3 Constraint-based models of sentence processing

Ken McRae and Kazunaga Matsuki

Introduction

A number of sentence comprehension models have been developed over the past 30 years or so. Many of these have been aimed at the central issue of explaining how people resolve temporary syntactic (and/or semantic) ambiguities. Because language is processed incrementally as spoken or written input is encountered over time, temporary ambiguities are rampant in natural language. A comprehender's ability to interpret sentences containing such ambiguities is an important subject of investigation, and provides insights into the comprehension system. The goal of this chapter is to review developments in constraint-based modeling. We begin by presenting some background, and then discuss advances in sentence-processing research that are due to researchers' attempts to specify the constraint-based approach. Consider the following four sentences.

(1) The horse raced past the barn fell.

(2) The landmine buried in the sand exploded.

(3) The actor forgot his lines were supposed to be spoken with a Scottish accent.

(4) The woman wished her husband was a better person.

The first two are examples of the main clause/reduced relative clause ambiguity that has been used as a tool in many studies. These sentences need to be understood as containing the less frequent and more syntactically complex reduced relative in which the initial noun phrase (*The horse* or *The landmine*) is the patient of the initial verb (i.e., they are being *raced* or *buried*), which is part of the reduced relative. Both (1) and (2) are temporarily ambiguous at the initial verb because many verbs have the same past tense and past participle inflection (as used in a main clause and in a reduced relative, respectively), and because English speakers can drop the relative pronoun and auxiliary (e.g., *that was*) prior to the initial verb.

Sentence (1) causes a great deal of difficulty because all of the initial cues point toward the incorrect main clause interpretation. Consider the moment when someone has read or heard up to *raced*. People's real-world knowledge about horses includes the fact that they often race, and thus *horse* is a great agent of *raced*. Although *raced* is ambiguous between a past tense and passive participle reading, it is usually used as a past tense verb. Therefore, comprehenders are likely to interpret the initial portion of (1) as if it will continue as a main clause, although it does not. In addition, the main clause reading carries smoothly through the prepositional phrase (PP, *past the barn*) because *raced* can be used intransitively (without a direct object, DO), and a horse racing past a barn is a plausible event. Furthermore, there is no context that contains multiple horses that might pragmatically be distinguished using a reduced relative (i.e., picking out the one that was raced past the barn). Therefore, even after *barn* is read or heard, it is very difficult to reject the main clause reading and to correctly interpret the temporarily ambiguous reduced relative. The main verb *fell* syntactically disambiguates (1) as having contained a reduced relative, but the sentence remains difficult to comprehend even at this point due to the strong constraints that all work together to cue the incorrect interpretation.

On the other hand, sentence (2) is quite easy to understand because the constraints point to a reduced relative interpretation. Because landmines do not bury things, *landmine* is a terrible agent for *buried*. Also, landmines are typically buried, and thus are a great fit as a patient. Furthermore, although *buried* is ambiguous, it is used more frequently as a passive participle than as a past tense verb. Thus, all of these cues support the reduced relative reading, even at the initial verb.

The latter two sentences are examples of the direct object/sentential complement ambiguity, another commonly used ambiguity in sentence comprehension research. In general, a sentential complement follows a verb less frequently than does a direct object. Sentences (3) and (4) are temporarily ambiguous because *that* following the verb can be dropped in English. Sentence (3) is quite difficult to understand for a few reasons. Given our knowledge of actors and what they do, we know that an actor can forget things. The potential direct object, *his lines*, is a great example of something that an actor might forget, and thus strongly supports the direct object interpretation. In addition, *forgot* is used much more often with a direct object than a sentential complement. Therefore, when a comprehender encounters *were supposed to*, which syntactically signals a sentential complement reading, difficulties arise because all of the other information prior to that point cues an interpretation in which the actor forgot his lines, rather than forgot something about his lines. In contrast, (4) is much easier than (3) because *wished* rarely takes a direct object, and is followed by a sentential complement with a high probability. Furthermore, because it makes perfect sense for a woman to wish something about *her husband*, and it makes little sense for *her husband* to be the direct object of *wished*, all of this information cues a sentential complement.

A major goal of sentence comprehension models is to explain why some temporary syntactic ambiguities cause more difficulty than do others. For a number of years, the garden-path model dominated (Frazier, 1987; Frazier & Rayner, 1982; see Chapter 2 in this volume). In this model, when each word is read or heard, a modular first stage of processing uses only the syntactic structure to that point, the major syntactic category of that word (i.e., noun, verb, etc.), and general syntactic rules to compute a single analysis. Outputting a single analysis designates it as a serial model. After the first stage is completed, the second stage (thematic processor) uses all available information (and so is not modular) to check the plausibility of the single first-stage analysis, and then to reanalyze if necessary. Van Gompel et al.'s race model (Van Gompel, Pickering, Pearson, & Liversedge, 2005; Van Gompel, Pickering, & Traxler, 2001) is also serial in that a single analysis wins a race and thus is passed on for further processing. When each word is read or heard, its major category is combined with all information regarding the previous words and discourse to produce a single candidate syntactic structure. If this analysis turns out to be incorrect, reanalysis is necessary. Finally, constraint-based models are not modular in that all available information and knowledge is used to weigh potential interpretations over time. They are also parallel in that multiple possible interpretations are entertained, and these interpretations often compete.

This chapter is organized as follows. First, we describe the general properties of constraint-based models. Next, we discuss constraints that have been hypothesized and tested, and how they have been measured. In the final section, we present various implemented constraint-based models to illustrate how constraints can be weighted and combined.

Properties of constraint-based models

The first principle underlying constraint-based models is that multiple sources of information (or "constraints") are used for comprehending sentences and resolving ambiguities. These constraints can include general syntactic biases, probabilistic lexically specific syntactic information, word meaning, selectional restrictions of verbs, knowledge of common events, contextual pragmatic biases, intonation and prosody of speech, and other types of information gleaned from intra-sentential and extra-sentential context, including both linguistic and visual contexts. This property is not unique to constraint-based models because all theories assume that all relevant information and knowledge is used eventually to interpret language.

In constraint-based models, it is assumed that there is little or no delay in information availability. This distinguishes them from the garden-path and race model. Computing some types of information might conceivably take longer than others, however. For example, computing information that requires conceptually combining the meanings of multiple words might take longer than information tied to a single word, such as verb subcategorization preferences.

The third property is that there is no delay in information usage once it becomes available. That is, once a constraint is computed or accessed, it is used immediately for comprehension. Therefore, there is no time during processing when only the major syntactic category of the current word is available.

Fourth, multiple potential alternative interpretations are activated probabilistically in parallel. An alternative way to state this is that, at any given moment, comprehenders activate (or construct) multiple relevant interpretations of the given sentential input, and these are weighted probabilistically. This contrasts with the serial models discussed above.

Finally, many constraint-based models include anticipation or expectation of structure and content. Elman's (1990) simple recurrent network, and models based on that architecture, are the clearest examples of this.

