Chapter 12
Syntactic Parsing

Martin J. Pickering and Roger P. G. van Gompel

When people hear connected speech or read texts, they start processing immediately. In a classic demonstration, Marslen-Wilson (1973, 1975) had listeners shadow speech, and found that their errors were constrained by prior semantic context even when the shadowing lag was only around 300 ms. This indicates that sentence interpretation can occur extremely rapidly. Many subsequent studies have demonstrated that lexical, syntactic, and aspects of semantic processing occur without appreciable delay, for example during reading (e.g., Just & Carpenter, 1980) or spoken language comprehension in the presence of a visual array (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Furthermore, recent evidence suggests that people may even anticipate properties of upcoming words in the sentence (e.g., Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005).

The phenomenon that each word in the sentence is interpreted immediately when it is encountered is referred to as *incrementality*. One aspect of this is that people appear to compute the grammatical structure of sentences incrementally. Most of the evidence for incremental syntactic processing or *parsing* comes from the study of potentially ambiguous sentences. For example, in sentence (1), people experience difficulty as soon as they encounter the phrase *by the lawyer* (Clifton et al. 2003; Ferreira & Clifton, 1986; Rayner, Carlson, & Frazier, 1983; Trueswell, Tanenhaus, & Garnsey, 1994):

1. The evidence examined by the lawyer turned out to be unreliable.

After *the evidence examined*, the sentence is temporarily (or locally) ambiguous between the correct "reduced relative" analysis in which *examined* is a past participle, and a "main clause" analysis in which *examined* is a past tense verb. But the phrase *by the lawyer* provides very good evidence that the main clause analysis is wrong and that the reduced relative analysis is correct. Difficulty reading *by the lawyer* therefore suggests that people initially select or favor the main clause analysis, and experience difficulty when they realize that this analysis is probably wrong. In psycholinguistic terminology, they are "garden pathed" (Bever, 1970).

Handbook of Psycholinguistics: 2nd Edition
ISBN: 0-12-369374-8

Such findings show that people do not wait until the end of the sentence before syntactically analyzing it. They are compatible with a *serial* account, in which people initially adopt one analysis (in this case, the main clause analysis). On this account, difficulty occurs when people realize this inconsistency and *reanalyze* (i.e., adopt a different analysis). They are also compatible with a *ranked parallel* account, in which people adopt more than one analysis, but rank one higher than any others. Difficulty occurs when later information causes people to re-rank their analyses. All current accounts assume that syntactic processing is either serial or ranked parallel.

Hence syntactic and semantic processing begin without any appreciable delay. This might suggest that the processor has an architecture in which different sources of information are integrated immediately. However, this assumption has proved extremely controversial, as we shall now see. We therefore first review theories and data relating to initial processing. Following this, we turn to questions of reanalysis and processing complexity, and then address a range of newer topics in parsing research that provide important links between the field and other areas of psycholinguistics and cognition more generally.

1. TWO-STAGE ACCOUNTS

For a long time, the study of parsing was almost entirely dominated by the question of how initial parsing decisions are made. The "big" question underlying this research was the question of *encapsulation*: the extent to which different knowledge sources are formally separated. Encapsulation is a critical property of *modularity*, which roughly amounts to the thesis that the mind consists of separate, specialized components that exist independently of a central store of general knowledge (J. A. Fodor, 1983). Although Fodor treated language as one module, parsing research has tended to ask whether specific aspects of language, such as syntax, are modular.

In practical terms, the investigation of encapsulation has addressed the issue of ambiguity resolution during initial processing. The earliest accounts of syntactic ambiguity resolution assumed that decisions were based on strategies such as interpret a string of words as agent verb patient (the *NVN* strategy) if possible (Bever, 1970; cf. Kimball, 1973). Such accounts are modular, because they make reference to syntactic information alone, and do not, for example, pay attention to the plausibility of the alternative analyses. However, people eventually make use of information like plausibility in choosing an analysis. Thus any modular account is *two-stage*, with initial processing being modular, but subsequent processing being usually not modular. This dichotomy is most clearly described in Rayner et al. (1983), where a separate (*thematic*) processor plays no role in the initial choice of analysis, but is used during reanalysis.

