Cues, constraints, and competition in sentence processing

Jeffrey L. Elman
University of California, San Diego

Mary Hare
Bowling Green State University

Ken McRae
University of Western Ontario


This work was supported by NIH grant MH6051701A2 to all three authors, NSERC grant OGP0155704 to the second author, and NFS DBSS92-09432 to the third author. We would like to thank Doug Roland for helpful discussions and comments, and for assistance with the corpus analyses.

INTRODUCTION

One of the hallmarks of human behavior is that it reflects the rapid integration of constraints from an often large number of sources, in a manner that is flexible, responsive to novel situations, and apparently effortless. This is true in virtually every sensorimotor activity in which we engage, whether it be vision, motor activity or audition, and is also particularly evident in the domain of language. There is now abundant evidence suggesting that processing of language requires the integration of information at a variety of levels: grammatical structure, meaning, discourse, world-knowledge, and so on (Bates & MacWhinney, 1989; Clifton, 1993; Jurafsky, 1996; McClelland, St. John, & Taraban, 1989; MacDonald, Pearlmutter, & Seidenberg, 1994; Spivey-Knowlton & Sedivy, 1995; Spivey-Knowlton, Trueswell & Tanenhaus, 1993; Swinney, 1979; Tanenhaus & Trueswell, 1995).

In the case of comprehension, the information a listener or reader has access to is rarely completely transparent. Rather, it is useful to think of the input as providing cues to structure and meaning. Some of these cues may be highly reliable. Others are only weakly informative. Some are frequent; others are rare. Sometimes the cues converge, but they may also conflict. Successful comprehension requires that all this information be computed, integrated, reconciled, and resolved. Furthermore, this is a dynamic process in which different interpretations may prevail at different points in time.

But how does this occur? The challenge for those who strive to understand such behaviors is that the complexities of multivariable systems—particularly those that exhibit nonlinear interactions and temporal dynamics—are formidable. Moreover, the standard computational
metaphor that has traditionally been invoked in analyses of human behavior (viz., the human brain as a digital computer) does not lend itself to modeling such systems.

Because of this, over the past several decades there have emerged a number of alternative frameworks that appear to offer more promise for understanding human behavior. One of the earliest and most influential has been the Competition Model of Bates and MacWhinney (1989; see also MacWhinney, this volume). The Competition Model (henceforth, CM) explicitly recognizes the fundamental importance in behavioral analyses of identifying the informational value of cues—distinguishing between their availability and reliability—and the potential of such cues to interact in complex ways. Although elegantly simple, the CM provides a rich conceptual framework not only for understanding behaviors in their (putative) end-state, but has also generated insights into their ontogeny as well as causes for variation, both at the level of individuals as well as across languages and cultures. One reason that the CM has had such an impact on language research is that it was one of the first models to focus on multiple cues and competition among them, countering a popular view at the time which focused instead on syntactic variables and a modular grammar.

Of course, there are many issues that remain to be resolved, and in domains such as sentence processing, significant controversies remain over the nature of the underlying mechanisms by which language is produced and processed. Indeed, one of the deepest divides in psycholinguistics concerns the degree to which sentence processing involves an autonomous component that is sensitive to only to syntactic relations and major grammatical categories, or whether the sorts of interactions that are posited by the CM (and others of its class) potentially occur at all stages in processing.

In sentence comprehension research, one of the most influential of the former theories is the two-stage serial model (sometimes also known as the “garden path” model) of Frazier and colleagues (Frazier, 1979, 1995; Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983). The two-stage theory assigns primary importance to syntactic considerations. Access to lexically-specific information, including structural information associated with specific words, detailed lexical semantics, discourse, and world knowledge, is assumed not to be available at this stage. With only impoverished information to guide the comprehender, it is assumed that she employs a set of heuristics to parse the incoming words (e.g., attach the new word to the syntactic tree in a way that minimizes number of nodes; or create an attachment that—all things being equal—is to a currently open phrase, rather than create a new phrase). These heuristics may not always work in the long run, and during a subsequent “second pass”, when additional information becomes available to the comprehender, revision of the initial parse may be necessary. However, the heuristics have the advantage that they are often correct and are quick and easy to employ.

In recent years, an alternative has emerged challenging the assumptions of the two-stage serial account of sentence processing. This contrasting approach, often termed the constraint-based (henceforth, CB) or expectation-driven model, emphasizes the probabilistic and context-sensitive aspects of sentence processing and assumes that comprehenders use idiosyncratic lexical, semantic, and pragmatic information about each incoming word to determine an initial structural analysis (Altmann, 1998, 1999; Altmann & Kamide, 1999; MacDonald, 1993; MacDonald et al., 1994; St. John & McClelland, 1990; Tanenhaus & Carlson, 1989; Trueswell, Tanenhaus, & Garnsey, 1994). On this account, a much broader range of information is assumed to be available and used at very early stages of processing. This information is often fine-grained and word-specific (McRae, Ferretti, & Amyote, 1997; McRae, Spivey-Knowlton, & Tanenhaus, 1998), may reflect lexical (MacDonald, 1994) and context-contingent frequency information (Allen & MacDonald, 1995;
Juliano & Tanenhaus, 1993), and may contribute to the immediate computation of semantic plausibility (Garnsey, Pearlmutter, Meyers, & Lotocky, 1997; Pickering & Traxler, 1998; Trueswell, Tanenhaus, & Kello, 1993).

Clearly, there is a close relationship, theoretically as well as historically, between the CB-approaches and the CM. The two approaches share core assumptions: (1) all sources of information matter; (2) these sources of information become available as soon as possible, but can differ in their time-course of availability; (3) information is used as soon as it becomes available; and (4) the influence of the various information sources depends on their strength and the relative strengths of other cues/constraints. In practice, the two approaches have tended to focus on different domains (CM on acquisition, cross-linguistic differences, and aphasia; CB on syntactic ambiguity resolution, the role of lexically-specific information in processing, and syntax-semantics interactions) and the approaches have also differed somewhat in methodologies. But there is considerable overlap, and the differences are historical more than theoretical.

