Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 48 (2003) 281–303
Journal of Memory and Language
www.elsevier.com/locate/jml

Sense and structure: Meaning as a determinant of verb subcategorization preferences

Mary Hare,[a,*] Ken McRae,[b] and Jeffrey L. Elman[c]

[a] *Department of Psychology, Bowling Green State University, Bowling Green, OH 43403-0228, USA*
[b] *University of Western Ontario, London, Canada*
[c] *University of California, San Diego, La Jolla, CA, USA*

Received 10 August 2001; revision received 14 June 2002

Abstract

Readers are sensitive to the fact that verbs may allow multiple subcategorization frames that differ in their probability of occurrence. Although a verb's overall subcategorization preferences can be described probabilistically, underlying non-random factors may determine those probabilities. One potential factor is verb semantics: Many verbs show sense differences, and a verb's subcategorization profile can vary by sense. Thus, although *find* can occur with a direct object (DO) or a sentential complement (SC), when it is used to mean 'locate' it occurs only with a DO, whereas in its 'realize' sense it is SC-biased, but can take either frame. We used corpus analyses to identify verbs that occur with both frames, and found that their subcategorization probabilities differ by sense. Off-line sentence completions demonstrated that contexts can promote a specific sense of a verb, which subsequently influenced subcategorization probability. Finally, in a self-paced reading time experiment, verbs occurred in target sentences containing either a structurally unambiguous or ambiguous SC, following a context favoring the verb's DO- or SC-biased sense. Sense-biasing context influenced reading times at *that*, and interacted with ambiguity in the disambiguating region. Thus, readers use sense-contingent subcategorization preferences during on-line language comprehension.

The nature and use of information regarding the syntactic arguments that a verb can take is crucial to language comprehension and production. A critical component of knowing how to use a verb involves knowing its combinatory properties with other words, including both the range of possible syntactic complements it permits (subcategorization profile; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Kello, 1993), and the semantic constraints on its possible arguments (selectional restrictions, or more precisely, thematic fit, McRae, Ferretti, & Amyote, 1997). In addition, verb subcategorization information has been of strategic importance in testing psycholinguistic theories because theories of sentence processing differ regarding the time-course of the availability and use of this information during on-line comprehension. In particular, the constraint-based or expectation-driven framework argues that a broad range of information is available and used by listeners or readers in determining even initial structural analyses (Altmann, 1998, 1999; MacDonald et al., 1994; MacWhinney & Bates, 1989; Spivey & Tanenhaus, 1998). In this paper, in concert with Roland (2002), Roland and Jurafsky (1998, 2002), and Roland et al. (2000), we argue that knowledge of lexically specific structural constraints may be quite fine-grained, and specifically that important aspects of subcategorization information are encoded not in relation to a verb, but to the verb's specific senses.

One way to determine the types of information computed during comprehension, and to identify

* Corresponding author. Fax: +419-372-6013.
E-mail address: hare@rowan.bgsu.edu (M. Hare).

0749-596X/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S0749-596X(02)00516-8

precisely when that information influences processing, is to study situations in which potentially available information sources are in conflict or underdetermine the interpretation of the input. Such situations arise whenever a word or a syntactic structure is temporarily ambiguous. For these reasons, ambiguity resolution has been a fruitful domain for testing theories of information use in sentence processing.

One such structural ambiguity, which is the tool used in the current study, is the so-called direct object/sentential complement (DO/SC) ambiguity. This ambiguity arises at the post-verbal noun phrase (NP) in sentences such as *The woman heard the dog had barked all night*. In sentences such as these, *the dog* (at the point where it occurs) could be either the DO of *heard* (as in *The woman heard the dog.*) or the subject noun of an SC, as it ends up being in the first example sentence. This ambiguity arises because the complementizer *that* may optionally be omitted in such constructions. The true structure is thus not revealed until the verb in the SC (at *had barked* in the sentence above). This is referred to as the disambiguation region.

Verb subcategorization biases and ambiguity resolution

Although some verbs occur in both SC and DO constructions with equal likelihood, many exhibit a bias. The notion that subcategorization preferences such as these might play a role in sentence processing has been considered by a number of researchers (Clark & Clark, 1977; Connine, Ferreira, Jones, Clifton, & Frazier, 1984; Ferreira & McClure, 1997; Fodor, 1978; Ford, Bresnan, & Kaplan, 1982; Fodor & Garret, 1967). Some researchers, such as Ford et al., have suggested that comprehenders use their knowledge of the relative probability with which a verb occurs with different subcategorizations to guide syntactic analysis, though others, such as Frazier (1987a, 1987b) and Ferreira and Henderson (1990) have reported patterns of results consistent with the notion that lexically specific knowledge is used only in the revision process.