So what guides initial processing? Although Bever (1970) and Kimball (1973) proposed sets of principles, it was Frazier (1979, 1987a) who established by far the most influential account, which came to be known as the "Garden-Path" theory.[1] In common with almost all two-stage accounts, it assumes serial processing. A single analysis is chosen on the basis of principles defined in terms of a phrase-structure tree. *Minimal Attachment* stipulates that an ambiguous phrase is attached to the preceding tree structure using the fewest number of nodes. Consider (2):

2. The spy saw the cop with binoculars but the cop didn't see him.

Frazier assumed that VP-attachment in the first clause (i.e., the spy used binoculars to see the cop) involved a flat tree structure [V NP PP]$_{VP}$ whereas NP-attachment (i.e., the cop had binoculars) involved an embedded complex noun phrase [V [NP PP]$_{NP}$]$_{VP}$. Since VP-attachment involves one fewer node than NP-attachment, it is the minimally attached analysis and is therefore adopted initially. Rayner et al. (1983) tested this prediction by contrasting it with (3), where the VP-attached analysis is implausible:

3. The spy saw the cop with a revolver.

Using eye tracking, they found that readers had more difficulty with sentences such as (3) than with sentences such as (2), and suggested that they attach to the VP in both cases, but subsequently revise their initial decision in (3). This reanalysis makes use of the thematic processor.

Minimal attachment was used to explain a number of other types of locally ambiguous sentences, including reduced relatives (4), object/complement ambiguities like (5), and ambiguities caused by head-final verbs in languages like Dutch (6):

4. The florist sent the flowers was very pleased.
5. The man realized his goals were out of reach.
6. … dat het meisje van Holland glimlachte.
   "… that the girl from Holland smiled."

In all of these cases, minimal attachment posits that the processor initially adopts the ultimately incorrect analysis. Notice, however, that it can explain preferences in globally ambiguous sentences like (2) as well as locally ambiguous sentences. In all the cases, some experimental evidence provided support for the principle (e.g., Frazier & Rayner, 1982; Frazier, 1987b; Rayner et al., 1983; Ferreira & Clifton, 1986; Ferreira & Henderson, 1990), but other evidence supports alternative accounts (see below).

[1] Note the capitalization. In our usage, garden path is a descriptive term. In a serial account, people garden path if they choose a syntactic analysis that later turns out to be incorrect. In a ranked parallel account, people garden path if their highest ranked analysis later turns out to be incorrect. Garden Path theory, on the other hand, refers to the two-stage serial account proposed by Frazier (1979, 1987a).

Late closure stipulated that incoming material was to be attached into the constituent currently being processed. It appeared to deal with two rather different types of ambiguity. First, it predicted that a phrase would form part of the current constituent rather than start a new constituent, as in (7). Second, it predicted that a phrase would form part of the most recent constituent possible, as in (8):

7. When Mary was knitting the socks fell to the floor.
8. John said that Sue left yesterday.

In (7), *the socks* can serve as the object of *knitting*, in which case it forms part of the verb phrase headed by *knitting*; or as the subject of a new clause, as turns out to be the case. Late closure predicts that people would initially treat *the socks* as the object of *knitting*. But this analysis becomes impossible after *fell* (because *fell* would have no subject). In accord with this, Frazier and Rayner (1982) found that people experienced difficulty after reading *fell*. In the globally ambiguous (8), late closure predicts that *yesterday* should modify *left* (i.e., Sue left yesterday) rather than *said* (i.e., John spoke yesterday). This amounts to a preference for "low" attachment over "high" attachment (Kimball, 1973). According to Frazier (1987a), minimal attachment takes precedence when the principles are in conflict. For example, the processor adopts VP-attachment in (2), in accord with minimal attachment, even though late closure would support NP-attachment. But in (7) and (8), Frazier claims that the two analyses do not differ in the number of nodes, and so late closure applies.

