Constraints on sentence comprehension

Edward Gibson[a] and Neal J. Pearlmutter[b]

May 15, 1998

[a]Department of Brain and Cognitive Sciences
E10-034C
Massachusetts Institute of Technology
Cambridge, MA 02139

[b]Psychology Department
125 NI
Northeastern University
Boston, MA 02115

Electronic mail: gibson@psyche.mit.edu, pearlmutter@neu.edu

Summary

The process of comprehending sentences involves the integration of a variety of different information sources, constrained by the available computational resources. This paper surveys the evidence for four types of constraints on sentence comprehension: 1) lexical constraints, 2) contextual constraints, 3) locality-based computational resource constraints, and 4) phrase-level contingent frequency constraints.

Research in sentence comprehension attempts to discover how people understand sentences. The process of understanding a sentence involves computing its meaning by combining the meanings of its individual words. Thus this research examines the representations and processes which are needed to connect the identification of individual words in speech or reading[1,2] with the mapping of sentence meanings onto mental models[3] or discourse representations[4,5]. We will focus on a specific current approach, according to which sentence interpretations are determined by a set of interacting informational constraints[6–8]; Altmann[9] provides a broader historical view of the field.

Recent results have suggested that constructing an interpretation for a sentence involves the moment-by-moment integration of a variety of different information sources, constrained by the available computational resources[6–8,10–13]. We assume that the foundations for the representations constructed during sentence comprehension are created by projection from the representations for individual words[6,14–18], and that the individual word representations are then combined to form phrases. See Box 1 for a summary of the kinds of representations involved. This combination process is constrained by phrase-formation information[14,15,18], which has two main influences: First, it delimits the range of possible combinations. For example, *the dog is happy* is a well-formed English sentence, but *dog the is happy* is not. Second, phrase-formation information partially determines the interpretations for the combinations. For example, *Eleanor loved Chris* and *Chris loved Eleanor* mean different things in English (despite involving the same individual word representations), because phrase-formation constraints require that the initial noun in a simple noun-verb-noun sentence be interpreted as the performer of the action specified by the verb, and that the later noun be interpreted as the entity on which the action is performed. We can categorize the remaining constraints into four broad categories which are relevant for both spoken and written sentence comprehension:

1. *Lexical*, or word-level, constraints, which depend on knowledge associated with particular words in a language;

2. *Contextual* constraints, which involve the communicative utility[23] and the plausibility of different interpretations, given knowledge about the state of the world;

3. *Computational resource* constraints, which depend on the availability of and access to working memory resources;

4. Phrase-level *contingent frequency* constraints, defined as the probability of phrases occurring in particular phrase structure contexts.

We describe the four kinds of constraints in more detail below, assuming a framework in which all the constraints apply freely. One additional constraint is *prosody*, which refers to properties of the speech signal beyond those related to the identity of words. These properties include variations in the pitch, amplitude, and duration of individual speech sounds and larger segments, as well as the placement of pauses. Although prosodic constraints are relevant in spoken sentence comprehension[24], space limitations require that we restrict our attention to those constraints which apply to both spoken and written comprehension. Note that even in spoken comprehension, it is not clear whether helpful prosodic information is reliably produced by speakers[25], so that an explanation for sentence comprehension will necessarily involve the four constraints listed above.

Lexical constraints

The influence on sentence-level comprehension of some word-level properties such as grammatical category (whether a word is a noun, a verb, etc.) has long been assumed[14], and the effect of additional word-level properties has been considered[16]. However, the variety and importance of such constraints in immediate moment-to-moment interpretation has been demonstrated only recently, in studies examining temporarily ambiguous sentences. One such type of ambiguous sentence is shown in (1).

(1) Amanda believed the senator…

a. …during the speech.

b. …was lying to the committee.

In (1), the relationship between the phrase *the senator* and the rest of the sentence is temporarily ambiguous: If the sentence continues as in (1a), then it is the senator who Amanda believes. If instead the sentence continues as in (1b), then *the senator* turns out to be the subject of the embedded sentence *the senator was lying to the committee.* The ambiguity arises because verbs like *believed* can license an object NP (the interpretation forced in (1a)) or an embedded sentence (forced in (1b)), and either possibility is acceptable prior to encountering one of the endings. Furthermore, the ambiguous region can be lengthened considerably (e.g., *Amanda believed the senator from the glorious Commonwealth of Massachusetts was lying...*), and thus the sentence processor must make some decision about a preferred interpretation during the ambiguous region.