One frequent criticism of CM/CB models has been that with so many free parameters and with no constraints on the nature of interactions, virtually any pattern of behavior might be explained. The criticism is not entirely fair, because an explicit goal of both CM and CB models has been to articulate precisely the conditions under which information becomes available, the factors that determine the informativeness and use of cues, and the principles that determine how multiple cues interact. Nonetheless, it is true that as CM/CB models increase in complexity, the ability to predict the fine detail of their behavior based only on first principles is severely limited (particularly when it is assumed that cue interactions involve nonlinearities). Furthermore, as the debate between two-stage and CM/CB models heated up, the role of differences in predictions regarding the time course of processing became increasingly significant. Often, experimental outcomes that differ by only a few hundred milliseconds are taken as discriminating between the two classes of models. With so much resting on such small differences, any sloppiness in the accuracy of a model’s predictions can not be tolerated.

What was needed, then was an implemented CM/CB model. Fortunately, at about the same time as the CM/CB models were being developed, there was a dramatic development on the computational front. The advent of neurally-inspired models of computation, in the form of connectionist simulations, provided a natural framework for understanding exactly the issues that the CM/CB models had pushed to the fore: What information is used in sentence processing, when it becomes available, and the nature of the (often nonlinear) interactions between information sources. These simulations furthermore made it possible to make precise predictions about the temporal dynamics of the interactions—predictions that were sometimes not what intuition might have predicted in the absence of simulation.

The family of connectionist models of language is large, and includes a variety of frameworks (see Bechtel & Abrahamsen, 2002; Christiansen & Chater, 2001; Reilly & Sharkey, 1992; and Smolensky, 2001 for recent overviews). In the remainder of this chapter, we focus on a specific architecture that comes closest in its explicit form to embodying the principles of the CM and related CB approaches: the Competition-Integration Model (McRae, Spivey-Knowlton, & Tanenhaus, 1998; McRae, Hare, & Elman, 2002; Spivey-Knowlton, 1996; Tanenhaus, Spivey-Knowlton, & Hanna, 2000). We begin by presenting the model in general terms, describe a set of empirical data involving the resolution of a syntactic ambiguity, and then show how the model implements the principles of the CM and CB framework.
THE COMPETITION-INTEGRATION MODEL

The Competition-Integration Model (Spivey-Knowlton, 1996; henceforth, CIM) had as its central principal that multiple constraints provide probabilistic support for possible syntactic alternatives. These cues thus compete in a manner that is analogous to Bates and MacWhinney's (1989) conception of a competition model. The model has matured over the years. Further developments included a mechanism for processing multiple words or regions of text, thus enabling simulations of reading time throughout the critical regions of sentences. Spivey-Knowlton and Tanenhaus (1998) incorporated an additional critical feature of the model in that they determined the strength of the various constraints using off-line norms and corpora analyses, so that these parameters were determined in a principled manner rather than being free. In both of these initial implementations, the influence of each constraint was determined by a weight connecting it to the possible interpretations. A critical issue left unresolved by these implementations, however, was constraint weighting: What determines the weight associated with a constraint? If left a free parameter, the constraint-based model itself becomes unconstrained.

Spivey-Knowlton and Tanenhaus dealt with this issue by equally weighting each constraint, but the reason for doing so was that it seemed least biased alternative. McRae et al. (1998) provided a solution to this problem, in that not only were all constraints determined by independent means, but in addition, the weight on each constraint was set by fitting the output of the model to human off-line completion norms, thus reducing the degrees of freedom in the model in a principled manner. The McRae et al. model is shown in Figure 1.

The goal of this model was to account for experimental data that had been collected in an attempt to understand how readers resolve the syntactic ambiguity that arises temporarily in sentences such as The man arrested... At the point when the verb is encountered, two possible continuations might be expected: One in which arrested is the main verb (MV) in its past tense form (The man arrested the criminal), and the other in which it is a past participle in a reduced relative (RR) construction (The man arrested by the police was found guilty).

The experimental data indicated that a number of constraints influenced readers’ preference for either the MV or RR interpretation as the sentence was read. These included the overall likelihood of such fragments being a main clause, given language-wide statistics; the relative frequencies with which specific verbs occur in their past participle versus past tense form; the constraining effect that a subsequent by has on the MV versus RR interpretation; and finally, the goodness of fit of the initial noun as a potential agent as opposed to patient of the specific verb being read (thematic fit). The data showed that readers are sensitive to the thematic fit of the noun to each of these roles, and that this influenced their ability to resolve the ambiguity, as all target sentences in this study contained the RR construction. For example, The cop arrested... favored a MV reading whereas The crook arrested... favored the RR reading.

McRae et al. (1998) used both corpus analyses (for the first three constraints) and off-line data from role/filler typicality ratings (for the thematic fit constraint) to estimate the degree to which each constraint supported each of the two interpretations. (Details of the implementation are given in the next section when we discuss a CIM implementation of another phenomenon.) Because the model could be given information incrementally, in the same way that the readers viewed successive pairs of words in the actual experiment, it was possible to measure the model’s changing interpretation of the sentence on a moment-by-moment basis. Moreover, a simple change in the model—the delay in availability of thematic-fit and lexically-specific structural information, so that only the configurational constraint operated initially—made it possible to simulate a two-stage serial account, and then contrast these predictions with those of a CM/CB model that allows...
all information to be made available and used as the words relevant to each constraint are read. Impressively, this latter condition provided a close quantitative fit to the reading time data from human subjects, whereas the two-stage serial version deviated significantly from the empirical data.

The MV/RR structural ambiguity is only one of several structural ambiguities that have been studied in the psycholinguistic literature. Another ambiguity arises when the nature of a verb’s complement (Direct Object or Sentential Complement) is temporarily unclear. As is true in the MV/RR ambiguity, this ambiguity affords a rich testing ground for probing comprehenders' sensitivity to various sources of information, and the time course with which that information is brought to bear. In the next section, we summarize previous work showing that interacting sources of information play a significant and early role in the resolution of this ambiguity, then turn to a series of studies that we have carried out that demonstrate that both discourse context and meaning have a significant effect as well. We then show how these effects are well-modeled by a CIM.