In fact, minimal attachment and late closure require certain additional assumptions to make the above predictions for (1)–(8). First, they depend on specific assumptions about tree structure. For example, minimal attachment only predicts VP-attachment of (2) if the VP-analysis has a flat structure, an assumption which is not adopted within modern transformational frameworks (e.g., Kayne, 1984; Chomsky, 1995). Second, the processor sometimes has to postulate structure associated with words it has not encountered yet. In (6), the noun phrase *van Holland* could constitute an argument of a following verb, such as *houdt* (liked), because such arguments precede past tense verbs in Dutch subordinate clauses. Minimal attachment predicts that *van Holland* is an argument rather than a modifier of *het meisje* so long as the processor postulates the verb at the noun phrase.

In later years, Garden-Path theory introduced another principle known as the *active-filler strategy* (Frazier, 1987b; Frazier & Clifton, 1989; Frazier & Flores D'Arcais, 1989) or the *minimal chain principle* (De Vincenzi, 1991). This was designed to deal with unbounded dependencies, such as *wh*-questions and relative clauses, in which the verb and its argument can be separated by clause boundaries:

9. Which girl do you believe John loves a lot?

In (9), *which girl* is the object of *loves*, but is found at the beginning of the sentence. Garden-Path theory assumed that such sentences involved transformations, so that the *filler* (*which girl*) moves from its canonical location after *loves* to the front of the sentence, and leaves a *gap* (known as a wh-*trace*) at its canonical location (e.g., Chomsky,

1981). Psycholinguistic accounts that assume transformational grammar (e.g., J. D. Fodor, 1978) claim that the processor associates the filler and the gap (*gap filling*), and then integrates the filler with the verb. The active-filler strategy assumes that the processor favors any analysis that allows gap filling over any analysis that does not. In (9), it is locally possible that *which girl* serves as the object of *believe*, if it is used transitively, so Garden-Path theory predicts that this analysis is initially adopted. It becomes impossible as soon as *John* is encountered, and therefore causes what has been termed a "filled gap" effect (Stowe, 1986).

The active-filler strategy makes an interesting prediction for word order ambiguities and relative clause ambiguities in languages like Dutch and German (Frazier, 1987b; Frazier & Flores D'Arcais, 1989). For example, it predicts that temporarily ambiguous Dutch subject relative clauses such as (10a) are easier to process than object relatives such as (10b).

10a. Karl hielp de mijnwerker die de boswachters vond.
"Karl helped the miners who found the forester."
10b. Karl hielp de mijnwerker die de boswachters vonden.
"Karl helped the miners who the forester found."

Because languages such as Dutch and German are assumed to have an underlying subject-object-verb order (e.g., Koster, 1975), the subject gap in (10a) caused by the extraction of the relative pronoun filler (*die*) precedes *de boswachters*, whereas the object gap in (10b) follows it. As a result, the subject gap can be filled earlier than the object gap, so subject relatives should be easier to process than object relatives. For similar reasons, the active-filler strategy also predicts that temporarily ambiguous sentences with a subject-verb-object (SVO) order in Dutch and German should be easier to process than sentences with an object-verb-subject (OVS) order. These predictions are supported by several studies (e.g., Bader & Meng, 1999; Frazier, 1987b; Frazier & Flores D'Arcais, 1989; Hemforth, 1993; Kaan, 2001; Mak, Vonk, & Schriefers, 2002; Schriefers, Friederici, & Kühn, 1995).

Notice that the active-filler strategy assumes the existence of gaps in accord with most versions of transformational grammar (e.g., Chomsky, 1981). However, alternative linguistic theories eschew gaps (e.g., Pollard & Sag, 1994, Chapter 9; Steedman, 2000), and in accord with such theories, Pickering and Barry (1991) proposed that the processor associates fillers directly with the verb (or other subcategorizer) without going via a gap. Such an account makes the same predictions as the active-filler strategy for sentences like (8). It does not account for ambiguity resolution preferences in word order ambiguities and relative clauses in Dutch and German, so they would have to be due to other factors, for example a preference for a particular information structure (e.g., Kaan, 2001). However, it does explain why people experience difficulty with (11) at the verb *shot*, even though the gap would follow *the hapless man* (Traxler & Pickering, 1996):

11. That is the very small pistol in which the heartless killer shot the hapless man yesterday afternoon.

The active-filler strategy can only predict this finding if the processor can postulate gaps before they occur (Gibson & Hickok, 1993).