One way of making a decision about how to interpret the ambiguity is to make use of some detailed information associated with the verb *believed*: In addition to knowing that it permits either an object NP or an embedded sentence, comprehenders might keep track of the relative frequency of these alternatives[6,7]. The relative frequency can be determined by having subjects write whole-sentence completions for fragments like *Amanda believed...*[26,27]. *Believed* is used more than three times as often with an embedded sentence (e.g., (1b)) as with an object NP (e.g., (1a)), and thus if the sentence comprehension mechanism relies on this kind of information, it will tend to prefer the embedded sentence alternative for the verb *believed*. Verbs with relative frequencies favoring the object NP alternative (e.g., *understood*, which licenses an object NP roughly eight times more often than an embedded sentence) would create the opposite preference—in favor of the object NP alternative. Comprehenders are sensitive to these relative frequencies, showing little or no difficulty in reading continuations like (1b) when the verb licensing the ambiguity more often takes an embedded sentence (e.g., *believed*), but showing substantial difficulty when the verb licensing the ambiguity more often occurs with an object NP (e.g., *understood*)[26,27].

Similar lexically-based frequency constraints have been shown to influence preferences in a variety of other ambiguities as well, examples of which are shown in (2) and (3). In (2), the verb *examined* can be the head of the main VP of the sentence as is required by (2a), but *examined* might instead initiate a relative clause modifying *the defendant*, as in (2b), where the main VP is *had always declared...*. Preferences in this ambiguity and comprehension difficulty in resolving it are predicted by a set of interrelated lexical properties of the triggering verb (*examined*), including how often it is used with an object NP and how often it is used in the past tense (required in (2a)) versus as a past participle (required in (2b))[6,7,28,29].

(2) The defendant examined...

  a. ...the confession and then declared his innocence.

  b. ...by the lawyer had always declared his innocence.

In (3), the PP *with the binoculars* can modify the seeing event (so that the spy is using them), but it might instead modify *the cop* (so that the cop has the binoculars rather than the spy). Preferences in this ambiguity have been argued to depend on the verb's preference to take PPs headed by particular prepositions as arguments[30], as well as the noun's preference to take such PPs and the preposition's preference to assign the particular interpretation required in each alternative[6,31].

(3) The spy saw the cop with the binoculars.

Contextual constraints

In addition to constraints associated with specific lexical representations, information which can be computed from the sentence or from the discourse context can also immediately influence sentence comprehension. This is particularly apparent in ambiguity resolution, where a variety of studies have examined the effects of contextual influences such as plausibility and referential complexity.

In the case of plausibility, most studies have examined the effect of making one alternative for an ambiguity implausible, typically by manipulating the semantic relationships between verbs and their arguments within a clause or sentence. For example, (4) contains the same temporary ambiguity as (2). The implausibility of evidence examining something leads to a preference for the relative clause interpretation[7,32–34].

(4) The evidence examined by the lawyer had always been considered circumstantial.

Plausibility influences in comprehension have also been shown with respect to the object NP versus embedded sentence ambiguity in (1)[27], the PP attachment ambiguity in (3)[35], the gerund/present-participle ambiguity in (5)[36], and filler-gap ambiguities[37].

(5) Visiting relatives...

a. ...are fun.

b. ...is fun.

Referential contexts have also been shown to influence many of the same ambiguities, in situations in which the alternatives for an ambiguity require reference to single versus multiple referents. For example, modification of a noun by a relative clause or PP typically involves a presupposition that the noun has multiple referents in the discourse, from which the relative clause or PP selects. Thus in the ambiguity in (2), the relative clause alternative (2b) requires a presupposition that the matrix VP alternative (2a) does not, making the relative clause alternative infelicitous. However, if the prior discourse context contains the presupposed multiple referents, the relative clause alternative may be preferred[11,38,39].

Referential context effects have also been demonstrated for PP attachment using preceding linguistic context[11,30], and using concurrent visual contexts[40]. In the latter case, for example, comprehenders heard the instruction in (6), which contains a temporary ambiguity involving the attachment of the PP *on the towel*.

(6) Put the apple on the towel in the box.

The participants in the experiment showed no difficulty in interpreting the ambiguity correctly when the instruction was presented concurrently with a display containing two referents for *apple*, one on a towel and one on a napkin. When only a single referent was visible, however, they had significantly more difficulty interpreting the instruction.