THE DIRECT OBJECT/SENTENTIAL COMPLEMENT AMBIGUITY

Some verbs occur in only one syntactic frame, for example, with only one type of possible complement (*She devoured hamburgers by the dozen; *She devoured; *She devoured that he would soon leave; etc.). However, many other verbs can be used in multiple frames (She believed John; She believed that John was telling the truth). Under certain circumstances, this fact may result in temporary ambiguity as a sentence is being read or heard. Consider, for example, the sentence The woman heard the dog had barked all night. The argument of the verb heard turns out to be a sentential complement (SC). However, if the sentence is read incrementally, at the point where the postverbal noun phrase (NP), the dog, is read, it is possible that rather than being the subject of the SC (as it turns out to be), the dog is the direct object (DO) of the verb (as it would be if the sentence ended at that point, e.g., The woman heard the dog). Thus, the role of the NP is temporarily ambiguous. Of course, the ambiguity can be avoided if the complementizer that occurs (*The woman heard that the dog… unambiguously signals that the NP is the subject of a SC). But in practice, the omission of the complementizer is common. In that case, the true structure is not revealed until the reader encounters the verb in the SC (had barked in the sentence above). This is referred to as the disambiguation region.

Resolving the DO/SC ambiguity: The role of verb bias and DO plausibility

Although some verbs occur in both SC and DO constructions with equal likelihood, many exhibit a bias toward one structure or the other. The notion that such differences (also referred to as subcategorization preferences) might play a role in sentence processing has been considered by a number of researchers (Clark & Clark, 1977; Connine, Ferreira, Jones, Clifton, & Frazier, 1984; Ferreira & McClure, 1997; Fodor, 1978; Fodor & Garrett, 1967; Ford, Bresnan, & Kaplan, 1983). Some researchers, such as Ford et al. (1983), have suggested that comprehenders use their knowledge of the relative probability with which a verb occurs with different subcategorizations to guide syntactic analysis. Others, such as Frazier (1987) and Ferreira and Henderson (1990), have claimed that lexically-specific knowledge of this sort is used only in the second stage of processing.

The role of verb bias in comprehension has been the focus of a number of studies involving the DO/SC ambiguity. Although some studies report late or no effects of verb bias (Mitchell, 1987; Ferreira & Henderson, 1990), more recent work has shown early effects (though see Kennison, 2001). Trueswell et al. (1993) contrasted sentence pairs that were structurally ambiguous at the
postverbal NP, and differed only in the bias of the main verb. These sentences were then contrasted with the reading time for structurally unambiguous versions (i.e., containing the complementizer *that*). The additional time taken to read the ambiguous version was referred to as the ambiguity effect. In a self-paced reading time experiment, Trueswell et al. found a large ambiguity effect for sentences containing DO-biased verbs at the point following the disambiguation toward a SC. In sentences involving SC-biased verbs, on the other hand, reading times for ambiguous sentences were similar to unambiguous controls.

In the Trueswell et al. study, the same sentence frame (with the same postverbal NP) was used for both a DO- and a SC-biased verb. This NP was always a plausible DO for the DO-biased verbs (e.g. *The waiter confirmed the reservation was made yesterday...*) but rarely or never plausible as DO for the SC-biased verbs (*The waiter insisted the reservation was made yesterday...*). As a result, bias was confounded with DO plausibility, and the results may have been influenced by the degree of commitment to a semantically plausible or implausible parse (cf. Pickering & Traxler, 1998).

Garnsey et al. (1997) addressed this issue in an eyetracking study in which both verb bias and DO plausibility were separately manipulated. Verbs of each bias type (including a third, equi-biased condition) appeared in sentences in which the post-verbal NP was either plausible or implausible as a DO. Thus for the implausible conditions, the NP was syntactically ambiguous, but semantically anomalous if interpreted as a DO. An effect of verb bias was found nonetheless: At the disambiguation, reading times were longer in the ambiguous than the unambiguous conditions for DO-biased, but not SC-biased verbs. In addition, with DO-biased verbs, sentences with plausible DOs yielded a significant ambiguity effect whereas the effect for those with implausible DOs was not statistically reliable. Finally, the clearest influence of plausibility was found when verbs were relatively equi-biased in terms of subcategorization preferences. Together, then these two studies offer compelling evidence that both verb bias and DO plausibility play a role in guiding the interpretation of ambiguous sentences.

**Resolving the DO/SC ambiguity: The role of verb meaning**

The studies reviewed above assume (as do many computational models) that subcategorization bias is computed across all instances of a verb. But verb bias effects, in our view, reflect the relationship between verb meaning and verb subcategorization (e.g., Fisher, Gleitman, & Gleitman, 1991). If this is true, then subcategorization bias should be sensitive to verb meaning – and thus for many verbs with multiple senses, the verb’s subcategorization bias may vary depending on which sense is intended. For example, the verb *find* must take a DO when it is used to mean ‘locate’, but when it is used to mean ‘understand’, a SC structure is more common. This follows logically from the different requirements that the two meanings impose on the arguments in each case. When one performs a concrete action (*He found the book on the table*), a patient is likely to be specified and to be realized as a DO. However, mental events or expressions of mental attitude are more typically followed by a proposition that describes the event or situation (*He found the plane had left without him*), and this is naturally realized by a SC (although note that a DO can sometimes be used as well: *He found nothing but confusion*). If the relevant relationship is between
structure and a specific sense of a verb, not structure and the verb in the aggregate, then patterns of verb bias would be better described by considering the specific sense that is used.

Linguistic research (e.g., Argaman & Pearlmutter, 2002; Grimshaw, 1979; Levin, 1993; Pesetsky, 1995; Pinker, 1989) has detailed a complex relationship between verb meaning and verb subcategorization. Thus, although earlier studies have assumed that the bias of the verb overall helps guide comprehension, we argue that subcategorization preferences are best measured with respect to specific senses of the verb. This view has been corroborated by corpus analyses (Hare, McRae, & Elman, in press; Roland & Jurafsky, 2002). Hare et al. demonstrated that a large set of individual verbs show significant differences in their subcategorization profiles across three corpora, but that cross-corpus bias estimates are much more stable when sense is taken into account. In addition, they showed that consistency between sense-specific subcategorization biases and experimenters’ classifications largely predicts results of recent experiments on the resolution of the DO/SC ambiguity.

Hare et al. (in press) and Roland and Jurafsky (2002) establish that there are systematic relations between verb sense and subcategorization preferences. Hare et al.’s results demonstrating the correspondence between sense-specific subcategorization preferences and the success of various experiments in obtaining an influence of verb subcategorization information suggested that comprehenders might learn and exploit meaning-form correlations at the level of individual verb senses, rather than the verb in the aggregate. However, these results did not directly demonstrate this.