Although Garden-Path theory has been very much the dominant two-stage account, there have been many alternative accounts in which initial decisions are based on some sources of information to the exclusion of others. Most of these accounts propose that the processor initially adopts an analysis in which the verb (or other element) can assign a thematic role to a new constituent (Abney, 1989; Crocker, 1995; Pritchett, 1992). This can also be interpreted as an initial preference for adjuncts over arguments. For example, Schütze and Gibson (1999) found a preference for NP-attachment in sentences similar to (2) when a prepositional phrase was ambiguous between being an argument of the noun phrase and an adjunct of the verb phrase (cf. Clifton, Speer, & Abney, 1991). Such research could be used to discriminate between different modular accounts, but in practice far more attention has been paid to the question of whether the processor is modular or not. Before reviewing relevant evidence, we briefly outline the main characteristics of interactive (i.e., non-modular) accounts.

2. INTERACTIVE ACCOUNTS

In contrast to modular sentence processing accounts such as the Garden-Path model, interactive accounts assume that all potentially relevant sources of information can be used immediately during sentence processing and can affect initial processing decisions. Current interactive sentence processing models, which have been developed from earlier interactive models (e.g., Bates & MacWhinney, 1989; Tyler & Marslen-Wilson, 1977; Taraban & McClelland, 1988), are often called constraint-based (or constraint-satisfaction) models (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Trueswell, Tanenhaus, & Kello, 1993; Trueswell et al., 1994); though there are also hybrid models in which some information can be delayed (e.g., Boland & Blodgett, 2001; Boland & Boehm-Jernigan, 1998). Interactive models generally assume that the processor activates all possible analyses of a sentence in parallel, and that the activation of the analyses depends on the amount of support they receive from the various sources of information. When one analysis receives much more support than its alternatives, processing is easy, but when two or more analyses receive about equal support, processing difficulty occurs. Additionally, they tend to be lexicalist, in that they assume that most or all syntactic information is stored with individual lexical items (e.g., MacDonald et al., 1994; Trueswell, 1996).

One of the difficulties with constraint-based models is that they tend not to be very predictive unless they have identified the full set of constraints that affect processing and have a precise model of how these constraints affect processing. However, some recent accounts have used computational modeling to derive more precise predictions. For example, Spivey and Tanenhaus (1998) and McRae et al. (1998) reported a model that explains how various sources of information affect phrase-by-phrase processing of reduced-relative ambiguities. In this model, all syntactic analyses of an ambiguous structure are activated in parallel, and their activation is determined by various constraints. At each word in the sentence, the activation of the analysis that receives most support from the constraints increases until it reaches a threshold level and the processor moves to the next word. Reading times are modeled as the number of cycles the processor has to go through before it reaches this threshold level of activation. Hence, if two or more analyses receive about equal support from the various constraints, reading times should be long, but if one analysis receives much more support than its alternatives, reading times should be short.

Tabor, Juliano, and Tanenhaus (1997) and Tabor and Tanenhaus (1999) developed a learning-based model to predict sentence processing difficulty. During learning, the model maps language input into a multidimensional space. Sentence fragments that have similar continuations (because they are syntactically and semantically similar) occupy positions close together in the space and form clusters that function as attractors. These attractors can be considered as different analyses of ambiguous sentence fragments. Reading times during subsequent sentence processing are modeled as the time it takes for a sentence fragment to reach one of the attractors. When a sentence fragment is very similar to a single cluster of previously encountered fragments (i.e., it has a similar syntax and semantics), it starts at a position very close to the cluster, so it will quickly reach the attractor and reading times are predicted to be fast. In contrast, when a sentence fragment is similar to more than one attractor, that is, the syntax and semantics of the fragment is consistent with more than one analysis, reading times are long.