Computational resource constraints: Locality

An important constraint affecting sentence comprehension which is not reducible to lexical or contextual constraints is *locality*[8,14,41–44]. In ambiguous structures, the locality constraint causes a preference for an interpretation associated with a local attachment over an interpretation associated with a less local attachment. For example, locality explains the existence of the strongly preferred interpretation of (7), in which the adverbial *yesterday* is associated to the clause *the suspect left the country*, rather than to the clause headed by *told*:

(7) The bartender told the detective that the suspect left the country yesterday.

In addition to applying in ambiguity resolution, locality is also an important factor in determining the processing difficulty of unambiguous structures. For example, in English a relative clause whose relative pronoun is coindexed with a gap in object position, such as (8a), is more complex on a variety of measures than a relative clause whose relative pronoun is coindexed with a gap in subject position, as in (8b)[45–47]. Furthermore, aphasic stroke patients cannot reliably answer comprehension questions about object-RCs, although they perform well on subject-RCs[48].

(8) a. The reporter$_i$ who$_i$ the senator attacked e_i admitted the error.

b. The reporter$_i$ who$_i$ e_i attacked the senator admitted the error.

This difference can be explained by locality considerations: the distance between the relative pronoun *who* and the gap e_i is longer with an object-gap, where the NP *the senator* intervenes, than with a subject-gap, where the two are adjacent. Locality constrains two central components of sentence comprehension:[8] 1) structural integration, consisting of integrating new input words into the currently existing syntactic and discourse structures; and 2) storage of the syntactic categories that are necessary to complete the current input string as a grammatical sentence. Locality constrains these components as follows: 1) the greater the distance between an incoming word and the head or dependent to which it attaches, the greater the integration cost; and 2) the longer a predicted category must be kept in memory before being encountered, the greater is the cost for maintaining that prediction. Locality is computed in terms of new *discourse referents*: phrases representing objects and events in the world that can later be referred to using a referential expression, such as a pronoun. Locality accounts for preferences in ambiguous attachment structures such as (7) under the assumption that the sentence comprehension mechanism attempts to minimize the integration cost at each step: Integrating *yesterday* to the VP headed by *told* involves crossing structures for the intervening material *that the suspect left the country*, whereas integrating *yesterday* to the VP headed by *left* involves crossing only the NP *the country*.

The severe complexity of *nested* structures[17,41,49–51] is also explained by locality[8]. A syntactic category A is said to be nested within another category B if B contains A, a constituent to the left of A, and a constituent to the right of A. The difficulty of processing nested clauses is illustrated in (9):

(9) a. [The scientist collaborated with the professor [who$_i$ e_i had advised the student [who$_j$ e_j copied the article]]].

b. [The student [who$_j$ the professor [who$_i$ the scientist collaborated with e_i] had advised e_j] copied the article].

(9a) contains no nested clauses and is relatively easy to understand. On the other hand, in (9b), the relative clause *who the scientist collaborated with* is nested within the embedded clause *the professor ... had advised*, and this clause is nested within the outer clause *the student ... copied the article*. The doubly nested structure of (9b) makes this sentence very difficult to understand.

According to the locality-based proposal, the processing difficulty associated with nested structures has to do with the distance between heads and their dependents. In non-nested structures, heads are close to their dependents, often adjacent to one another, whereas some heads and dependents are widely separated in nested structures. For example, in the non-nested structure in (9a), the relative pronouns and their associated gaps are adjacent. In contrast, the relative pronouns and their associated gaps are widely separated in the nested structure in (9b), especially for the outer pair. This difference in head-dependent distance leads to both integration and memory cost differences in the processing of these two structures.

Interestingly, doubly-nested structures like (9b) are easier to understand when a first- or second-person pronoun is in the subject position of the most embedded clause (e.g., *I* or *you* replacing *the scientist* in (9b))[8,51]. The discourse-basis for the computation of locality explains this result. Because all discourses implicitly include a speaker / writer and a hearer / reader, first- and second-person pronouns are old referents in every discourse. Thus the distance between the relative pronouns and their associated gaps is less when *I* replaces *the scientist* in (9b), and the structure is correspondingly less complex. Complexity rating experiments across a range of nested constructions in different languages support the predictions of the locality-based theory[8,52].

The memory cost component of the theory also provides an account of comprehension effects involving locally ambiguous filler-gap dependencies. It has been observed in a variety of languages that, given a filler, people prefer to posit a gap for the filler as soon as possible[53–55]. This preference follows from the memory cost component of the theory: By postulating a gap as soon as possible, the processor minimizes the number of syntactic predictions that it needs to store. In addition, the position in which a gap is posited is modulated by plausibility as well as the lexical properties of the licensing head (e.g., a verb)[37].