Development of constraint-based models

During the late 1980s and early 1990s, researchers began using constraint-based models to account for experimental data and to motivate psycholinguistic experiments, in large part by contrasting them with the garden-path model. At that time, constraint-based models essentially consisted of the statement that all types of contextual information that are relevant to interpretation are used rapidly. This is, of course, a vague theoretical stance. As Tanenhaus and Trueswell (1995) stated, constraint-based models were highly underspecified in a number of ways. This was nicely summed up by MacDonald (1994), who stated, "There is little evidence available about the range of probabilistic constraints that affect ambiguity resolution, the relative strength of these constraints, or how they interact with one another" (p. 160). As a consequence of this underspecification, constraint-based models were criticized for being unfalsifiable (Frazier, 1995). If a theory corresponds to stating that all types of information matter, that information types are differentially manipulated in any specific experiment, and that constraints can be weighted differently, then the theory is unduly malleable, and such criticism is justifiable. On the positive side, valid criticism often results in progress, and it did in this case. To address underspecification, and to move the field forward, a major challenge was to make constraint-based models more specific in multiple ways. Below are four major issues that needed to be addressed, and on which progress has been made.

1. What constraints are relevant to particular contexts, ambiguities, and sentences? For understanding and interpreting experimental results, what constraints were manipulated in a particular experiment, and what other constraints matter, regardless of whether or not they were purposely manipulated?
2. What are the values or strengths of the relevant constraints? Answering this question demands careful, valid measurement. Part of the issue concerns how best to quantify constraints, including over what elements they are best conditionalized (grain size).

3. What are the relative weights on the constraints in particular contexts/sentences? That is, how does a researcher specify how strongly each constraint influences potential interpretations? Do weights differ by the linguistic environment stipulated by a specific context and construction? How do they differ over time as a sentence is processed incrementally?
4. How are constraints combined? What mechanism can be used to combine the influence of multiple constraints? Should this mechanism combine constraints in a linear or a non-linear fashion?

The constraints

A major aspect of elucidating constraint-based theory involves identifying important constraints. Over approximately the past 25 years, progress has been made in that a substantial number of constraints have been identified and tested. The Appendix lists many of them. These constraints cover a large range of information types and vary along numerous dimensions. For example, some are syntactic (subcategorization preferences), whereas others are semantic or pragmatic (referential pragmatics). Some are tied to single words (transitivity) whereas others are conditionalized over combinations of words (syntactic probabilities given *verbed by*; event-specific thematic fit). Finally, some are based on the physical linguistic signal (prosody of speech) whereas others are not linguistic at all (aspects of the visual environment accompanying an utterance).

At a general level, there are global syntactic biases, such as the subject–verb–object (SVO) bias in English (Bever, 1970). Such global biases, which take into account only the major syntactic category of each word, can be viewed as corresponding to, for example, the principle of minimal attachment (i.e., build the syntactically simplest possible structure consistent with the sentence fragment up to that point; Rayner, Carlson, & Frazier, 1983). In constraint-based models, however, such biases are probabilistic, rather than binary principles.

A number of constraints are tied to single words, partly due to well-developed theories of lexical representations, and partly due to the relative ease of identifying them. A central aspect of constraint-based models has been the constraint-based lexicalist account (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, 1996). In this approach, syntactic ambiguity resolution is similar to, and depends crucially on, lexical ambiguity resolution (i.e., resolving ambiguities inherent to single words). This view has spurred a great deal of investigation into lexically specific syntactic biases (Carlson & Tanenhaus, 1988; MacDonald et al., 1994; Trueswell, Tanenhaus, & Kello, 1993). One is a verb's bias toward being used in its past tense versus passive participle form. For example, Trueswell (1996) showed that the relative frequency with which an ambiguous verb is used as a past tense verb versus a past participle influences resolution of the main clause/reduced relative ambiguity (e.g., *searched* has low passive participle bias whereas *selected* has high passive participle bias).

Another well-studied issue concerns people's knowledge of the probability that a specific verb is followed by alternative structures. Many verbs can be used in multiple structures, and verbs vary in terms of the relative frequencies with which they appear in different structures. For example, verbs such as *insist* are never used transitively (with a direct object) and are often followed by a sentential complement (*She insisted she was right.*), whereas verbs such as *confirm* are strongly biased toward being followed by a direct object (*She confirmed her reservation.*). Some studies of verb subcategorization preferences have illustrated their rapid influence on resolving the direct object/sentential complement ambiguity (Garnsey, Pearlmutter, Meyers, & Lotocky, 1997; Trueswell et al., 1993), whereas other have found null or delayed effects (Ferreira & Henderson, 1990; Kennison, 2001). Another subcategorization bias involves verb transitivity, i.e., whether or not a verb takes a direct object. Several studies have shown that these verb biases are used rapidly to resolve ambiguity during reading (Garnsey et al., 1997; MacDonald, 1994; Tanenhaus, Boland, Garnsey, & Carlson, 1989; Trueswell et al., 1993; Staub, 2007; but cf. Mitchell, 1987).

In some cases, the manner in which a constraint is viewed has been refined in systematic ways. For example, verbs such as *admit* have multiple senses. Roland and Jurafsky (2002) noted that subcategorization biases can differ by verb sense. In the case of *admit*, the "let in" sense is strongly direct object biased, whereas the "acknowledge" mental process sense is biased toward being followed by a sentential complement. Hare, McRae, and Elman (2003) found that such sense-contingent syntactic biases rapidly influence direct object/sentential complement ambiguity resolution. In addition, Hare, McRae, and Elman (2004) showed that conditionalizing subcategorization biases over verb sense reconciled previously conflicting results (rapid effects in Trueswell et al., 1993, and Garnsey et al., 1997, versus null effects in Ferreira & Henderson, 1990, and Kennison, 2001).

Structural constraints also have been conditionalized over combinations of words. For example, Spivey-Knowlton and Sedivy (1995) measured the probability that a *with*-prepositional phrase (PP) includes a noun that modifies either a verb (*The fireman smashed down a door with an axe*.) or a noun (*The teacher despised the student with the bad attitude*.). Using the Brown corpus (Kuçera & Francis, 1967), they measured overall attachment bias of a verb followed by a *with*-PP, the biases of both action and psych verbs followed by a *with*-PP, and the same biases when the direct object noun phrase (NP) contained either a definite (*the*) or indefinite (*a*) determiner. They found that *with*-PPs have a moderate overall bias toward verb phrase-attachment (VP-attachment), which is largely due to the overwhelming bias toward VP-attachment when the direct object is a definite NP. When the direct object is an indefinite NP, there is actually a moderate statistical bias toward NP-attachment. Moreover, the bulk of that pattern is due to action verbs (which are frequent) rather than verbs of perception (which are comparatively rare). In Spivey-Knowlton and Sedivy's corpus analysis, action verbs showed a bias toward VP-attachment even when the direct object was an indefinite NP, and verbs of perception showed a bias toward NP-attachment even when the direct object was a definite NP. They presented self-paced reading results demonstrating people's sensitivity to these contingencies.

Constraints involving thematic roles of verbs have played a major role in multiple ways. The thematic roles that a verb assigns to its arguments, such as agent, patient, and instrument, are assumed in many theories to be stored in a verb's argument structure (Levin & Rappaport, 1986; Tanenhaus & Carlson, 1989). Research focusing on thematic roles has dealt with thematic fit, the fit between noun concepts and the potential thematic roles assigned by the verbs. As was the case with verb biases, the scope of this constraint has been refined in systematic ways. Many studies have investigated verb-general thematic fit (i.e., not tied to any specific verb) in the form of binary selectional restrictions, such as the fact that agents of actions tend to be animate (Caplan, Hildebrandt, & Waters, 1994; Clifton et al., 2003; Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). Some researchers have viewed thematic roles as verb-specific or event-specific concepts that are continuous rather than binary in nature (McRae, Ferretti, & Amyote, 1997). That is, thematic role assignment is based on people's general knowledge about the roles that specific entities and objects play in specific types of events denoted by the verbs. McRae, Spivey-Knowlton, and Tanenhaus (1998) demonstrated that the relative likelihood of a sentence-initial noun being the agent or patient of the event denoted by a verb (as in *The cop arrested* . . . versus *The crook arrested* . . .) immediately influences resolution of the main clause/reduced relative ambiguity.