Computational models make constraint-based accounts more testable and have highlighted the different sources of information that appear to affect sentence processing. We now turn to these sources of information.

3. FREQUENCY EFFECTS

An obvious possibility is that choice of analysis is affected by frequency. It is therefore not surprising that many early accounts suggested that the processor initially favors frequent analyses over infrequent ones (Clifton, Frazier, & Connine, 1984; J. D. Fodor, 1978; Ford, Bresnan, & Kaplan, 1982; Mitchell & Holmes, 1985). Such accounts almost inevitably clash with Garden-Path theory, because principles like minimal attachment make no reference to frequency. On the other hand, frequency plays a very natural role within interactive accounts, so evidence for early effects of frequency would support them.

However, any frequency-based account needs to answer the question: frequency of what? The early accounts largely assumed that the processor counted the frequency with which particular verbs (or other elements) were used in a particular construction. Some verbs are most commonly used transitively (e.g., *read*), whereas others are most commonly used intransitively (e.g., *sing*), and the processor might simply adopt the most frequent analysis for each verb in cases of ambiguity. Alternatively, the processor might ignore the frequency of individual verbs, and simply adopt the most frequent construction. For example, verbs are more commonly used transitively than intransitively (in English at least) and so the processor might always adopt the transitive analysis, even when the verb is most commonly used intransitively. Between these extremes, it could "group" similar verbs together (e.g., assessing the frequency with which verbs of physical transfer are used transitively). It also has to resolve the question of how "construction" is defined (e.g., do transitives in main clauses and subordinate clauses get "counted" together?). Mitchell, Cuetos, Corley, and Brysbaert (1995) refer to such issues as questions of *grain size*. Fine-grained accounts pay more attention to individual properties of sentences, but may face the "sparse data problem" if many categories are counted. Coarse-grained models make predictions that are simply very inaccurate for some ambiguities.

Mitchell et al. (1995) assumed that the processor would make a specific choice about grain size, and proposed that it used only coarse-grained information during initial processing (i.e., information independent of lexical items). In contrast, constraint-based theories claim that the processor employs fine-grained information associated with individual lexical items (e.g., Garnsey, Pearlmutter, Myers, & Lotocky, 1997; MacDonald et al., 1994; Trueswell et al., 1993; Trueswell, 1996). In fact, they also tend to assume that the processor uses various types of frequency information during syntactic ambiguity resolution including coarse-grained frequency information that is independent of lexical items, lexically specific frequency information, and perhaps even more fine-grained information associated with combinations of lexical items (e.g., McRae et al., 1998).

To explore the role of frequency, let us consider the resolution of complex noun-phrase ambiguities like (12), in which the Spanish sentence (12b) is a translation of the English sentence (12a):

12a. The journalist interviewed the daughter of the colonel who had the accident.
12b. El periodista entrevisto a la hija del coronel que tuvo el accidente.

This construction has been particularly important in assessing the Garden-Path theory. According to late closure, people should initially assume that *who had the accident* attaches "low" to *colonel* (i.e., so that he had the accident) rather than "high" to *daughter* (i.e., so that she had the accident). Although Cuetos and Mitchell (1988) found evidence for low attachment in (12a), they found evidence for *high* attachment in (12b). More recent evidence suggests that English shows either a weak preference for low attachment or no clear preference at all (e.g., Carreiras & Clifton, 1993, 1999; Traxler, Pickering, & Clifton, 1998) and there is some evidence for a similar low attachment preference in Italian (De Vincenzi & Job, 1995; but cf. Frenck-Mestre & Pynte, 2000), but other languages such as French, German, and Dutch show a high attachment preference (e.g., Brysbaert & Mitchell, 1996; Hemforth, Konieczny, & Scheepers, 2000; Zagar, Pynte, & Rativeau, 1997).

Several accounts explaining cross-linguistic differences in relative clause attachment have been proposed. Frazier and Clifton (1996) proposed Construal theory, which is a hybrid processing model in that it claims that structural parsing principles such as minimal attachment and late closure operate for structures involving primary syntactic relations (roughly, involving arguments), whereas for non-primary relations, as found in relative clause attachment, the processor immediately uses non-syntactic information.