In what follows, we present such a demonstration. We begin with a close examination of the relationships that exist between verb meaning and structure for a specific set of verbs. We find that a reliable correlation exists for these verbs in corpus analyses, and that comprehenders make use of such relationships in off-line sentence completion norms when verb sense is biased by preceding context. Next, we demonstrate that this knowledge plays a role in how readers resolve the DO/SC ambiguity in an on-line reading task. All of this sets the stage for the final section, in which we use the CIM architecture described above to simulate and understand the precise mechanism by which multiple constraints interact in order to resolve the DO/SC ambiguity.

**Sense and structure**

How might one evaluate whether or not verb sense plays a role in a reader’s expectations regarding which syntactic frame is likely (and by extension, how DO/SC ambiguities are resolved)? Two steps are involved. First, we need to know what the pattern of usage is for the various senses of a set of candidate verbs, and the syntactic structures that are typically used with each sense. This can be determined through analysis of large-scale corpora. Second, a set of experimental stimuli are needed in which, for each verb, the same initial sentence fragment may be preceded by different contexts, one which biases toward the sense associated with a DO structure, and the other which biases toward the sense associated with a SC structure. The question then becomes whether, in an on-line task (such as reading), there is evidence that comprehenders process the same sentence fragments differently, depending on the prior sense-biasing context.

---

1 In what follows we follow the standard practice of referring to different meanings of a polysemous word as *senses* of the word, to distinguish them from the distinct meanings of homonyms like *bank*. 


**Corpus analyses.** We began with 20 verbs that could occur with both DO and SC arguments, and which were categorized in WordNet (Miller, Beckwith, Fellbaum, Gross, & Miller, 1993) as having more than one sense. Typically, the DO associated sense involved a concrete event (e.g., an action) and the SC associated sense involved a more abstract one (e.g., a mental event or statement making). All sentences containing these verbs were extracted from three written and one conversational corpora: the Wall Street Journal (WSJ), Brown Corpus (BC), WSJ87/Brown Laboratory for Linguistic Information Processing (BLLIP), and Switchboard (SWBD), respectively. All corpora are available from the Linguistic Data Consortium at the University of Pennsylvania.

Because these corpora were parsed, we were able to classify each sentence according to 20 subcategorization frames expanded from the set used by Roland and Jurafsky (2002). The fine-grained parse categories were then collapsed into the more general categories of DO, SC, and Other (see Hare et al., 2003, for details). This analysis provided information about the overall frequency of occurrence with which each verb occurs in the DO, SC, and Other frames. We then computed sense-contingent bias counts to test whether these would differ from the overall probabilities. For each of the 20 verbs, we identified two senses that appear to be sufficiently distinct, that we believe are known to undergraduates, and that allow different subcategorization frames according to WordNet. We searched WordNet’s Semantic Concordance for the two senses of each target verb. The Concordance consists of a subset of the Brown corpus, with all content words tagged for sense. This allowed us to extract all sentences containing the relevant verb senses. The result of this analysis demonstrated a probabilistic relationship between a verb’s sense and its subcategorization preferences.

**Experimental results.** But are comprehenders sensitive to such differences, and does a verb’s sense affect how a comprehender will interpret the DO/SC ambiguity? To test this, we constructed a set of items based on the 20 polysemous verbs. Each item consisted of a context sentence intended to promote either the SC-biased or DO-biased sense (as determined by the corpus analyses), followed by a target sentence that contained the verb and continued with a SC. For example, the verb admit occurred in the same sentence but preceded by two different contexts (all target sentences were identical to at least one word past the postdisambiguation region):

1. **SC-biasing:**
   The intro psychology students hated having to read the assigned text because it was so boring.
   They found (that) the book was written poorly and difficult to understand.

2. **DO-biasing:**
   Allison and her friends had been searching for John Grisham's new novel for a week, but yesterday they were finally successful.
   They found (that) the book was written poorly and were annoyed that they had spent so much time trying to get it.

In (1), the context biases toward the ‘become aware of’ sense of found, which is highly associated with an SC structure; in (2) the context biases toward the DO associated ‘discover’ sense. To verify that the two contexts did in fact promote the desired senses of each verb, a series of norming studies were carried out in which subjects completed various sentence fragments (with the postverbal noun, without the postverbal noun, and with and without the prior context).
In a final study, subjects read target sentences, presented one word at a time on a computer screen, and preceded (for any given subject) by one of the two contexts. After reading each word, subjects pressed a key that revealed the next word. In order to assess the specific difficulty introduced by the ambiguity, some subjects saw versions in which the complementizer *that* was present, and others saw the ambiguous version without the complementizer. Differences in reading times between these two conditions provided a measure of the ambiguity effect.

The results are shown graphically in Figure 2. There are no effects of either context (DO versus SC biasing) or ambiguity (presence or absence of *that*) up through the first word of the disambiguating region (*was* in the example shown). At the second word of the disambiguating region (*written*), there is a significant interaction between context and ambiguity. This occurs because in the DO-biasing context, reading times are longer for ambiguous sentences than for unambiguous sentences (i.e., there is a significant ambiguity effect); but no such effect occurs in the SC-biasing context. There are also main effects for context (reading times in the SC-biasing context are shorter than in the DO-biasing context) and for ambiguity (longer reading times for ambiguous sentences than for unambiguous sentences.) But by the first word of the post-disambiguating region (*poorly*), the context by ambiguity interaction disappears, and there is no longer a significant main effect of either ambiguity or context.

Our interpretation of the context by ambiguity interaction in the disambiguating region is that the DO-biasing contexts led readers to expect the sense of the verb that typically occurred with DOs; SC-biasing contexts had exactly the opposite effect. In the SC condition, therefore, subjects were primed to expect a sentential complement, whether or not the complementizer was present. When subjects encountered the disambiguating words indicating that in fact there was a sentential complement, these were consistent with their expectations. However, in the DO condition, the absence of the complementizer reinforced the expectation that the postverbal NP was a DO. The inconsistency between this interpretation and the evidence provided by the words disambiguating towards an SC led to elevated reading times, relative to when the complementizer was present as an early warning of an SC structure.

**Modeling the DO/SC ambiguity resolution**

We have implemented a CIM to evaluate whether the experimental results of Hare et al. (2003) are consistent with the predictions one might make, given the CM/CB approaches described at the outset of this chapter. We have already described one such model; here we give a more detailed explanation of the model that we used to account for the effects that sense and other constraints have on resolving the DO/SC ambiguity. A graphical depiction of the model is shown in Figure 3. After describing the model’s architecture, we present results of its behavior when it was used to simulate the actual stimuli used in the human experiment. We then provide our interpretation of how the model’s behavior captures the principles of the CM/CB framework.