Locality in ambiguity resolution interacts with a second factor. This can be seen in the processing of Spanish structures like (10)[56,57]:

(10) El astrónomo predijo la órbita del planeta [que se observó desde el satélite].
"The astronomer predicted the orbit of the planet [that was observed from the satellite]."

Spanish comprehenders read the relative clause *que se observó desde el satélite* faster when it is disambiguated towards the non-local site *órbita* than when it is disambiguated towards the local site *planeta*[57], thus demonstrating an anti-locality bias in this construction. One account of this result is that a second factor competes with locality, *predicate proximity*, which favors attachments which are structurally closer to verbs[44]. Because *órbita* is structurally closer to a verb (*predijo*) than *planeta*, predicate proximity competes with locality in this construction, and predicate proximity wins. Evidence for the two-factor hypothesis comes from the construction in (11), which is formed from (10) by adding a third potential NP attachment site *cambio*[57]:

(11) El astrónomo predijo el cambio de la órbita del planeta [que se observó desde el satélite].
"The astronomer predicted the change of the orbit of the planet that was observed from the satellite."

Because attachments to non-local sites become difficult with increasing distance, locality weighs strongly against the attachment of the RC to the first site *cambio*. The higher locality cost outweighs the competing non-local attachment factor, so that local attachment (to *planeta*) is preferred. Attachment to the first site is the next easiest, because the non-local attachment factor favors this attachment. Attachment to the intermediate site (*órbita*) is not favored by any constraints, and it is correspondingly the most difficult attachment to make[44,57]. Furthermore, locality and the second factor cannot be lexically-based, because both the two- and three-site ambiguities involve the same attaching phrase (an RC) as well as two of the same potential attachment sites immediately preceding it, yet one site is favored in the two-site ambiguity (*órbita* in (10)) while the other site is favored in the three-site ambiguity (*planeta* in (11)).

Interestingly, the non-local attachment preference observed in the Spanish two-site ambiguity in (10) is not universally present across languages. In particular, English displays a local attachment bias in corresponding English items[56]. To account for the cross-linguistic variability, it has been proposed that the cost associated with violating the non-local attachment factor varies across languages[44,58]. However, the nature of the factor favoring non-local attachment in these constructions is uncertain, as is the source of these cross-linguistic differences.

Phrase-level contingent frequency constraints

An alternative possible explanation of cross-linguistic and cross-structure preference differences relies on keeping track of the frequencies of ambiguity resolutions contingent on different lexical and phrase structure environments[59–63]. This contingent frequency approach differs from the use of purely lexical constraints in that the frequencies involved are assumed to be tabulated over syntactic constructions rather than individual lexical items (e.g., the NP Prep NP RC construction in (10), and the NP Prep NP Prep NP RC construction in (11)). However, there is evidence that ambiguity resolution frequencies in corpora do not always match comprehension preferences[64]. If the corpora that have been analyzed thus far are representative of the kinds of corpora that people are normally exposed to, then the contingent frequency approach may have difficulty in accounting for such a discrepancy.

The strongest evidence for the need to keep track of contingent frequencies of phrase structures in sentence comprehension is provided by a set of experiments involving the grammatical category ambiguity of the word *that*[61,62]. For example, in (12a) and (13a) the word *that* is a demonstrative article, modifying the noun *hotel.* In (12b) and (13b) on the other hand, the word *that* is a complementizer, introducing an embedded sentence (*cheap hotels were clean...*).

(12) a. That cheap hotel was clean and comfortable to our surprise.
b. That cheap hotels were clean and comfortable surprised us.

(13) a. The lawyer insisted that cheap hotel was clean and comfortable.
b. The lawyer insisted that cheap hotels were clean and comfortable.

In sentence-initial contexts like (12), there is a preference to resolve the ambiguity in favor of the article interpretation (12a), whereas in post-verbal contexts like (13), there is a preference to resolve the ambiguity in favor of the complementizer interpretation (13b)[61,62]. Keeping track of contingent frequencies is one way to account for these results, but there are others. In particular, memory resource factors may favor the article interpretation in the sentence-initial phrasal context[8], but not differentiate the two in the post-verbal context, so that the higher frequency of the complementizer interpretation (lexical frequency independent of phrase-level context) may determine the preference in this environment.