Constraints that cross sentence boundaries also have been shown to influence ambiguity resolution (see Chapter 5 in this volume). Most language is understood within a broader context, both linguistic (other sentences; a conversation) and physical (the sights, sounds, and smells of a real-life context). The incremental interactive theory of sentence comprehension developed by Crain and Steedman (1985) and Altmann and Steedman (1988) focused on contextual factors, most notably, referential pragmatics. These researchers stressed that the context in which an ambiguity was embedded could strongly influence preferences. What is particularly crucial are the ways in which potential syntactic structures best match the pragmatic constraints of the discourse model. For example, one discourse function of a relative clause is to select among alternatives. If there were two actresses, and one was the favorite of the director, and the other of the producer, it would make pragmatic sense to begin a sentence with *The actress favored by the director*, and such a reduced relative is more easily interpreted (see Spivey & Tanenhaus, 1998).

Finally, although a great deal of research testing constraint-based models has been conducted with written language, some studies have used spoken language. A number of studies have shown rapid influences of visual context on ambiguity resolution (Chambers, Tanenhaus, & Magnuson, 2004; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), including developmental studies (Trueswell & Gleitman, 2004), using the visual-world paradigm in which participants' eye movements to the visual scene or objects are monitored while sentences are heard. Using this paradigm, the influence of pictures that match one or another interpretation of an utterance has been illustrated, often in terms of predicting or anticipating the structure and semantic content of sentences. These can include the number of referents, location of referents, or implied events in a scene. Other research using spoken language has demonstrated that various aspects of prosody and intonation of speech are important constraints for resolving syntactic ambiguities (Beach, 1991; Speer, Kjelgaard, & Dobroth, 1996).

In summary, the field of sentence processing has moved over the years from vague statements of the type that all constraints matter, to investigations in which many constraints have been identified and shown to rapidly influence syntactic ambiguity resolution.

Measuring constraints

In constraint-based models, the vast majority of constraints are viewed as probabilistic. Therefore, carefully measuring the relevant probabilities is critical to understanding their influence. Predicting the speed and strength with which a specific constraint affects comprehension requires valid measures of not only the constraint of primary interest, but also the other constraints at work in experimental items. This is important both for generating hypotheses and for understanding precisely why specific results were obtained. In general, rather than speculating about the influence that various constraints might have, or using intuition to explain why certain results were obtained, quantifying all relevant factors produces a deeper understanding of the empirical phenomena.

It is interesting to note the parallel between a major influence of connectionist modeling in general, and that of constraint-based models of sentence comprehension (which are often connectionist models themselves, or incorporate many of the same principles). For connectionist models, one consequence of the emphasis on learning environmental distributional statistics was a focus on the importance of quantifying relevant aspects of the environment (Daugherty & Seidenberg, 1992). The development of constraint-based models was likewise tied to quantifying linguistic and extra-linguistic constraints. Critical to this endeavor was the development of many types of norming methods designed to tap people's linguistic and world knowledge. Furthermore, advances in computational linguistics have played a vital role in constraint-based modeling, and in language research in general.

A major method for measuring constraint strength is to use human judgments or productions. In many off-line production norming tasks, participants are given sentence fragments such as *The defendant examined* and are asked to complete them. Researchers then count the frequencies of the syntactic structures that participants have produced. This type of norming provides quantitative estimates for experimental items. Given that much of the debate surrounding constraint-based versus garden-path models hinges on precisely when a constraint influences comprehension, off-line norms provide estimates of people's knowledge and its influences when they have time to consider the options. Results also can be considered to be an estimate of the degree to which people predict various structures given some context/fragment. Finally, it is possible to gauge the influence of the words in a sentence incrementally by using gated completion norms in which completions for successively longer fragments are collected (McRae et al., 1998).

There are a number of issues regarding completion norms, however. Sentence completion is a production task, so it necessarily combines comprehension of the given fragment with production of the completion. Depending on one's view of the extent to which comprehension and production rely on independent processes or representations, this may be an issue. Because shorter, less complex structures are easier and faster to produce, completion tasks may overestimate the frequency of shorter, less complex structures. Another issue is that sentence fragments sometimes combine multiple constraints. For example, participants might be given a context sentence or sentences, an initial NP, and a verb. Because it is unclear which constraint or set of interacting constraints are influencing participants' completions, such completion tasks do not necessarily provide information about the influence of specific constraints.

Potential production biases can be alleviated via rating tasks. For example, participants might be given fragments of sentences and be asked to rate the degree to which they are grammatical or plausible. One potential issue is that it can be difficult to disentangle structural influences from the influences of people's knowledge about real-world events. Each of the two judgments may be influenced by variables from the other domain. For example, two sentences describing the same event may be given different plausibility ratings because syntax leaks into the judgment.

For these reasons, particularly when measuring people's semantic or event knowledge, researchers have tapped this knowledge outside of particular sentences. For example, thematic fit ratings have been used in numerous studies (McRae et al., 1998; Pado, Crocker, & Keller, 2009). People are asked to rate, for example, how common it is for a spoon to be used for stirring. One can also ask participants to list things that people use to stir.

In corpus analyses, researchers calculate probabilities of interest by either examining parsed corpora (e.g., Penn Treebank), or randomly sampled sets from unparsed corpora (Jurafsky, 1996). The assumption is that corpora reflect subsets of inputs that comprehenders experience as language learners, and from which they acquire their knowledge regarding the distributional properties of words or phrases within certain syntactic contexts. Corpus analyses have played a crucial role in quantifying constraints for experiments and modeling.

Estimating constraint weights from corpora is not without its challenges, as discussed by Roland, Dick, and Elman (2007). There is a quality versus quantity issue, because parsed corpora tend to be smaller in size, and the choice of automatically, semi-automatically, or manually parsing corpora, or random sampling, is associated with a tradeoff between accuracy and costs. This is a potential issue when a researcher's goal is to obtain estimates that are conditionalized over multiple items. For example, less familiar verbs occur rarely in corpora, and if the goal is to estimate their occurrence in various structures in specific environments, estimates will be based on sparse data. Finally, corpora vary in genre and thus their content; compare Wall Street Journal corpora to those based on internet postings. Corpora also differ in register (from relatively free spoken conversation to heavily edited written newspapers or books). Researchers such as Gahl, Jurafsky, and Roland (2004) have shown that such differences can lead to systematic cross-corpus variability.

In summary, a great deal of progress has been made on estimating the strengths of numerous constraints, and this progress has had a major positive influence on sentence processing research. Norming methodologies have played a crucial role, and the possible types of norms are limited only by one's imagination. Advances in corpus techniques have enabled greatly improved experimental predictions and modeling. Every method has associated complications, but researchers continue to develop an appreciation for precisely what is being measured and how estimates can be influenced by specific factors.