**Architecture.** As Figure 3 shows, we assumed there are at least seven constraints that interact and contribute to the relative strength of a DO versus SC interpretation in Hare et al.’s (2003) experimental items. The two octagons in the center, marked “DO” and “SC”, are metaphorical interpretation nodes that have associated activation values ranging between 0.0 and 1.0. These activations are intended to indicate the model’s confidence in a DO or SC interpretation; these interpretations change over time as new input is received and as different constraints interact.

The seven constraints themselves are shown as rectangles surrounding the interpretation nodes. Each constraint has a strength that was estimated in a manner appropriate to it. The mean constraint values across the verbs are presented in Table 1. (1) The *sense-specific*
**subcategory bias** reflects the influence of the context sentences in promoting each of the two targeted senses of each verb, where one sense is used predominantly with a DO, and the other is used predominantly with an SC. The strength of this constraint was estimated from completion norms in which subjects completed a sentence fragment that included only the subject and verb and was preceded by one of the two context sentences. The completions were scored for the sense of the verb, regardless of the actual syntactic structure the subject used, and the percentage of times that subjects used each target sense of the verb was then used as the value of the constraint. As shown in Table 1, on average, the SC-context promoted the SC-correlated verb sense for 89% of the completions, and the DO-context promoted the DO-correlated sense 76% of the time. This constraint varies by verb and by context. (2) The **sense-independent subcategorization bias** was estimated for each verb from the Brown Corpus, and reflected the verb’s overall frequency of usage—regardless of sense—with DO and SC arguments. We assume that the over-all usage of a verb might influence a comprehender’s expectations of its usage in any given context, independent of its likely sense in that context. Thus, this constraint varies by verb, but not by context. (3) The **the bias** was taken from completion norms out of context (i.e., subjects saw only fragments of the form *She admitted*). We counted the number of times a postverbal *the* was used in DO versus SC completions. Other work involving corpus analyses (Roland, Elman, & Ferreira, 2003) had indicated that the presence of a postverbal *the* (in the absence of a complementizer) is a strong predictor that the postverbal noun is a DO. Consistent with that research, the *the* bias strongly supported a DO-interpretation on average. This constraint varied by verb, but not by context. (4) The **thematic fit of noun as DO vs. SC subject** allowed the model to take into account the goodness of fit of the postverbal noun as either a DO or SC subject. For example, *goals* can be either a DO or SC subject in the sentence fragment *She realized her goals*..., whereas *shoes* is a very implausible DO and thus might be expected to promote a SC interpretation (cf. Garnsey et al., 1997; Pickering & Traxler, 1998). Thematic fit was estimated based on a combination of completion norms, WordNet semantic representations, and Resnik’s (1993) measure of selection association. Thematic fit varied by both verb and context, although Hare et al. purposely did not manipulate this variable and avoided using highly plausible or implausible NPs. (5) The **structural bias of SC auxiliary or copula** was estimated in the following manner. This word was the first in what is usually termed the disambiguation region because the words in this region provide strong structural cues that a SC is being read. The most obvious possibility is to provide full support to the SC interpretation. However, in 16 of the 20 target sentences, the first word of the disambiguation was an auxiliary or copula, which carries no semantic content, and is the short type of word that is often skipped (not fixated) in free reading (as measured by eyetracking studies). In contrast, in 15 of 20 cases, the second word of the disambiguation region was a content word such as the main verb of the SC, a predicate adjective, predicate nominal, or adverb. Because of these facts, the **structural bias of SC auxiliary or copula** provided 75% of its support to the SC, and 25% to the DO interpretation. (6) and (7) The influence of both the **structural bias of SC main verb** and the **structural bias of next word in SC** was estimated as fully supporting the SC interpretation.

**Normalized recurrence.** Constraints were integrated using a three-step normalized recurrence mechanism developed by Spivey-Knowlton (1996). First, each of the *c* informational constraints (two for fragments that ended at the verb) was condensed into its normalized probabilistic support for the *a* relevant competing alternatives (i.e., SC and DO).

\[
S_{c,a}(\text{norm}) = \frac{S_{c,a}}{\sum S_{c,a}}
\]
$S_{c,a}$ represents the activation of the $c^{th}$ constraint that is connected to the $a^{th}$ interpretation node. $S_{c,a}(\text{norm})$ is $S_{c,a}$ normalized within each constraint. Constraints were then integrated at each interpretation node via a weighted sum based on Equation 2.

$$I_a = \sum [w_c \cdot S_{c,a}(\text{norm})] \quad (2)$$

$I_a$ is the activation of the $a^{th}$ interpretation node. The weight from the $c^{th}$ constraint node to interpretation node $I_a$ is represented by $w_c$. Equation 2 was applied to each interpretation node and summed across all constraint nodes that supported it. Finally, the interpretation nodes sent positive feedback to the constraints commensurate with how responsible the constraints were for the interpretation node's activation, as in Equation 3.

$$S_{c,a} = S_{c,a}(\text{norm}) + I_a \cdot w_c \cdot S_{c,a}(\text{norm}) \quad (3)$$

These three steps (Equations 1, 2, then 3) comprised a single cycle of competition. On the basis on these equations, as competition cycles progress, the difference between the two interpretation nodes gradually increases. For example, when the verb is presented, we assume that information from the sense-specific and sense-independent subcategorization biases becomes available. The DO/SC values within a constraint are normalized so that internal to each constraint they sum to one. This normalization produces competition; for example, for the DO value to become larger, the SC value must decrease. Each of these two constraints then sends input to the interpretation nodes: This input is equal to the relevant (DO or SC) value of the constraint times the strength of the weight that connects it to the interpretation node. The value of each interpretation node is now the sum of its two current inputs. Finally, each interpretation node sends positive feedback to the sense-specific and sense-independent bias constraints; this is equal to its current activation times the strength of the relevant weight times the current relevant constraint strength. This changes the values of those constraints, so that incoming information on the next cycle of competition will differ.