Conclusion

We have discussed four types of constraints which are operative during sentence comprehension: lexical constraints, contextual constraints, a locality-based computational resource constraint, and phrase-level contingent frequency constraints. We discussed a variety of evidence for the first three of these but noted that the status of contingent frequency constraints is less clear. These sources of information, in combination with phrase-formation constraints and (in speech) prosodic constraints, apply rapidly during normal sentence comprehension to determine the interpretation(s) for an incoming string of words. When the incoming string is consistent with only a single interpretation, the constraints determine the difficulty of maintaining that interpretation. When the incoming string is consistent with multiple interpretations (in cases of ambiguity), the constraints also determine the relative preferences for the different interpretations. We propose that this set of constraints will be sufficient to explain sentence comprehension results.

Outstanding questions

- What is the relative timing of the constraints? This question has often been conflated with that of *modularity*[65], so that in some models a delay in the influence of lexical and contextual constraints is justified in terms of architectural modularity[66–68]. However, even if the architecture of the system permits free interaction among information sources[6,13,61,69], as we assume, the timing of the availability of information and the application of constraints needs to be specified. Evidence from speed-accuracy tradeoff studies suggests that phrase-formation constraints may apply more rapidly than some kinds of lexical and contextual constraints[70], but more work on this issue is needed.

- What is the relative strength of the constraints? Some evidence exists that lexical constraints are stronger than contextual constraints[6,27,29,30] and that lexical constraints are stronger than the locality constraint in certain circumstances[37]. However, it remains unclear whether these are general properties of the different constraints or whether they are specific to the cases examined. In addition, it remains to be seen how phrase-formation constraints are weighted with respect to the other constraints discussed here[71].

- What is the source of individual differences in sentence comprehension? Differences in working memory capacity[8,47,72,73], in processing efficiency[34,72], and in amount of exposure to language[34,74] have all been considered. Similarly, what is the relationship between individual differences in language comprehension and in other cognitive systems[72–74]?

- What is the nature of the representations underlying the constraints? For example, in language production, the existence of phrase-level priming phenomena[75] has been used to argue for the existence of phrase-level representations. There is more limited evidence for the existence of phrase-level priming in comprehension[76].

- Are phrase-level contingent frequency constraints necessary to explain comprehension performance, or are the remaining types of constraints sufficient? If phrase-level contingent frequency constraints are necessary, can they subsume the effects of other constraints (e.g., locality)?

- How is distance determined with respect to locality? New discourse referents seem to be an important contributing factor[8], but whether there are additional components (e.g., other discourse factors, intervening words) is still unclear. In addition, the specific function relating these factors to difficulty remains to be determined.
- What is the source of cross-linguistic attachment preference differences? Differences in corpus frequencies[59], in word order[44,66], and in lexical pronoun frequencies[58] have each been proposed as possible explanations.

Box 1: The hierarchical structure of sentences

Insert Figure 1 about here

The meaning of a sentence is formed from a combination of the meanings of smaller parts of the sentence called *constituents* or *phrases*. The meaning of each phrase is itself determined by a combination of the meanings of its constituents, down to the word level and below. A phrase structure for the sentence *The reporter who the senator attacked defended the report on the web* is provided in Figure 1, with phrases shown down to the word level. Each phrase contains a particular word-level category, called the *head* of the phrase, which is obligatorily present in all instances of that phrase type. For example, the head of a verb phrase (VP) is a verb, the head of a noun phrase (NP) is a noun, etc. Thus, the head of the NP *the report* is the noun *report* and the head of the VP *defended the report on the web* is the verb *defended*. All other categories are referred to as *dependents* of the head of the phrase which immediately dominates them. For example, the article *the* is a dependent of the noun *report* in the NP *the report*. The NP *the report* and the PP *on the web* are dependents of the verb *defended* in the phrase structure in Figure 1.

The sentence (S) in Figure 1 is composed of two phrases at the top level: 1) an NP *the reporter who the senator attacked*; and 2) a VP *defended the report on the web*. This initial NP is referred to as the subject of the verb *defended*, and it is interpreted as the individual performing the action specified by the verb. The VP is further broken down into three constituents: 1) a verb *defended*; 2) an NP immediately following the verb (the verb's object) *the report*; and 3) the prepositional phrase (PP) *on the web*. The object NP is interpreted as the entity on which the action is performed, and the PP is interpreted as the location where the action takes place. This division represents one of two possible interpretations of the VP, one in which the act of defending the report took place on the web. The NP *the report* is divided into two further constituents: an article *the* and a noun *report*. The PP is also divided into two constituents: a preposition *on* and an NP *the web*.