Weighting and combining constraints

Even when the relevant constraints have been identified and their strengths estimated, there remain the critical issues of how they are weighted relative to one another, and how they are combined or integrated. A great deal of debate has centered on precisely how and when various constraints influence ambiguity resolution. The simplest view is that if a researcher manipulates one constraint (e.g., thematic fit) in a way that biases readers toward the less frequent, syntactically more complex interpretation of an ambiguous sentence (e.g., the reduced relative interpretation of a sentence containing a main clause/reduced relative ambiguity), then ambiguity effects that would otherwise occur without such a manipulation should be eliminated. This view makes perfect sense at first glance, and numerous studies have used this logic. However, there are two issues.

First, the theoretical stance that all relevant constraints matter entails that predicting self-paced reading times or eye-movement latencies requires taking into account the influences of all constraints and their strengths. If, for example, a researcher manipulates a single constraint, but all other constraints strongly oppose it, it is unlikely that the manipulated constraint will have a major influence on comprehension, and almost certainly will not eliminate comprehension difficulty. Such a result could be interpreted as supporting the garden-path model, and often has been (Ferreira & Clifton, 1986; Rayner, Carlson, & Frazier, 1983). However, such a result could also be interpreted as supporting the constraint-based theory because the majority of constraints favor a specific analysis.

The second issue is that even if the manipulation of a non-syntactic constraint eliminates the ambiguity-driven difficulty in self-paced reading or early eye-tracking measures such as first fixation duration or first-pass reading times, proponents of garden-path or race models can argue that the second stage of processing (or reanalysis) kicks in with a small, but real, time delay. This logic has been used as well (Clifton & Ferreira, 1989; Frazier, 1995). Thus, the field was left in the position that there was no signature data pattern that clearly discriminated constraint-based from other models (McRae et al., 1998).

One potential solution is to use computational modeling. Computational modeling forces researchers to make explicit decisions regarding parameters and mechanisms. Implementation requires decisions regarding the constraints to incorporate, their strengths, how they are weighted, and how they are combined. Simulations can help to adjudicate among theories because the output can be compared to human performance, and that output does not always match researchers' intuitions. In general, simulations increase a theory's falsifiability. Of course, computational modeling is not a panacea because there are typically some free parameters that influence a model's performance, but the exercise of implementing a model is highly beneficial in terms of clarifying what choices have to be made, how parameters can be set, and how parameter values influence the model's behavior.

Implemented constraint-based models

In this section, we present implemented constraint-based models. A number of researchers have conducted simulations to generate and test predictions for on-line reading-time experiments. As discussed in detail by Hintzman (1991), there are a number of advantages of implemented models. These include overcoming hindsight bias ("Of course my model would account for that result."), the latitude in explanation that is available when using verbally described models, and the extreme difficulty of using intuition to predict the performance or output of a fully interacting non-linear system that changes over time.

We begin with the competition–integration model because it was the first implementation of a constraint-based model and it has been used most frequently in the literature. We then describe Tabor's dynamical systems model (Tabor, Juliano, & Tanenhaus, 1997). Finally, we present the coordinated interplay account network that has been used to simulate visual-world eye-tracking data (Mayberry, Crocker, & Knoeferle, 2009).

The competition–integration model

Spivey and colleagues (McRae et al., 1998; Spivey & Tanenhaus, 1998; Spivey-Knowlton, 1994; Spivey-Knowlton, 1996) developed a model that simulates on-line reading latency data. The competition–integration model has been implemented in a number of studies (Binder, Duffy, & Rayner, 2001; Elman, Hare, & McRae, 2004; Ferretti & McRae, 1999; Green & Mitchell, 2006; Hanna, Spivey-Knowlton, & Tanenhaus, 1996; McRae et al., 1998; Spivey &

Tanenhaus (1998); Tanenhaus, Spivey-Knowlton, & Hanna, 2000). It has been used to study the main clause/reduced relative ambiguity, direct object/sentential complement ambiguity, agentive/locative prepositional phrase ambiguity, and relative clause attachment. A number of structural, lexically syntactic, thematic, and referential constraints have been included. Thus, it has been applied reasonably widely to account for self-paced reading and eye-tracking data.

Although some details differ, there are mostly commonalities across the various implementations. The competition–integration model consists of input constraint nodes, output interpretation nodes, and weights connecting them. McRae et al.'s (1998) implementation is shown in Figure 3.1. The two hexagons in the center represent the interpretations of interest, main clause and reduced relative. With one exception (Ferretti & McRae, 1999), the competition–integration model has simulated competition between two interpretations. The activation values of interpretation nodes vary between 0 and 1 to capture the probability of, or the model's confidence in, competing interpretations, which changes over time as constraints interact and new input arrives. The model does not generate syntactic alternatives on its own. Instead, it simulates resolving a syntactic ambiguity once it is encountered. It accomplishes this by evaluating the relevant constraints, and using them to support various alternatives. Therefore, the competition–integration model was designed to simulate ambiguity effects when matched ambiguous and unambiguous sentences are compared (Tanenhaus et al., 2000).

The rectangles surrounding the hexagons in Figure 3.1 are input nodes corresponding to relevant constraints. The circles within each type of input represent the constraint values specific to one interpretation. How those values are estimated varies across constraints, and values are not necessarily in the same scale, but are treated as similar to probabilities during computations. That is, they are transformed to range from 0 to 1 and add up to 1 for each type of input. As shown in Figure 3.1, not all input nodes exert an influence from the beginning of a sentence or region; only the relevant constraints are included in computations, and inputs are added as they become applicable. In McRae et al. (1998), thematic fit of the initial NP, the main clause bias, the *by* bias, and the verb tense/voice constraint are operative at the verb+by region. 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.

Constraints are integrated using a normalized recurrence algorithm developed by Spivey-Knowlton (1996). Each cycle of competition consists of three steps:

1. Each value within each input node is divided by the sum of those values to normalize activation within each constraint.
2. The activation of each interpretation node is the sum of supporting input activations scaled by the connecting weights.
3. The input nodes receive feedback from the interpretation nodes proportional to the activation of that input in Step 2.

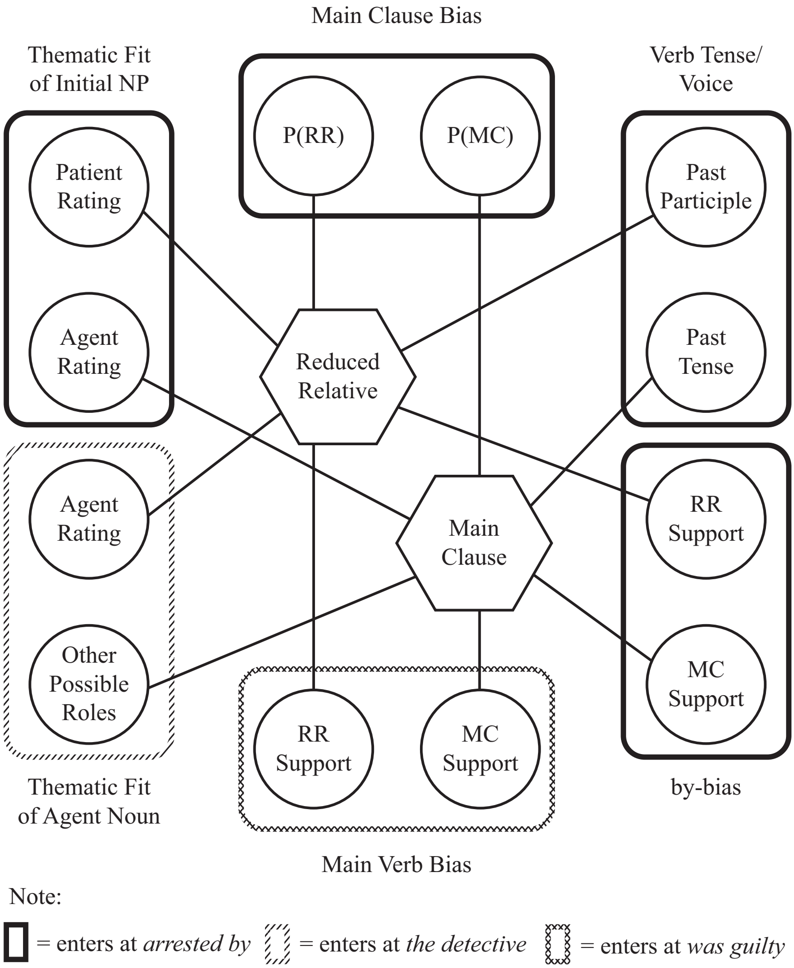

Figure 3.1 Schematic of McRae et al.'s (1998) competition–integration model.