**Estimating weights.** The sense-general verb bias and sense bias of context weights were determined by fitting the weights to off-line data (sentence completion norms) in the same manner as McRae et al. (1998). For simulating completions, because the basic model stops processing only when the activation of one interpretation node reaches 0 and the other reaches 1, we halted competition after various numbers of cycles and sampled the interpretation nodes' activations. Because it is not clear how to determine the number of cycles that best simulate subjects’ behavior in the completion task, we simulated completions given the context plus verb using 20 to 40 cycles, with a step size of 2 (i.e., 20, 22, 24, ... 40)$^2$. For each number of cycles, to estimate the weights for the two constraints, we varied each weight from .01 to .99 using a step size of .01. The activation of the SC interpretation node was used to estimate the proportion of human subject SC completions. For each simulation, we calculated root mean square error between the activation of the interpretation nodes and the proportion of SC completions. As in McRae et al., we averaged the weights over 110 simulations; the best-fitting ten models at each number of cycles. This process provided weights of .512 for the sense-independent subcategorization bias and .488 for the sense-specific subcategorization bias. These same weights were fixed and used in the simulation

---

$^2$ This range of cycles was chosen following McRae et al. (1998), where it was shown to produce reasonable behavior.
of the reading time experiment. Because all other constraints entered into processing one at a time at each subsequent word (e.g., the, book, was, written, poorly), the weight associated with each of them was fixed at 1.0.

**Simulating on-line reading times.** To simulate reading time data, competition continued at each word, beginning at the verb, until one of the interpretation nodes reached a criterion level of activation. All activations were retained as the initial state for competition at the following word. The criterion within a region was dynamic and was a function of the cycle of competition, according to Equation 4.

\[
dynamic \ criterion = 1 - \Delta crit \* cycle \quad (4)
\]

The constant controlling the rate of change of the dynamic criterion is represented by \( \Delta crit \). The current cycle of competition in a certain region is represented by \( cycle \). According to Equation 4, as the duration of competition in a particular region increases, the criterion for stopping competition and moving to the next region becomes more lenient. Competition necessarily terminates in a region when the dynamic criterion becomes .5 because the activation of at least one of the interpretation nodes must be greater than or equal to .5. For example, if \( \Delta crit = .005 \), the maximum number of competition cycles is \( .5/\Delta crit = .5/.005 = 100 \). A dynamic criterion is necessary for modeling reading across multiple regions of a sentence because fixation durations are partially determined by a preset timing program (Rayner & Pollatsek, 1989; Vaughan, 1983). In other words, a reader will spend only so long on a fixation before making a saccade. Presumably, this same logic holds for self-paced reading in that readers attempt to resolve competition at each one-word segment for only so long before pressing the space bar for more information. It does not make sense for readers to expect ambiguity to be fully resolved at each point in a sentence; reading processes are presumably sensitive to the fact that language contains numerous local ambiguities that are typically disambiguated by subsequent input.

In the simulation, for a specific value of \( \Delta crit \), competition began at the verb (found), where only the first two constraints, sense-independent subcategorization bias and sense-specific subcategorization bias were involved. At the, the the bias entered competition and the three weights were normalized. At the noun (book), the thematic fit of noun as DO vs. SC subject entered competition and the weights were again normalized. This continued through the third word of the disambiguation region. Because it is not entirely obvious what \( \Delta crit \) is the most appropriate, 40 simulations were conducted in which this parameter was varied from .005 to .0089 in steps of .0001. This range of \( \Delta crit \) provides reasonable behavior in the model because competition is not halted too quickly nor allowed to go on for long periods of time. The mean number of cycles of competition were mapped onto differences in reading times between the versions of the SC target sentences that included versus excluded the postverbal that, under the assumption that the versions with that provide the proper baseline against which to measure competition effects. Thus reading time differences due to processes not simulated by the model are factored out, so there is no need to incorporate variables such as word length and frequency. The activation of the DO and SC interpretation nodes were also recorded.

**Results.** The results of the model’s performance are shown in Figure 4 (left axis); to facilitate comparison with the sense-biasing experiment, human data from Figure 2 are reproduced (right axis). As is obvious, there is a close quantitative match between the human data and the performance of the model. Why does this occur?

In this framework, available constraints combine in a nonlinear fashion to produce competition among alternative interpretations. Differences between ambiguous and unambiguous
versions of sentences are considered to result from competition among constraints that support alternative interpretations. When the constraints strongly support a single interpretation, this interpretation is activated highly and there is little competition among alternatives, corresponding to the prediction of little or no ambiguity effects. At the other extreme, when the constraints are balanced among different alternatives, the activation levels of those interpretations are more equal and there is a great deal of competition. This produces large ambiguity effects. Let us consider now these effects in more detail, taking into account the sources of constraint that are active at each word, and what their net effect is. At each region, we consider the constraints that are operative and the ways in which their interactions are different in the two contexts.

At the point when the verb (e.g. found) is read, two sources of information are available to combine and influence the interpretation. The sense-independent subcategorization bias of the verb amounts to a weak transitivity bias because most of our verbs—and arguably, most in the language as a whole—tend to occur somewhat more often in DO structures. This is the same for both the DO- and the SC-biasing contexts. The sense-specific subcategorization bias of each verb in each context promoted a DO interpretation in the DO-biasing contexts, and an SC interpretation in the SC-biasing contexts. Therefore, for DO-biasing contexts, both constraints support the DO interpretation so that there is little competition at the verb and the DO interpretation node is highly activated (.87; i.e., on average, the model is 87% sure that a DO will follow these items). For SC-biasing contexts, the sense-specific bias tend to dominate, producing a small amount of competition on average with an SC interpretation being favored (SC node activation of .66; i.e., on average, the model is 66% sure an SC will follow these items).

When the is encountered, it tends to strongly support a DO interpretation for virtually all items. Therefore, for DO-biasing contexts, the model strongly tends to reinforce a DO interpretation (resulting mean DO interpretation node activation of .92), and little competition occurs. For the SC-biased items, because the the bias is contrary to the current SC interpretation, some competition results and the mean activation of the SC interpretation node is lowered (.41).

When the noun, book, is encountered, the thematic fit of noun as DO vs. SC subject becomes relevant. Because it was computed taking sense into account, it overall supports the DO interpretation for the DO-biased items and SC interpretation for the SC-biased items. This provides yet more reinforcement for the (eventually incorrect) DO interpretation for the DO-biased items, and little competition results (DO interpretation node activation of .94). On the other hand, the SC-biased items receive support for the (eventually correct) SC interpretation, and the mean activation is raised back over.5 (SC interpretation node activation of .56).