Frazier (1990) proposed that low attachment in (12a) is preferred in English because the Saxon genitive (the colonel's daughter) is more common, so when the Norman genitive (the daughter of the colonel) is used, it indicates that relative clause attachment is low. By contrast, Spanish does not have a Saxon genitive.

Gibson, Pearlmutter, Canseco-Gonzalez and Hickok (1996) and Gibson, Pearlmutter, and Torrens (1999) argued that the same parsing strategies are used in English and Spanish. They investigated relative clauses that could be attached to one of the three noun phrases and observed that both in English and Spanish, attachment to the most recent, third noun phrase was easiest to process, followed by an attachment to the first noun phrase, while attachment to the middle noun phrase was very hard to process. They argued that attachment to the third noun phrase is preferred as a result of recency, while attachment to the first noun phrase is relatively easy due to *predicate proximity*: a preference to attach as close to the head of a predicate phrase as possible. The latter principle is assumed to be different in strength across languages.

Mitchell et al. (1995) and Cuetos, Mitchell, and Corley (1996) argued instead that the observed cross-linguistic differences in relative clause attachment are due to differences in the frequency of occurrence of high and low attachment between languages. They suggested that low attachment is preferred in English for sentences such as (12), because across all relative clauses involving two potential attachment sites, low attachment is frequent than high attachment. In contrast, they argued that high attachment is more preferred in other languages, because in these languages, it is more frequent than low attachment. However, such a coarse-grained frequency account is inconsistent with a number of studies. First, several studies (e.g., Gilboy, Sopena, Clifton, & Frazier, 1995; Traxler et al., 1998) have shown that relative clause attachment preferences depend on the preposition in the complex noun phrase. In particular, when the preposition is *with* rather than *of* (as in 12), the preference for low attachment is much stronger. If this is due to frequency information, this indicates that the processor takes into account information from (closed class) lexical items (but see Frazier & Clifton, 1996 for a different account explaining relative clause attachment preferences). Second, Brysbaert and Mitchell (1996) showed a preference for high attachment in Dutch, even though low attachment is more frequent (Mitchell & Brysbaert, 1998). While Desmet and Gibson (2003) criticized some of the evidence against coarse-grained frequency accounts on methodological grounds (Gibson & Schutze, 1999), Desmet, De Baecke, Drieghe, Brysbaert and Vonk (in press) argued for a more fine-grained account to explain relative clause attachment preferences in Dutch. They showed that relative clause attachment preferences were affected by the animacy and concreteness of the noun phrases: Attachment to animate and concrete noun phrases was preferred to attachment to inanimate and abstract noun phrases. These online attachment preferences corresponded to relative clause attachment corpus frequencies when the animacy and concreteness of the noun phrases was taken into account (see also Desmet, Brysbaert, & De Baecke, 2002).

Other research has more directly pitted lexical frequency-based explanations against Garden-Path theory. For example, Trueswell et al. (1993) investigated sentences such as

(13), which are temporarily ambiguous because the main verb (*forgot/hoped*) can occur with a direct object (e.g., *the man forgot the solution*) or a sentence complement, which is the correct analysis.

13a. The student forgot the solution was in the book.
13b. The student hoped the solution was in the book.

Trueswell et al. compared these sentences with unambiguous sentences disambiguated by the complementizer *that* following the main verb. When the verb occurred more often with a direct object than a sentence complement (e.g., *forgot*, as measured by sentence completions), the temporarily ambiguous sentences were harder to process than unambiguous controls. However, no difference was observed when the verb occurred more frequently with a sentence complement than a direct object (e.g., *hoped*). Hence, they concluded that verb subcategorization frequencies have an immediate effect on sentence processing. Recent evidence by Hare, McRae, and Elman (2003) indicates that the semantics of the verb affects subcategorization preferences. For example, *find* occurs with a direct object when it means *locate*, whereas it occurs with a sentence complement when it means *realize*. Hare et al. showed that a context that instantiated one or the other sense of the verb affected processing difficulty at the disambiguation.