Source: Reprinted from *Journal of Memory and Language*, *38 (3)*, K. McRae, M. J. Spivey-Knowlton, & M. K. Tanenhaus, Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension, 283–312, © 1998, with permission from Elsevier.

These three steps comprise one cycle of competition and are repeated until the activation level of one of the output nodes reaches a criterion value. The criterion changes dynamically as a function of number of cycles and a constant called Δcrit, which makes the threshold more lenient as the number of competition cycles increases, and ensures that competition terminates. This dynamic criterion is needed because fixation durations are partially determined by a preset timing program so that a reader spends only so long on a fixation before making a saccade (Rayner & Pollatsek, 1989; Vaughan, 1983). This logic holds for self-paced reading in that readers attempt to resolve competition at each segment for only so long before pressing the key for more information. Because it is not clear what value of Δcrit is the most appropriate, researchers either choose a fixed Δcrit after exploring the parameter space with multiple data sets (Spivey & Tanenhaus, 1998), or choose a range of values and average across simulations (McRae et al., 1998). These models use a Δcrit for which competition is not halted too quickly, to allow for differences among conditions to be observed. After simulating the interpretation of each sentence, the mean number of cycles is compared to the ambiguity effect found in behavioral data (e.g., the mean difference in reading time in each region between the ambiguous and unambiguous conditions). All published models have used a linear transformation to compare human ambiguity effects and the model's cycles of competition.

The model receives information incrementally. Thus, its changing interpretation can be measured on a moment-by-moment basis. When simulating ambiguity effects across multiple regions, the model takes various constraints assumed to be available at each region as its inputs, and iterates until the criterion is reached before moving on to the next region. The model then adds new constraints and associated weights, while retaining all activations from the previous region as the initial state. The constraints for the word or region currently being read are given one-half of the overall weight. Weights are normalized for each region. At the initial region, if the constraints support the interpretations relatively equally, substantial competition results. At subsequent regions, competition increases with the degree to which new constraints oppose the carried-over interpretation and constraint values.

McRae et al. (1998) simulated data from a self-paced reading study using the main clause/reduced relative ambiguity (Figure 3.2). The human data indicated that a number of constraints influenced interpretation: the main clause bias, the verb tense/voice constraint, and the *by* bias. In particular, readers were sensitive to the goodness of fit of the initial noun as a potential agent or patient of the specific verb (thematic fit). For example, *The cop arrested* . . . favored a main clause reading whereas *The crook arrested* . . . favored the reduced relative. McRae et al. used both off-line data from role/filler typicality ratings (thematic fit) and corpus analyses (for the other constraints) to estimate the degree to which each constraint supported each interpretation. They also used off-line gated sentence completion data to determine the weights. They tested a large set of weight configurations and averaged over a subset of models that yielded the smallest root mean square error values to the off-line completion proportions. With these inputs and parameters, the model provided a close quantitative fit to the reading time data. Additionally, they delayed the availability of thematic fit and lexically specific syntactic information, so that only the configural constraint (main clause bias) operated initially. The delayed version deviated significantly from the empirical data, supporting the view that all constraints immediately influence ambiguity resolution.

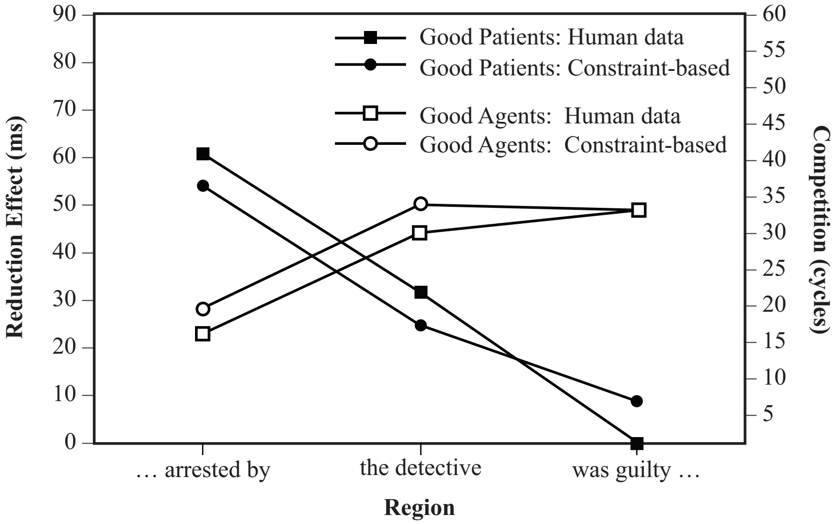

Figure 3.2 Predicting ambiguity effects (reduction effect: the difference between ambiguous reduced and unambiguous relative clauses) with the competition–integration model.

Source: Reprinted from *Journal of Memory and Language*, *38 (3)*, K. McRae, M. J. Spivey-Knowlton, & M. K. Tanenhaus, Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension, 283–312, © 1998, with permission from Elsevier.

The competition–integration model has been used to investigate other constraints and ambiguities. Spivey and Tanenhaus (1998) simulated data from several studies showing the influence of extra-sentential referential context on the main clause/reduced relative ambiguity (Murray & Liversedge, 1994; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993). They determined constraint values using off-line norms and corpora analyses. For practical reasons, Spivey and Tanenhaus used equal weights for each constraint rather than estimating them from off-line norming data. Their model simulated both eye-tracking and self-paced reading data. The competition–integration model also has been used to simulate the direct object/sentential complement ambiguity (Elman et al., 2004; Ferretti & McRae, 1999), and the agentive/locative *by*-phrase ambiguity (Hanna et al., 1996; Tanenhaus et al., 2000). Furthermore, Tanenhaus et al. (2000) present an illuminating set of simulations in which they successfully simulated data from a range of experiments that had been interpreted as evidence for both constraint-based and garden-path models.

There have been challenges to the competition–integration model. It has been claimed that the model predicts (prolonged) processing difficulty under the circumstance in which constraints are balanced between alternative interpretations because competition cannot be easily resolved under such circumstances (Traxler, Pickering, & Clifton, 1998; Van Gompel et al., 2001; Van Gompel et al., 2005). In contrast, in sentences with adjunct or relative clause modifiers, reading times were shorter when structural ambiguity was present than when it was absent. For example, Van Gompel et al. (2005) used materials like the following.

(5) I read that the governor of the province retiring after the troubles is very rich.

(6) I read that the province of the governor retiring after the troubles is very rich.

(7) I read that the bodyguard of the governor retiring after the troubles is very rich.