From this point forward, for all items, evidence accumulates that strongly supports an SC interpretation. Because the SC-biased items already show, on average, an SC interpretation in the model, competition is relatively minimal, properly predicting the small ambiguity effects obtained in the human data (the activation of the SC interpretation node is .63 at was, 84 at written, and 94 at poorly). However, this information is contradictory for the DO-biased items for which the model has constructed a strong DO interpretation up to the postverbal noun book (i.e., the SC interpretation was activated only at .06). The first word of the disambiguation (was, auxiliary/copula structural constraint) causes somewhat increased competition for the DO-biased items, as in the human data. However, the model continues to reflect a DO-interpretation (SC interpretation node activated only at .13). This inconsistency is intensified when the second word of the disambiguation (written) is encountered because of the presence of this additional word that is an extremely strong cue for a SC. The competition between alternative interpretations produces large competition effects in the model that are similar to the ambiguity effects found in
the human data (this is the point at which the crucial interaction occurred). Due to this competition, the activation of the SC interpretation node for DO-biased items jumps over .5 to .62 on average. Because the model has now interpreted even the DO-biased items as a SC (albeit without a great deal of confidence, on average), the next word in SC simply provides a further strong cue for a SC structure, and there is little competition even for these items (as mirrored in the lack of an ambiguity effect in the human data at poorly). By the end of competition at poorly, the activation of the SC interpretation is now .86 on average.

**CONCLUSIONS**

There is no question that the Competition Model of Bates and MacWhinney has played a pivotal role in re-thinking the mechanisms of language processing. At the time when the theory was first proposed, many psycholinguists, following the lead of theoretical linguists who argued for autonomous syntactic and semantic components in the grammar, believed that processing was highly modular, and that there were minimal interactions between processing components. On this account, interactions between meaning and structure were not to be expected.

The Competition Model challenged this assumption, and led to a program of research that revealed a far richer view of language processing in which a myriad of cues compete to constrain interpretation. To be sure, there are patterns of behavior that may primarily reflect constraints having to do with structure, and others that primarily affect meaning. But there is no firewall that prohibits interactions between constraints from different domains.

In the time that has followed, the empirical evidence in favor of this view has become increasingly compelling. A number of new accounts have been proposed in response to these data. But although the names and details many differ, the Competition Model stands as probably the earliest and most influential of such accounts. With a richer understanding of the empirical phenomena, it has also become apparent that a more precise and detailed description of the processing mechanism is desirable. This role has been filled by computational simulations, most often using connectionist networks, of the sort we have described here. These computational models make possible detailed and precise predictions about human behavior. One lesson of such models is that the verbal description of a model often does not lead to clear or reliable predictions about what behavior should be expected. The models may be conceptually simple, but their behavior is not. A second role played by such computational models is that they lay the foundation for a more formal analysis of the deeper principles that underlie processing. Finally, simulations make it possible to explore a broader space of possibilities than may have been revealed by existing empirical data, and can serve to generate hypotheses regarding human behavior that has yet to be studied. They thus lay the groundwork for future research.

What are questions that remain to be answered? With regard to the specific phenomenon we have considered here—viz., the interpretation of sentences that are temporarily ambiguous—the model suggests that the following issues should be pursued.

**Plausibility.** As we described earlier, several studies have found evidence that the plausibility of the postverbal NP as a potential DO versus SC subject may play affect how comprehenders interpret these NPs. But at least two questions can be asked.

First, what are the conditions under which plausibility matters? The empirical results are unclear. Some studies report that plausibility does play a role in resolving temporary syntactic ambiguities (e.g., Pickering & Traxler, 1998), but the effect may be complex. Garnsey et al. (1997) found that, with DO-biased verbs, ambiguity effects (i.e., difficulties in processing that
arise when the complementizer is absent) occur only when the postverbal NP is a plausible DO. Implausible DOs, on the other hand, did not generate ambiguity effects (as if the absence of the complementizer were compensated for by the implausibility of the NP). In the experiment we carried out to study the effect of verb sense on subcategorization expectations, we attempted to hold the plausibility of the NP constant. In the modeling these data, plausibility plays a role only with respect to context. But the model provides a powerful tool for asking what effect differences in plausibility might have on the DO/SC ambiguity. That is, one can easily manipulate plausibility as a variable that is crossed with the other cues, in order to generate predictions about the cue effects and cue interactions that would be predicted of human data.

A second question concerns exactly what is meant by plausibility. Although it is often invoked in the experimental literature, this is a term that remains somewhat fuzzy. Many studies use off-line data to operationalize the measure of DO plausibility. However, differences in the precise nature of the off-line tasks may yield very different estimates. For example, if one asks subjects to judge the plausibility of a sentence—in its entirety—it is possible that different subjects may focus on different aspects of the sentence. If a subject judges the sentence *The Pope deveined the shrimp* to be implausible, is this because the Pope is an unlikely agent of this activity? Or because the activity itself is rare? Or because it is unusual to devein shrimp (as opposed to something else)? The problem here is that although the sentence as a whole may in fact be implausible, the postverbal NP itself is an excellent direct object for this verb. Thus, despite the off-line characterization of this as an implausible sentence, one might expect that in an on-line reading experiment, reading times to *shrimp* would be consistent with the NP being a very plausible DO.

For these reasons, we believe that, at least in the context of the DO/SC ambiguity, it is more fruitful to think of the postverbal NP plausibility in terms of the NP’s goodness of fit to the competing thematic roles of patient or agent (realized in this case as DO or SC subject). In that case, it makes more sense to use sentence completion, role-filler judgments, or corpus data to directly assess the degree to which the NP is a likely filler of the two roles. However, this remains an area that is yet to be fully explored.

The importance of other cues? At present, the model incorporates seven cues that we believe play a role in resolving the DO/SC ambiguity. We believe these are important cues, but we do not imagine they are the only cues that are relevant. There is good reason to suspect that other cues play a role as well.