In more recent years, the role of verb bias in comprehension has been subject to careful scrutiny in a number of studies involving the DO/SC ambiguity. Although some studies report late or no effects of verb bias (Ferreira & Henderson, 1990; Mitchell, 1987), more recent work has shown early effects (though see Kennison, 2001). Trueswell et al. (1993) contrasted sentence pairs that were structurally ambiguous at the post-verbal NP, and differed only in the bias of the main verb. These sentences were then contrasted with structurally unambiguous versions (i.e., containing the complementizer *that*), and the additional time taken to read the ambiguous version was referred to as the ambiguity effect. In a self-paced reading time experiment, Trueswell et al. found a large ambiguity effect for sentences containing DO-biased verbs at the point following the disambiguation toward an SC. In contrast, with SC-biased verbs, reading times for ambiguous sentences were similar to unambiguous controls. The bias by ambiguity interaction was replicated in an eyetracking experiment, although in this case it was found at the disambiguation itself, rather than in the post-disambiguation region. Finally, although there was a 34 ms ambiguity effect for SC-biased verbs in first-pass reading times, this was nonsignificant. The increased reading times were attributed to a *that*-preference effect because they were correlated with lexically specific *that*-preference.[1]

Trueswell et al. (1993) used the same sentence frame for both a DO- and an SC-biased verb. The post-verbal NP was always a plausible DO for the DO-biased verbs (e.g., *The waiter confirmed the reservation was made yesterday...*) but rarely or never a plausible DO for the SC-biased verbs (*The waiter insisted the reservation was made yesterday...*). As a result, bias was confounded with DO plausibility, and the results may have been influenced by the degree of commitment to a semantically plausible or implausible parse (cf. Pickering & Traxler, 1998). Garnsey et al. (1997) addressed this issue in an eyetracking study crossing bias with DO plausibility. Verbs of each bias type (including a third, equi-biased condition) appeared in sentences in which the post-verbal NP was either plausible or implausible as a DO. Thus for the implausible conditions, the NP was syntactically ambiguous, but semantically anomalous if interpreted as a DO. An effect of verb bias was found nonetheless: At the disambiguation, reading times were longer in the ambiguous than the unambiguous conditions for DO-biased, but not SC-biased verbs. In addition, with DO-biased verbs, sentences with plausible DOs yielded a significant ambiguity effect whereas the effect for those with implausible DOs was not statistically reliable. Overall, then, recent results suggest that verb bias is one relevant source of information in guiding sentence interpretation.

Sense and structure

In the literature reviewed above, and in many computational models, bias is measured across all instances of a verb—for example, it has been proposed that a verb's lexical representation specifies conditional probabilities for each of the possible subcategorization frames in which it might occur (Jurafsky, 1996). Although this may be an accurate description of the end

[1] Garnsey, Pearlmutter, Meyers, and Lotocky (1997) did not find a correlation between *that*-preference and the ambiguity effect at either the post-verbal NP or the disambiguation.

result, it provides little insight into the question of how these probabilities may have developed. In addressing that question, it might be more accurate to view the probabilities as at least partially determined by the relationship between meaning and structure.

Although it seems unlikely that structural choice can be reduced entirely to verb semantics, linguistic research has shown that there is a complex and detailed relationship between verb meaning and verb subcategorization (Grimshaw, 1979; Levin, 1993; Pesetsky, 1995). As one example, Pinker (1989) has pointed out that subtle semantic distinctions (''narrow conflation classes'') between otherwise similar verbs often determine the sorts of syntactic structures in which the verbs may appear. Thus although both *throw* and *pull* may appear in the dative construction, differences in the manner of the motion specified by these two verbs result in *throw* but not *pull* being permitted in the double object construction. In a set of correlational analyses, Argaman and Pearlmutter (2002) have also shown that the subcategorization probabilities of a set of verbs and their derived nominals, which are presumed to share core semantic features, have similar subcategorization probabilities, supporting the view that subcategorization choice and semantics are associated.

In sum, linguists have carefully investigated interactions between meaning and subcategorization, and have shown that this knowledge plays a role in a competence grammar. This does not require, of course, that it affects performance as well; that is, comprehenders may or may not exploit this information in normal production and comprehension. In the present paper, we use off-line production and on-line comprehension to test whether this is so.

If verb bias effects do indeed reflect the comprehender's awareness of meaning-structure relationships, it is a rather striking fact that when a verb allows two subcategorization possibilities, the meaning often differs in the two cases as well. Thus, although *find* must take a DO when it is used to mean 'locate', when it is used to mean 'understand' it also—even preferentially—allows a SC. In terms of thematic roles, when one performs a concrete action (*He found the book on the table.*) a patient is likely to be specified. However, mental events or expressions of mental attitude are more typically followed by a propositional SC that describes the event or situation (*He found the plane had left without him.*), although a DO can be used in some circumstances (*He found nothing but confusion*). If the interesting relationship is between structure and a specific sense of a verb, not structure and the verb in the aggregate, then patterns of verb bias would be better described by considering the specific sense that is used. Previous research has found this using corpus analyses: Roland and Jurafsky (2002) analyzed three verbs that differed in bias between the Wall Street Journal and Brown, and found that their distribution of senses varied between the two corpora as well. Systematic differences in subcategorization probabilities were associated with the verbs' senses, resulting in different estimates of subcategorization bias. This suggests that subcategorization biases were based on the particular sense of the verbs as they were used in those contexts. Thus although studies to date have assumed that overall verb bias plays a role in guiding processing, we suggest that subcategorization preferences are sense-specific, and are not always insightfully measured by calculating a statistical profile over all senses of a verb. Before testing this suggestion, we first clarify what we mean by verb sense, and how it might be expected to influence processing.

