ [_{V'}\ [_{V}$ authors $]\ [_{NP}\ ]\ ]]]$
b. $[_{IP}\ [_{NP}$ the $[_{N'}\ [_{N'}$ editor$_i$ $]\ [_{CP}\ [_{NP}\ O_i\ ]\ [_{IP}\ [_{NP}$ authors $]\ ]]\ ]]]$

The word *authors* is ambiguous between nominal and verbal interpretations. The structure including the verbal reading is associated with x_{TA} PLUs since the θ-grid for the verb *authors* includes an unassigned role.

Structure (41b), on the other hand, includes three noun phrases, each of which is in a position that may be linked to a thematic role but currently is not linked to any θ-role. Thus the load associated with structure (41b) is $3 * x_{TR}$ PLUs. Since the difference between the loads associated with structures (41b) and (41a) is so high ($3 * x_{TR} - x_{TA}$ PLUs), only the inexpensive structure, structure (41a), is maintained.

5 PROCESSING OVERLOAD

The Properties of Thematic Assignment and Reception also give a plausible account of the unacceptability of sentences with an abundance of center-embedding. Recall that I assume that a sentence is unacceptable because of short term memory overload if the combination of memory associated with properties of the structures built at some stage of the parse of the sentence is greater than the allowable processing load K. Consider (42):

(42) # The man that the woman that the dog bit likes eats fish.

Having input the noun phrase *the dog* the structure for the partial sentence is as follows:

(43) $[_{IP}\ [_{NP}$ the $[_{N'}\ [_{N'}$ man$_i$ $]\ [_{CP}\ [_{NP}\ O_i\ ]$ that $[_{IP}\ [_{NP}$ the $[_{N'}\ [_{N'}$ woman$_j$ $]\ [_{CP}\ [_{NP}\ O_j\ ]$ that $[_{IP}\ [_{NP}$ the dog $]\ ]]]$

In this representation there are three lexical noun phrases that need thematic roles but lack them. Furthermore, there are two non-lexical NPs, operators, that are in positions that may prospectively be linked to thematic roles. Thus, under my assumptions, the load associated with this representation is at least $5 * x_{TR}$ PLUs. I assume that these properties are responsible for the unacceptability of this sentence, resulting in the inequality in (44):

(44) $5 * x_{TR} > K$

Note that sentences with only one relative clause modifying the subject are acceptable, as is exemplified in (45)

(45) The man that the woman likes eats fish.

Since (45) is acceptable, its load is below the maximum at all stages of its processing. After processing the noun phrase *the woman* in (45), there are three noun phrases that currently lack θ-roles but may be linked to θ-roles as future input appears. Thus we arrive at the inequality in (46):

(46) $3 * x_{TR} \leq K$

Thus I assume that the maximum processing load that people can handle lies somewhere above $3 * x_{TR}$ PLUs but somewhere below $5 * x_{TR}$ PLUs. Although these data are only suggestive, they clearly make the right kinds of predictions. Future research should establish the boundary between acceptability and unacceptability more precisely.

6 CONCLUSIONS

Since the structural properties that are used in the formation of the inequalities are independently motivated, and the system of inequalities is solvable, the theory of human sentence processing presented here makes strong, testable predictions with respect to the processability of a given sentence. Furthermore, the success of the method provides empirical support for the particular properties used in the formation of the inequalities. Thus a theory of PLUs, the preference factor P and the overload factor K provides a unified account of 1) acceptability and relative acceptability; 2) garden-path effects; and 3) preferred readings for ambiguous input.

7 ACKNOWLEDGEMENTS

I would like to thank Robin Clark, Dan Everett, Rick Kazman, Howard Kurtzman and Eric Nyberg for comments on earlier drafts of this work. All remaining errors are my own.

8 REFERENCES

Abney, Stephen P. 1987 The English Noun Phrase in its Sentential Aspect. Ph.D. Thesis, MIT, Cambridge, MA.

Berwick, Robert C. and Weinberg, Amy S. 1984 *The Grammatical Basis for Linguistic Performance.* MIT Press, Cambridge, MA.

Bresnan, Joan 1982 *The Mental Representation of Grammatical Relations.* MIT Press, Cambridge, MA.

Chomsky, Noam 1970 Remarks on Nominalization. In R. Jacobs and P. Rosenbaum (eds.), *Readings in English Transformational Grammar*, Ginn, Waltham, MA:184-221.

Chomsky, Noam 1981 *Lectures on Government and Binding.* Foris, Dordrecht, The Netherlands.

Chomsky, Noam 1986 *Barriers.* Linguistic Inquiry Monograph 13, MIT Press, Cambridge, MA.

Clark, Robin and Gibson, Edward 1988 A Parallel Model for Adult Sentence Processing. In: *Proceedings of the Tenth Cognitive Science Conference*, McGill University, Montreal, Quebec:270-276.

Fodor, Jerry A. 1983 *Modularity of Mind.* MIT Press, Cambridge, MA.

Fodor, Jerry A.; Garrett, Merrill F. and Bever, Tom G. 1968 Some Syntactic Determinants of Sentential Complexity. *Perception and Psychophysics* 2:289-96.

Frazier, Lyn 1979 On Comprehending Sentences: Syntactic Parsing Strategies. Ph.D. Thesis, University of Massachusetts, Amherst, MA.

Frazier, Lyn 1985 Syntactic Complexity. In Dowty, David, Karttunen, Lauri, and Arnold Zwicky (eds.), *Natural Language Processing: Psychological, Computational and Theoretical Perspectives*, Cambridge University Press, Cambridge, United Kingdom:129-189.

Frazier, Lyn 1987 Syntactic Processing Evidence from Dutch. *Natural Language and Linguistic Theory* 5:519-559.

Frazier, Lyn and Fodor, Janet Dean 1978 The Sausage Machine: A New Two-stage Parsing Model. *Cognition* 6:291-325.

Fukui, Naoki and Speas, Margaret 1986 Specifiers and Projections. *MIT Working Papers in Linguistics* 8, Cambridge, MA:128-172.

Gibson, Edward 1987 Garden-Path Effects in a Parser with Parallel Architecture. In: *Proceedings of the Fourth Eastern States Conference on Linguistics*, The Ohio State University, Columbus, OH:88-99.

Gibson, Edward 1989 Parsing with Principles: Predicting a Phrasal Node Before Its Head Appears. In: *Proceedings of the First International Workshop on Parsing Technologies*, Carnegie Mellon University, Pittsburgh, PA:63-74.

Gibson, Edward and Clark, Robin 1987 Positing Gaps in a Parallel Parser. In: *Proceedings of the Eighteenth North East Linguistic Society Conference*, University of Toronto, Toronto, Ontario:141-155.

Gorrell, Paul G. 1987 Studies of Human Syntactic Processing: Ranked-Parallel versus Serial Models. Ph.D. Thesis, University of Connecticut, Storrs, CT.

Kurtzman, Howard 1985 Studies in Syntactic Ambiguity Resolution. Ph.D. Thesis, MIT, Cambridge, MA.

Marcus, Mitchell P. 1980 *A Theory of Syntactic Recognition for Natural Language*. MIT Press, Cambridge, MA.

Pritchett, Bradley 1988 Garden Path Phenomena and the Grammatical Basis of Language Processing. *Language* 64:539-576.

Rayner, Keith; Carlson, Marcia and Frazier, Lyn 1983 The Interaction of Syntax and Semantics during Sentence Processing: Eye Movements in the Analysis of Semantically Biased Sentences. *Journal of Verbal Learning and Verbal Behavior* 22:358-374.