The main verb subject itself is probably one such cue. Several empirical findings lead us to this conclusion. First, we have observed from corpus analyses (Hare et al., 2003; in press) that the animacy of the subject of some verbs is highly correlated with the sense of the verb. For example, the verb *worry* can mean ‘to cause concern in’ (as in *The illness worried her mother*) as well as ‘to be concerned about’ (as in *The daughter worried her mother was sick*). When the subject is inanimate, only the first sense is possible. When the two senses are also correlated with different subcategorization frames, as is the case with *worry*, the subject animacy then might be interpreted as a cue to verb sense, which in turn will influence the interpretation of the postverbal noun when there is a DO/SC ambiguity.

Second, in a related vein, we know from experimental work studying the MV/RR ambiguity (e.g., *The man arrested...*) that comprehenders are sensitive to the thematic fit of the subject to either the agent or patient roles (McRae et al., 1998). When the subject is a good agent (e.g., *The cop arrested...*), readers tend initially to prefer a MV reading of the verb. When the subject is a good patient (*The crock arrested...*), the RR relative interpretation is preferred.
We believe that the same kinds of considerations regarding verb-specific thematic role filler expectations may interact with the sense/structure relationship that plays a role in the DO/SC ambiguity. This is because senses of the same verb often have different preferred fillers for the same thematic role. The verb *admit* has multiple senses, including ‘let in’ and ‘acknowledge.’ These two senses generate different expectations about what are likely fillers of the agent role: *doorman* is a good agent for the ‘let in’ sense, whereas *criminal* is a better agent for the ‘acknowledge’ sense. The subject may potentially provide a source of information about which sense of the verb was intended, and this in turn could lead to different expectations about whether a DO or SC will follow. If this is true, then we predict that *The doorman admitted the man had spoken*... would lead to larger ambiguity effects at *had spoken* than would *The criminal admitted the man had spoken*....

Nor does this exhaust the potential sources of cues that might be used in resolving the DO/SC ambiguity. Corpus analyses of sentences that contain DO/SC ambiguities have revealed a strong correlation between the length of the postverbal NP and a DO structure (Roland et al., 2003). Other analyses suggest that there are conventional expressions that are highly associated with specific senses of some verbs (*Mr. & Mrs. Smith proudly announce*... is highly predictive of a DO, whereas other usages predict an SC). It has yet to be demonstrated that comprehenders use such correlations as a cue to structure, and if they do, what the strength of the cues might be. This is all fertile ground for future empirical work as well as modeling.

**The final frontier: Can we learn the constraints?**

Finally, we turn to what we see as a longer-term challenge (or, more optimistically, opportunity) for the CIM: Can models be devised that are able to learn the constraints from scratch?

The goal of the model we have described, and the broader class of models like it, is to understand in detail the ways in which constraints affect outcomes. As our empirical knowledge of domains such as sentence processing has deepened, we have gained a richer appreciation of the multiplicity of the constraints that effect processing and of the complexity of their interactions. Being able to simulate these effects has become a crucial tool in developing and testing our theoretical accounts.

Identifying the constraints that apply in a domain has largely occurred through the process of hypothesis-and-test, often by carrying out off-line experiments in order to independently motivate the existence and strength of a candidate constraint. But can one imagine a model whose goal is to simultaneously identify constraints, and then to learn their effects? This is not simply a matter of trying to short-cut the modeling process; it is actually the task that confronts novices in any domain. We would argue that the two-year old who is learning language does not bring to the task the full set of constraints as an adult with many years of experience as a language user. (Obviously, we part company here with those who believe that the significant core of constraints relevant to language is in fact part of a child’s innate endowment.) Being able to model the acquisition of constraints, as well as their effect, would thus allow us to study a number of developmental phenomena that are currently not easily modeled by the current version of the CIM.

There is good reason to believe that such a goal can be achieved. To a large extent, this is precisely what has fueled the interest in learning within the connectionist community. We now know, as a result of this work, that there is much more information in the environment than may at first be obvious, and that there are relatively simple learning mechanisms that are able to extract this information. This in turn has led to a re-evaluation of the so-called ‘poverty of the stimulus’ arguments in favor of linguistic nativism (e.g., Lewis & Elman, 2002).
But another lesson of this work is that as knowledge domains increase in complexity, so too must the architecture of the mechanisms that process them. A model that is architecturally homogenous is no more likely to be able learn language than would a brain that is architecturally homogenous. It is for good reason, therefore, that models such as the CIM build complexity into their architecture at the outset. In those models, the structure of constraint knowledge is directly reflected in the structure of the models. The challenge, if one wants to model the process by which those constraints are learned, thus in part becomes the challenge of modeling the developmental process by which initially simple mechanisms architectures become more complex over time as a result of experience. We believe this can be done, and represents the next significant advance in constraint-based models. But it is, as we have said, a challenge.

REFERENCES


Hare, M., McRae, K., & Elman, J. L. (in press). Admitting that admitting verb sense into corpus analyses makes sense. *Language and Cognitive Processes*.


Figure 1. A schematic of the Competition-Integration Model. Thematic fit of the initial NP, the main clause bias, the by bias, and the verb tense/voice constraint become operative at the verb+by region. The thematic fit of the agent NP comes into play at the agent NP region and the main verb bias becomes operative at the main verb region. From McRae, Spivey-Knowlton, and Tanenhaus, (1998).

Note:

- ENTERS AT arrested by
- ENTERS AT the detective
- ENTERS AT was guilty
Figure 2. Results of the sense-biasing experiment. Ambiguity effect from the verb through the first word of the postdisambiguating region.
Figure 3. The CIM used to model the effects of sense on the DO/SC ambiguity. The (mean) strength of each constraint is shown in the table at the bottom.

<table>
<thead>
<tr>
<th>Constraint</th>
<th>SC Context</th>
<th>DO Context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SC</td>
<td>DO</td>
</tr>
<tr>
<td>sense-specific subcategorization bias</td>
<td>M = .89</td>
<td>M = .11</td>
</tr>
<tr>
<td>sense-independent subcategorization bias</td>
<td>M = .16</td>
<td>M = .45</td>
</tr>
<tr>
<td>“the”</td>
<td>M = 1.6</td>
<td>M = 14.4</td>
</tr>
<tr>
<td>thematic fit of NP</td>
<td>M = 9.0</td>
<td>M = 4.1</td>
</tr>
<tr>
<td>auxiliary/copula</td>
<td>.75</td>
<td>.25</td>
</tr>
<tr>
<td>SC verb</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>next word of SC</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 4. Comparison of human reading times from sense-biasing experiment (right axis) with model completion times (left axis).