Sentences (5) and (6) are initially temporally ambiguous until *retiring* is encountered, but the ambiguity can be resolved using thematic fit in favor of high-attachment in (5) or low-attachment in (6). In both cases, *the governor* is, but *the province* is not, a semantically plausible subject of *retiring*. Sentence (7) is globally ambiguous because the relative clause can modify either *the bodyguard* or *the governor*. Because there is no strong bias toward one of the interpretations before or after *retiring*, Van Gompel et al. (2005) argued that competition models predict greatest competition in (7). However, participants spent less time reading the potentially disambiguating word (*retiring*) in (7) than in (5) or (6), referred to as an ambiguity advantage.

Green and Mitchell (2006) argued that Traxler et al.'s (1998) and Van Gompel et al.'s (2005) assumptions are based on a misinterpretation of the model. Green and Mitchell investigated how McRae et al.'s (1998) model behaves given various inputs. One issue concerned whether the model displays maximal competition given a new input that is balanced between two interpretations. The second issue concerned whether the model can display an ambiguity advantage.

The competition–integration model exhibits maximal competition when the bias of new inputs opposes the model's values inherited from the previous region, whereas there is less competition when new inputs are balanced. Balanced new inputs produce a high degree of competition when both the inputs and inherited activations are perfectly balanced. Green and Mitchell showed that this seldom occurs because the model amplifies early biases in a region, so that tiny biases in early phases of competition become large values by the time the model moves to the next phase of competition. They concluded that "Balanced legacies are very rare indeed" (p. 10). Second, the model displays an ambiguity advantage much like that reported by Traxler et al. and Van Gompel et al. in some circumstances. For instance, when there is an inheritied bias from previous regions toward some interpretation, which Green and Mitchell argued is extremely likely, a balanced ambiguity leads to faster processing. For Van Gompel et al.'s studies, it could be argued that the activations in the competition model, as currently implemented, would be balanced because the ambiguity begins at *retiring*. However, this may be an implementational rather than a theoretical issue. In a complete model of this type, one would assume that evidence for numerous interpretations builds throughout a sentence. Therefore, from the initial word onward, various interpretations would be favored, producing an unbalanced state when entering any subsequent region. This has not been simulated definitively, however.

In summary, the implemented competition–integration model has contributed to sentence-processing research because it forced researchers to make specific decisions regarding the four key issues defined above. Simulations enabled explicit testing of ideas regarding the relevant constraints, their strengths, how they are weighted, and how they combine. It also made it possible to simulate experiments that had been interpreted as support for two types of verbally described models, thus moving the field forward. However, there are clear limitations with current implementations. The model does not construct potential interpretations, rather, it simulates competition among assumed possible constructions. Given that simulations have included only a small number of alternatives, there are open issues regarding scalability. Furthermore, it does not compute the actual meaning of utterances, whereas computing meaning is presumably the major reason why people process language. Finally, potentially free parameters exist, such as the precise mapping of cycles of competition to human reading times.

The visitation set gravitation model

Several researchers have implemented connectionist models to capture the computational dynamics of sentence comprehension using simple recurrent networks (SRN; Elman, 1990; see Figure 3.3). An SRN consists of three feedforward layers (input, hidden, and output) as well as a context layer that stores a copy of activations of the hidden layer at the previous time step, and feeds its activation to the hidden layer during the current time step. The network can be trained on English-like sequences of words, each represented by a random vector and presented one at a time, and teaching it to predict the next word at each step. One of the interesting aspects of the model is that the activation of its hidden layer represents the "mental space" or "parse state" of the model. Analyses of hidden layer activation reveal that the patterns evoked in response to each word correspond to category membership of the words (e.g., verbs versus nouns, transitive versus intransitive verbs, animate versus inanimate nouns; Elman 1991). When the hidden layer's activation is viewed as a vector or a point in a metric space, the model forms, over the course of training, a set of clusters (the centers of which are called *attractors*) within the space where contextually (i.e., syntactically, thematically, or semantically) similar patterns are placed near one another. For a given sequence of input (a sentence), the network displays continuous movement (or a trajectory) through the multidimensional space of mental states (Elman, 1993). Thus, sentence comprehension can be characterized as the behavior of a dynamical system (Elman, 1995).

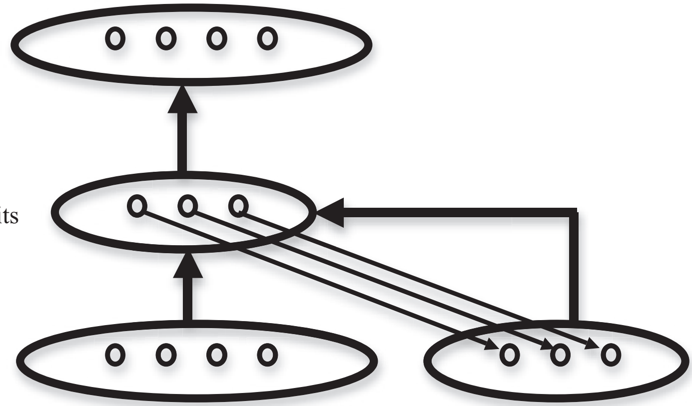

Figure 3.3 The architecture of a simple recurrent network (SRN). The input units are typically fully interconnected with the hidden units, which are then fully interconnected with the output units. The input and output units are the same. Weights from input and context units to hidden units, and from hidden units to output units are trained using the back propagation learning algorithm. The hidden unit activations are copied back to the context units.

Building on the SRN and the dynamical system metaphor, Tabor and colleagues (Tabor, Juliano, & Tanenhaus, 1997; Tabor & Tanenhaus, 2001) developed the visitation set gravitation (VSG) model to explore syntactic ambiguity resolution. In the VSG, the hidden layer representations of a trained SRN are analyzed to map representational states onto structural interpretations and to generate reading-time predictions. Gravitation time, which is the time required for a given hidden-layer state to move toward or "gravitate" into an attractor, corresponds to reading time for a word. When a sentence is ambiguous, multiple attractors pull with different gravitational strengths that are determined by their distance from the starting point. The model predicts processing disruption as a function of the relative proximity of the starting point of a trajectory to the attractors, and the relative strengths of the attractors' gravitational pull, both of which are determined by the model's experience during training. Tabor and Tanenhaus (2001) used the model to simulate the influence of thematic fit in the main clause/reduced relative ambiguity. Tabor et al. (1997) simulated the general pattern of contingent frequency effects in the determiner/complementizer ambiguity of *that*. The relative frequency with which *that* is used as a determiner (*The lawyer insisted that cheap hotel would be safe*) versus as a complementizer (*The lawyer insisted that cheap hotels would be safe*) is contingent on whether it appears before (35 percent versus 11 percent) or after the main verb of the sentence (93 percent versus 6 percent), and reading times to *that–adjective–noun* reflect this.

One important aspect of this modeling is that, unlike the competition–integration model whose inputs and weights are estimated using norms and corpora analyses, the model acquires the relevant statistics and constraints from the input corpus. This is a crucial step toward a more complete constraint-based model of language comprehension, although the SRN part of the VSG model has to date been trained on relatively simplified sets of input sentences.

The coordinated interplay account network (CIANet)

Mayberry, Crocker, and Knoeferle (2009) developed a model of situated language comprehension named CIANet, which also is based on an SRN. Unlike the models discussed so far, which focus exclusively on linguistically available constraints, an explicit aim of CIANet is to simulate the role of visual context during sentence comprehension. Before describing this model in some detail, we first review studies by Knoeferle and her colleagues (Knoeferle & Crocker, 2006; Knoeferle et al., 2005) to illustrate the type of phenomena the model was designed to simulate.