Trueswell (1996) manipulated the frequency of the verb as a past participle or past tense in reduced relative clause ambiguities such as (14).

14a. The message recorded by the secretary could not be understood.
14b. The room searched by the police could not be understood.

The verb *recorded* is most frequently used as a past participle, so this supports the reduced relative analysis, whereas *searched* is usually a past tense, consistent with the main clause analysis. Trueswell (1996) observed that the disambiguating *by*-phrase in sentences such as (14a) was harder to process than unambiguous controls, but there was no difference in sentences such as (14b), where the verb was biased toward the past participle analysis. However, this pattern of results depended on the animacy of the first noun phrase; when it was inanimate, difficulty in the ambiguous sentences occurred regardless of the verb bias, suggesting that both animacy and verb bias had to support the reduced relative analysis in order to eliminate any processing difficulty.

A number of other studies also provide evidence that lexical frequency information affects syntactic ambiguity resolution (e.g., Boland, Tanenhaus, Garnsey, & Carlson, 1995; Clifton et al., 1984; Garnsey et al., 1997; Holmes, Stowe, & Cupples, 1989; Mitchell & Holmes, 1985; Snedeker & Trueswell, 2004; Stowe, Tanenhaus, & Carlson, 1991). Importantly, some of these studies suggest that the effect of subcategorization frequency is very early (e.g., Garnsey et al., 1997; Snedeker & Trueswell, 2004). Clearly, this poses difficulties for Garden-Path theory or any other modular model in which the use of frequency information is delayed and supports constraint-based models or other accounts in which frequency information is used immediately during sentence processing.

However, a number of other studies suggest that the processor does not entirely rely on subcategorization frequency information. For instance, Mitchell (1987) had participants read (15) in two segments, indicated by the slash:

15a. After the young Londoner had arrived his parents/prepared to celebrate their anniversary.
15b. After the young Londoner had visited his parents/prepared to celebrate their anniversary.

They spent longer reading the first fragment of (15a) than (15b). This suggests that they initially adopted the transitive analysis and experienced difficulty when they subsequently discovered that this is inconsistent with subcategorization information of the verb. In contrast, they had more difficulty reading the second fragment of (15b) than (15a), suggesting that they had already reanalyzed (15a) but not (15b). Van Gompel and Pickering (2001) found similar results using eye tracking, and therefore countered the claim that Mitchell's effects might have somehow been the results of unusual presentation conditions (Adams, Clifton, & Mitchell, 1998).

Similarly, Pickering, Traxler, and Crocker (2000) observed that readers experienced difficulty with temporarily ambiguous sentence complement clauses even though their verbs were biased toward the sentence complement analysis, and Kennison (2001) showed that temporarily ambiguous sentence complements were harder than unambiguous controls regardless of the verb bias. Finally, McKoon and Ratcliff (2003) failed to find any evidence that frequency information affected the processing of reduced relatives, and argued instead that verb complexity accounts for differences in processing difficulty for reduced relatives (though cf. McRae, Hare, & Tanenhaus, in press).

It is probably safest to conclude that lexical frequency information has some influence on syntactic ambiguity resolution, but in many cases does not neutralize difficulty with temporarily ambiguous sentences. This formulation is consistent with constraint-based theories. Lexical frequency information is just one of the many constraints that affect sentence processing, so even when lexical frequency information supports a particular analysis, the alternative analysis may be activated by other sources of information such as a coarse-grained frequency preference. Without knowing all the constraints that affect sentence processing and their weights, it is difficult to test and falsify constraint-based theories. One possibility is to use sentence fragment completion data to estimate the influence that the different constraints together have (e.g., Garnsey et al., 1997; McRae et al., 1998), essentially adopting the view that the sentence processor takes into account extremely fine-grained frequency information related to combinations of different words. However, it is uncertain whether such completion preferences always correctly predict online preferences in syntactic ambiguity resolution (e.g., Binder, Duffy, & Rayner, 2001; Kennison, 2001; Pickering et al., 2000).

The results from studies investigating the use of frequency information can also be explained in modular accounts which assume that the use of frequency information is delayed