Under a second interpretation of the VP, the report was on the web, and the reporter defended this report using an unstated medium (speaking, perhaps). The constituent structure for this alternative interpretation of the VP would consist of a verb and an NP, with the NP being further subdivided into an NP *the report* and the PP *on the web*. There is said to be an *attachment ambiguity* at the point of processing the PP *on the web*: This PP can either attach as part of the VP, or as part of the NP initiated by *the report*.

The subject NP *the reporter who the senator attacked* is subdivided into two constituents: a simple NP *the reporter* and a *relative clause* (RC) *who the senator attacked*, which is a sentence modifier (S′) for the NP. The S′ is further subdivided into a relative pronoun (Rel-pro) *who*, which refers to the NP *the reporter*, and an S *the senator attacked*. The coreference between the relative pronoun *who* and the NP *the reporter* is indicated by the subscripted i on each. Within the RC, the relative pronoun *who* is referred to as a *filler* and is interpreted as the entity which is acted on by the verb *attacked*. This relationship is identified in the phrase structure by the presence of a co-indexed empty element e_i (a *gap*) in the object position of the verb.

Note that although most of the phrase structure hypotheses implicit here are standard across most syntactic and semantic theories[15,18,19,20], some of the assumptions that we make are less universally accepted. For example, the inventory of categories and their relationship to one another are debatable[18,19,21], as is the implicit claim that there exist empty-category positions mediating long-distance dependencies[21,22]. The specific assumptions that we make with respect to these controversial issues are for convenience only, so that we have a consistent notation to discuss sentence meaning.

References

1. Cutler, A. (1995). Spoken word recognition and production. In J.L. Miller & P.D. Eimas (Eds.), *Speech, language, and communication* (pp. 97-136). San Diego: Academic Press.

2. Balota, D.A. (1994). Visual word recognition: The journey from features to meaning. In M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 303-358). San Diego: Academic Press.

3. Johnson-Laird, P.N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

4. Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psych. Rev.*, 95, 163-182.

5. van Eijck, J. & Kamp, H. (1997). Representing discourse in context. In J. van Bentham & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 179-237). Amsterdam: Elsevier.

6. MacDonald, M.C., Pearlmutter, N.J. & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psych. Rev.*, 101, 676-703.

7. Trueswell, J.C., Tanenhaus, M.K. & Garnsey, S.M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *J. Mem. Lang.*, 33, 285-318.

8. Gibson, E. (in press). Linguistic complexity: Locality of syntactic dependencies. *Cognition.*

9. Altmann, G. (in press). Ambiguity in sentence processing. *Trends in Cognitive Sciences.*

10. Marslen-Wilson, W.D. & Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.

11. Altmann, G. & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.

12. MacWhinney, B., & Bates, E. (Eds.) (1989). *The crosslinguistic study of sentence processing.* Cambridge: Cambridge University Press.

13. McClelland, J.L., St. John, M. & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Lang. Cog. Proc.*, 4, SI287-335.

14. Frazier, L. (1979). On comprehending sentences: Syntactic parsing strategies. Ph.D. dissertation, Bloomington, IN: Indiana University Linguistics Club.

15. Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht, The Netherlands: Foris.

16. Ford, M., Bresnan, J. & Kaplan, R.M. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 727-796). Cambridge, MA: MIT Press.

17. Gibson, E. (1991). A computational theory of human linguistic processing: Memory limitations and processing breakdown. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA. Available as Center for Machine Translation Technical Report CMU-CMT-91-125.

18. Pollard, C. & Sag, I.A. (1994). *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.

19. Chomsky, N. (1995). *The minimalist program.* Cambridge, MA: MIT Press.

20. Chierchia, G. & McCollell-Ginet, S. (1991). *Meaning and grammar.* Cambridge, MA: MIT Press.

21. Steedman, M. (1996). *Surface structure and interpretation.* Cambridge, MA: MIT Press.

22. Pickering, M. & Barry, G. (1991). Sentence processing without empty categories. *Lang. Cog. Proc.*, 6, 229-259.

23. Grice, P. (1975). Logic and conversation. In P. Cole and J.L.Morgan (Eds.), *Studies in syntax* (pp. 41-58). New York: Academic Press.

24. Warren. P. (Ed.) (1996). *Prosody and parsing. Lang. Cog. Proc.*, 11, 1-215.

25. Allbritton, D., McKoon, G. & Ratcliff, R. (1996). The reliablity of prosodic cues for resolving syntactic ambiguity. *JEP: LMC*, 22, 714-735.