Knoeferle et al. (2005) monitored participants' eye-movements to a depicted scene while listening to sentences describing the event. For example, given a scene depicting a pirate, a princess carrying a bucket and a sponge, a fencer carrying a paint pallet and brush, and a few other common objects (not depicted in Figure 3.4), they listened to one of the following sentences:

(8) SVO sentence: Die Prinzessin wäscht offensichtlich den Pirat.
(The princess$_{\text{AMBIGUOUS}}$ washes apparently the pirate$_{\text{ACC}}$.)

(9) OVS sentence: Die Prinzessin malt offensichtlich der Fechter.
(The princess$_{\text{AMBIGUOUS}}$ paints apparently [is apparently painted by] the fencer$_{\text{NOM}}$.)

In German, the nominative and accusative case are different for masculine noun phrases such as *Pirat* (*der* and *den* respectively), but are identical for feminine noun phrases such as *Prinzessin* (*die* in both cases). Because German allows scrambling of nouns, without any context, sentences (8) and (9) are temporarily ambiguous in terms of the thematic role of the initial noun phrase (*Die Prinzessin*) until the case-marked determiner of the second noun phrase is encountered (*den* vs. *der*). However, sentence (8) can be disambiguated much earlier at the verb because the combined information in the scene and in the sentence up to the verb constrains the possible role of the mentioned referent (i.e., the princess is more likely the agent of washing because she is carrying the bucket and sponge in the picture). Eye-movement patterns indicated that participants made use of this constraint to resolve the ambiguity. More specifically, immediately after the verb was heard, participants looked more to the soon-to-be-mentioned thematic role fillers of the event denoted by the verb than to other objects in the scene (in sentence (8), *the pirate* fulfills the patient

Figure 3.4 Example visual scene from Knoeferle et al. (2005).

Source: Reprinted from *Cognition*, *95(1)*, P. Knoeferle, M. W. Crocker, C. Scheepers, & M. J. Pickering, The influence of the immediate visual context on incremental thematic role assignment: evidence from eye-movements in depicted events, 95–127, © 2004, with permission from Elsevier.

role of the washing event, whereas in sentence (9), *the fencer* is the agent of painting).

Knoeferle and Crocker (2006) used the same methodology to investigate how prior knowledge about events in general comes into play. Of particular interest is the relative importance of immediate visual context and general thematic role knowledge of verbs. Participants were shown a scene with, for example, a wizard with a monocle, and a detective with some food and a pipe, facing toward a pilot, while they listened to one of the sentences below:

(10) Den Piloten bespitzelt gleich der Detektiv.
(The pilot$_{ACC}$ spies-on soon the detective$_{NOM}$.)

(11) Den Piloten bespitzelt gleich der Zauberer.
(The pilot$_{ACC}$ spies-on soon the wizard$_{NOM}$.)

Whereas general world-knowledge would predict the detective to be the more likely agent of *spying*, the visual scene depicts *the wizard* to be more likely. At the verb, the wizard was fixated more than the other objects, suggesting that participants relied more on the immediate visual context than general knowledge. Only when the linguistic input revealed otherwise at the second noun phrase in (10) did people fixate the contextually conflicting but thematically more likely object (i.e., *detective*).

Based on these results, Knoeferle and Crocker (2006) proposed the coordinated interplay account (CIA) of situated language comprehension. Its central tenet is that incremental and anticipatory processing of linguistic input guides people's attention to a simultaneously present visual context. Furthermore, this guided attention toward mentioned and/or yet-to-be mentioned visual referents influences how the relevant visual context is used during sentence comprehension. The account also assumes that the visual contextual information takes priority over more general event knowledge when they conflict.

CIANet is the computational instantiation of the CIA. Mayberry et al. (2009) used CIANet to simulate the key findings of Knoeferle et al. (2005), Knoeferle and Crocker (2006), and several other visual-world experiments. In addition to including an SRN, CIANet consists of an input layer representing the visual input of a depicted scene coded in terms of event constituents (e.g., agent, action, and patient). During simulations, two events from the same scene are presented (e.g., *painting* and *washing* event from Figure 3.4), only one of which is described by the linguistic input. CIANet contains a layer of sigma-pi units (units that have multiplicative rather than additive connections) that gate attention to event constituents, enabling attention shifting. That is, the more the model attends to one event, the less it attends to the other, producing a competitive aspect to processing. The model's performance is analyzed quantitatively by comparing its outputs to the pattern of the behavioral data from the visual-world paradigm, or by generating a neurobehavioral equivalent from the pattern of hidden layer activation (see Crocker, Knoeferle, & Mayberry, 2010, for mapping of the model's hidden layer patterns to ERP data). Simulating Knoeferle et al.'s (2005) results, CIANet displayed predictive ambiguity resolution at the same point in time in the sentence (i.e., immediately after the verb). The model also simulated the human data from Knoeferle and Crocker's (2006) research, where participants relied on both constraints from the depicted event as well as their knowledge about common events, but relied more on the former when the two sources conflicted. Also note that as in the VSG, CIANet learned the constraints and their relevance from the input over the course of training. A final contribution of CIANet is that it simulates processing in a language other than English, and thus incorporates different constraints and investigates issues such as the on-line influence of case-marking.

Conclusions

Constraint-based models have been used to demonstrate how constraints are integrated over time both within and across the regions of an unfolding sentence. They have been used to simulate data originating from a number of syntactic ambiguities, methods, and languages. Constraint-based models have evolved from a general statement that all types of constraints matter, to multiple implemented models with carefully measured constraints, which has reduced theoretical degrees of freedom, and made them falsifiable. These explicit implementations were an important advance because of the difficulty in distinguishing among similar verbally described theories.

Acknowledgments

This work was supported by National Institutes of Health grant HD053136 and Natural Sciences and Engineering Council Grant OGP0155704 to Ken McRae.

References

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

Beach, C. (1991). The influence of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory & Language*, *30*, 627–643.

Bever, T. G. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognitive development of language* (pp. 279–362). New York: Wiley.

Binder, K. S., Duffy, S. A., & Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language*, *44*, 297–324.

Caplan, D., Hildebrandt, N., & Waters, G. S. (1994). Interaction of verb selectional restrictions, noun animacy, and syntactic form in sentence processing. *Language and Cognitive Processes*, *9*, 549–585.

Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic roles and language comprehension. In W. Wilkins (Ed.), *Thematic relations* (pp. 263–288). New York: Academic Press.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 687–696.

Clifton, C. Jr., & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes*, *4*, 77–104.

Clifton, C. Jr., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, *49*, 317–334.

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: Cambridge University Press.

Crocker, M. W., Knoeferle, P., & Mayberry, M. R. (2010). Situated sentence processing: The coordinated interplay account and a neurobehavioral model. *Brain & Language*, *112*, 189–201.

Daugherty, K., & Seidenberg, M.S. (1992). Rules or connections? The past tense revisited. In *Proceedings of the 14th annual meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Elman, J. L. (1991). Representation and structure in connectionist model. In G. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. (pp. 345–382). Cambridge, MA: MIT Press.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*, 71–99.

Elman, J. L. (1995). Language as a dynamical system. In R. Port & T. van Gelder (Eds.), *Mind as motion* (pp. 195–223). Cambridge, MA: MIT Press.