26. Trueswell, J. C., Tanenhaus, M. K. & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *JEP: LMC*, 19, 528-553.

27. Garnsey, S. M., et al. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *J. Mem. Lang.*, 37, 58-93.

28. MacDonald, M.C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Lang. Cog. Proc.*, 9, 157-202.

29. Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *J. Mem. Lang.*, 35, 566-585.

30. Spivey-Knowlton, M. & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, 55, 227-267.

31. Schütze, C. & Gibson, E. (1996). Argument preference in English PP attachment. Paper presented at the Linguistic Society of America, San Diego, CA.

32. McRae, K., Ferretti, T.R. & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Lang. Cog. Proc.*, 12, 137-176.

33. Pearlmutter, N.J. & MacDonald, M.C. (1992). Plausibility and syntactic ambiguity resolution. *Proc. 14th Conf. Cog. Sci. Soc.*, pp. 498-503.

34. Pearlmutter, N.J. & MacDonald, M.C. (1995). Individual differences and probablistic constraints in syntactic ambiguity resolution. *J. Mem. Lang.*, 34, 521-542.

35. Taraban, R. & McClelland, J.R. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *J. Mem. Lang.*, 27, 597-632.

36. Tyler, L.K. & Marslen-Wilson, W.D. (1977). The on-line effects of semantic context on syntactic processing. *J. Verb. Learn. Verb. Behav.*, 16, 683-692.

37. Boland, J.E. et al. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *J. Mem. Lang.*, 34, 774-806.

38. Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural*

language processing: Psychological, computational and theoretical perspectives (pp. 320-358). Cambridge, UK: Cambridge.

39. Ni, W., Crain, S. & Shankweiler, D. (1996). Sidestepping garden-paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Lang. Cog. Proc.*, 11, 283-334.

40. Tanenhaus, M.K. et al. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.

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

42. Frazier, L., & Fodor, J.D. (1978). The Sausage Machine: A new two-stage parsing model. *Cognition*, 6, 291-325.

43. Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *J. Psych. Res.*, 23, 295-322.

44. Gibson, E. et al. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, 59, 23-59.

45. Wanner, E. & Maratsos, M. (1978). An ATN approach in comprehension. In M. Halle, J. Bresnan & G. Miller (Eds.) *Linguistic theory and psychological reality* (pp. 119-161). Cambridge, MA: MIT Press.

46. Holmes, V.M., O'Regan, J.K. (1981). Eye fixation patterns during the reading of relative clause sentences. *J. Verb. Learn. Verb. Behav.*, 20, 417-430.

47. King, J. & Just, M.A. (1991). Individual differences in syntactic processing: The role of working memory. *J. Mem. Lang.*, 30, 580-602.

48. Caramazza, A. & Zurif, E. (1976). Dissociation of algorithmic and heuristic processes in sentence comprehension: Evidence from aphasia. *Brain and Lang.*, 3, 572-582.

49. Miller, G.A. & Chomsky, N. (1963). Finitary models of language users. In R.D. Luce, R.R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology* (vol. 2), (pp. 419-491). New York: John Wiley.

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

51. Bever, T.G. (1970). The cognitive basis for linguistic structures. In J.R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-362). New York, NY: John Wiley.

52. Bach, E., Brown, C. & Marslen-Wilson, W. (1986). Crossed and nested dependencies in German and Dutch: A psycholinguistic study. *Lang. Cog. Processes*, 1, 249-262.

53. Stowe, L. (1986). Evidence for on-line gap location. *Lang. Cog. Proc.*, 1, 227-245.

54. Frazier, L. (1987). Syntactic processing evidence from Dutch. *Nat. Lang. Ling. Theory*, 5, 519-559.

55. De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian.* Dordrecht: Kluwer.

56. Cuetos, F., & Mitchell, D.C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*, 30, 73-105.

57. Gibson, E., Pearlmutter, N., & Torrens, V. (1998). Recency and lexical preferences in Spanish. Submitted manuscript.

58. Hemforth, B., Konieczny, L., & Scheepers, C. (in press). Syntactic attachment and anaphor resolution: Two sides of relative clause attachment. In M. Crocker, M. Pickering, & C. Clifton, Jr., *Architectures and mechanisms for language processing.* Cambridge: Cambridge University Press.

59. Mitchell, D.C., et al. (1996). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *J. Psych. Res.*, 24, 469-488.

60. Elman, J.L. (1991). Distributed representations, simple recurrent networks and grammatical structure. *Machine Learning*, 7, 195-225.