Elman, J. L., Hare, M., & McRae, K. (2004). Cues, constraints, and competition in sentence processing. In M. Tomasello & D. Slobin (Eds.), *Beyond nature–nurture: Essays in honor of Elizabeth Bates* (pp. 111–138). Mahwah, NJ: Lawrence Erlbaum Associates.

Ferreira, F., & Clifton, C. Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye-movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 555–568.

Ferretti, T. R., & McRae, K. (1999). Modeling the role of plausibility and verb-bias in the direct object/sentence complement ambiguity. In *Proceedings of the twenty-first annual conference of the Cognitive Science Society* (pp. 161–166). Hillsdale NJ: Erlbaum.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559–586). Hillsdale, NJ: Erlbaum.

Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research*, *24*, 437–468.

Frazier, L. & Clifton, Jr., C. (2005) The syntax-discourse divide: Processing ellipsis. *Syntax*, 8(2), 121–174.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye-movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *143*, 178–210.

Gahl, S., Jurafsky, D., & Roland, D. (2004). Verb subcategorization frequencies, American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, & Computers*, *36*, 432–443.

Garnsey, S. M., Pearlmutter, N., Meyers, E., & Lotocky, M. A. (1997). The contribution of verb-bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Green, M. J., & Mitchell, D. C. (2006). Absence of real evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *55*, 1–17.

Hanna, J. E., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1996). Integrating discourse and local constraints in resolving lexical thematic ambiguities. In G. Cottrell (Ed.), *Proceedings of the eighteenth annual conference of the Cognitive Science Society* (pp. 266–271). Hillsdale, NJ: Erlbaum.

Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*, *48*, 281–303.

Hare, M., McRae, K., & Elman, J. L. (2004). Admitting that admitting verb sense into corpus analyses makes sense. *Language and Cognitive Processes*, *19*, 181–224.

Hintzman, D. L. (1991). Why are formal models useful in psychology? In W. E. Hockley and S. Lewandowsky, (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 39–56). Hillsdale, NJ: Erlbaum.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137–194.

Kennison, S. M. (2001). Limitations on the use of verb information during sentence comprehension. *Psychonomic Bulletin & Review*, *8*, 132–138.

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye-tracking. *Cognitive Science*, *30*, 481–529.
Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*, 95–127.
Kuçera, H., & Francis, W. N. (1967). *A computational analysis of present-day American English*. Providence, RI: Brown University Press.
Levin, B., & Rappaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry*, *17*, 623–661.
MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*, 157–201.
MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 483–506.
McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, *12*, 137–176.
McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*, 283–312.
Mayberry, M. R., Crocker, M. W., & Knoeferle, P. (2009). Learning to attend: A connectionist model of situated language comprehension. *Cognitive Science*, *33*, 449–496.
Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 601–618). Hillsdale, NJ: Erlbaum.
Murray, W., & Liversedge, S. (1994). Referential context and syntactic processing. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 359–388). Hillsdale, NJ: Erlbaum.
Pado, U., Crocker, M. W., & Keller, F. (2009). A probabilistic model of semantic plausibility in sentence processing. *Cognitive Science*, *39*, 794–838.
Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. *Journal of Verbal Learning and Verbal Behavior*, *22*, 358–374.
Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.
Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*, 348–379.
Roland, D., & Jurafsky, D. (2002). Verb sense and verb subcategorization probabilities. In S. Stevenson & P. Merlo (Eds.), *The lexical basis of sentence processing: Formal, computational, and experimental issues* (pp. 325–346). Philadelphia, PA: John Benjamins Publishing.
Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, *25*, 249–271.
Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 1521–1543.
Spivey-Knowlton, M. (1994). Quantitative predictions from a constraint-based theory of syntactic ambiguity resolution. In M. C. Mozer, D. S. Touretzky, & P. Smolensky (Eds.), *Proceedings of the 1993 connectionist models summer school* (pp. 130–137). Hillsdale, NJ: Lawrence Erlbaum.

Spivey-Knowlton, M. (1996). Integration of linguistic and visual information: Human data and model simulations. Unpublished doctoral dissertation, University of Rochester.

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

Spivey-Knowlton, M., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology*, *37*, 276–309.

Staub, A. (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 550–569.

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. *Language and Cognitive Processes*, *12*, 211–271.

Tabor, W., & Tanenhaus, M. K. (2001). Dynamical systems for sentence processing. In M. H. Christiansen & N. Chater (Eds.), *Connectionist psycholinguistics* (pp. 177–211). Westport, CT: Ablex Publishing.

Tanenhaus, M. K., Boland, J. E., Garnsey, S. M., & Carlson, G. N. (1989). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research: Special Issue on Sentence Processing*, *18*, 37–50.

Tanenhaus, M. K., & Carlson, G. N. (1989). Lexical structure and language comprehension. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 529–561). Cambridge, MA: MIT Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Tanenhaus, M. K., Spivey-Knowlton, M. J., & Hanna, J. E. (2000). Modeling thematic and discourse context effects on syntactic ambiguity resolution within a multiple constraints framework: Implications for the architecture of the language processing system. In M. Pickering, C. Clifton, & M. Crocker (Eds.), *Architecture and mechanisms of the language processing system* (pp. 90–118). Cambridge: Cambridge University Press.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In P. Eimas & J. Miller (Eds.), *Handbook in perception and cognition*, *Volume 11: Speech, language, and communication* (pp. 217–262). San Diego, CA: Academic Press.

Traxler, M. J., Pickering, M. J., & Clifton, C. Jr. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*, 558–592.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, *35*, 566–585.

Trueswell, J. C., & Gleitman, L.R. (2004). Children's eye-movements during listening: Evidence for a constraint-based theory of parsing and word learning. In J. M. Henderson & F. Ferreira (Eds.), *Interface of language, vision, and action: Eye-movements and the visual world* (pp. 319–346). New York: Psychology Press.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285–318.

Trueswell, J. C., Tanenhaus, M. K., & Kello, K. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 528–553.

Van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*, 284–307.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*, 225–258.

Vaughan, J. (1983). Control of fixation duration in visual search and memory search: Another look. *Journal of Experimental Psychology: Human Perception & Performance*, *8*, 709–723.

Appendix: Some of the constraints that have been identified and tested

- probabilistic global syntactic biases
  - subject–verb–object (SVO) construction
  - passive constructions

- lexically specific syntactic biases
  - verb subcategorization frames/argument structure
    - direct object/sentence complement
    - sense-specific verb subcategorization biases
    - transitivity biases
    - dative alternations
    - verb tense probabilities (past tense versus past participle)
  - prepositions (*by*, *with*, *on*, *in*)
  - *that* preferences

- combinatorial preferences
  - verb + preposition (verbed + *by*)
  - psych verb + *with*
  - action verb + *with*
  - verb + *that*

- thematic biases
  - thematic grids: probabilities of thematic roles of a verb
  - verb-general selectional restrictions (e.g., ±animacy as subject or direct object)
  - verb-specific selectional restrictions (e.g., *eat*: ±edible)
  - verb-specific thematic fit (e.g., *cop* versus *crook* as agent or patient of *arrested*)
  - sense-specific thematic fit (e.g., *brakes* versus *spelling* as patient of *mechanic checked* versus *editor checked*)

- physical aspects of speech
  - intonation, prosody, contrastive stress, duration of, e.g., final syllable of a verb

- extra-sentential linguistic context
  - reference
  - pragmatics

- visual context
  - contexts of possible referents
  - depicted actions and relations among characters in a visually presented scene