61. Tabor, W., Juliano, C. & Tanenhaus, M.K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Lang. Cog. Proc.*, 12, 211-272.

62. Juliano, C. & Tanenhaus, M.K. (1994). A constraint-based lexical account of the subject/object attachment preference. *J. Psych. Res.*, 23, 459-472.

63. Christiansen, M.H. & Chater, N. (in press). Toward a connectionist model of recursion in human linguistic performance. *Cog. Sci.*

64. Gibson, E., Schütze, C. & Salomon, A. (1996). The relationship between the frequency and the processing complexity of linguistic structure. *J. Psych. Res.*, 25, 59-92.

65. Fodor, J.A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

66. Frazier, L. & Clifton, C. Jr. (1996). *Construal.* Cambridge, MA: MIT Press.

67. Mitchell, D.C. (1989). Verb guidance and other lexical effects in parsing. *Lang. Cog. Proc.*, 4, SI123-154.

68. Ferreira, F. & Henderson, J.M. (1990). The use of verb information in syntactic parsing: A comparison of evidence from eye movements and word-by-word self-paced reading. *JEP: LMC*, 16, 555-568.

69. Spivey-Knowlton, M.J. & Tanenhaus, M.K. (in press). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *JEP: LMC.*

70. McElree, B., & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. *JEP: LMC*, 24, 432-460.

71. Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cog. Sci.*, 20, 137-194.

72. Just, M.A. & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psych. Rev.*, 99, 122-149.

73. Waters, G.S. & Caplan, D. (1996). The capacity theory of sentence comprehension: A critique of Just and Carpenter (1992). *Psych. Rev.*, 103, 761-772.

74. MacDonald, M.C. & Christiansen, M.H. (in press). Individual differences without working memory: A reply to Just & Carpenter and Waters & Caplan. *Psych. Rev.*

75. Bock, J.K. (1986). Syntactic persistence in language production. *Cog. Psychol.*, 18, 355-387.

76. Frazier, L. et al. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Mem. Cog.*, 12, 421-430.

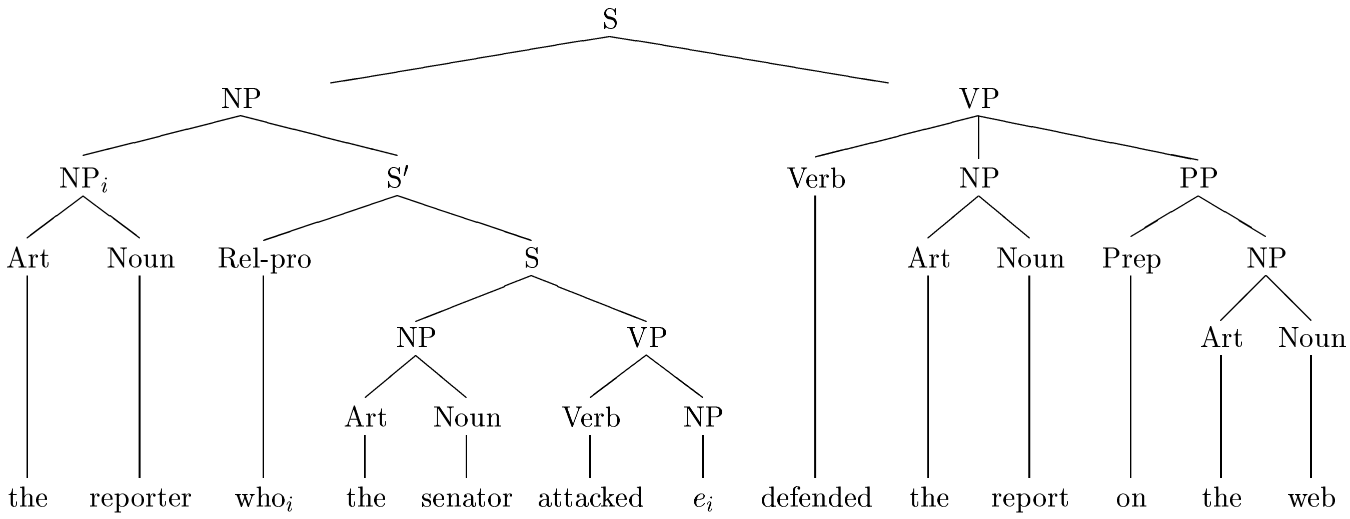

Figure 1: A phrase structure for the sentence *The reporter who the senator attacked defended the report on the web.*