Verb sense

Although there may be exceptions, the great majority of cases where different meanings of a verb are associated with different subcategorization frames involve polysemy. That is, these verbs exhibit highly related meanings, often with a more concrete physical sense and extensions to more abstract or metaphorical uses (Lakoff, 1987; Rice, 1992). As one example, the verb *grasp* has the concrete sense 'grip' (*She grasped the handrail.*) and the more abstract sense of 'come to understand' (*She grasped that he wanted her to be quiet.*). Different meanings of a polysemous word are referred to as *senses* of the word, to distinguish them from the distinct meanings of homonyms like *bank* (Klein & Murphy, 2001; Rodd, Gaskell, & Marslen-Wilson, 1999), although the distinction between senses and meanings is a matter of debate. We use the term *sense* as well since our interest is in whether different but related meanings of a verb differ in their structural biases.

If subcategorization profiles vary with verb sense, we might ask whether comprehenders are aware of such differences, and whether this knowledge plays a role in the processing of temporarily ambiguous sentences. This would be consistent with recent work arguing that structural ambiguity resolution is not only similar in principle to lexical ambiguity resolution, but that in fact structural ambiguities often depend on lexical ambiguities of various sorts (e.g., MacDonald, 1993; MacDonald et al., 1994). As one example, the main clause/reduced relative ambiguity hinges on whether a morphologically ambiguous verb is taken to be the past tense or a passive participle. The sense of the verb, on the other hand, changes minimally if at all in this case. In contrast, in many cases of the DO/SC ambiguity, resolving a structural ambiguity hinges on resolving the verb sense ambiguity. For example, understanding the sentence *She grasped her friend wanted to make a good impression on the new teacher* requires resolving the sense ambiguity concerning *grasp*. We focus on such cases, and argue that

understanding these temporarily ambiguous sentences depends both on activating the correct sense of the verb and on resolving the structural ambiguity. We return to this point after the self-paced reading experiment, with a discussion of how the various senses and associated subcategorization probabilities may be represented, and how the sense ambiguity may be resolved.

In the studies presented below, we first use corpora to calculate the subcategorization biases for specific senses of verbs, and show how these might differ from the verb's overall bias. We then use off-line completions and on-line self-paced reading to test whether comprehenders' expectations regarding the structure that follows a verb are influenced by the particular sense used in a given context.

Corpus analyses

We first investigated whether overall subcategorization preferences differed from sense specific preferences for a set of candidate verbs. This was begun with a multi-corpus analysis of the verbs' overall subcategorization preferences, followed by an analysis of subcategorization preferences contingent on verb sense.

Method

Twenty verbs were chosen that could take both SC and DO arguments, that were categorized in WordNet (Miller, Beckwith, Fellbaum, Gross, & Miller, 1990) as having more than one sense, and that were judged by the experimenters to show clear sense differences.[2] In nearly all cases, the difference involved a concrete sense (an action) contrasted with a more abstract one (e.g., mental event or statement-making) which was generally an extension of the first. All sentences containing these verbs were extracted from three written and one conversational corpora: the Wall Street Journal (WSJ), Brown Corpus (BC), WSJ87/Brown Laboratory for Linguistic Information Processing (BLLIP), and Switchboard (SWBD), respectively. The four corpora vary in size and genre: BC is a 1-million word balanced corpus, with content from a number of written corpora. WSJ is the same size (1 million words) but consists exclusively of Dow Jones newswire stories. Both the WSJ and BC corpora were parsed as part of the Penn Treebank Project (Marcus, Santorini, & Marcinkiewicz, 1993). SWBD is a spoken corpus (Godfrey, Holliman, & McDaniel, 1992) of which 1.4 million words was parsed as part of Treebank. BLLIP is approximately 30-million words, comprised of the three year Wall Street Journal collection from the ACL/DCI corpus. This corpus was parsed using methods developed by Eugene Charniak, Don Blaheta, Niyu Ge, Keith Hall, John Hale, and Mark Johnson of BLLIP. The first three parsed corpora, and the raw data for the fourth, are available from the Linguistic Data Consortium at the University of Pennsylvania.

The extraction was done automatically, using *tgrep* scripts modified from those generously provided by Doug Roland (Roland, 2002). The verbs were classified according to 20 subcategorization frames expanded from the set used by Roland and Jurafsky (1998). The fine-grained parse categories were then collapsed into the more general categories of DO, SC, and Other according to the summary given in Table 1.

Note that only finite embedded clauses (with or without the complementizer *that*) were counted as SCs. Infinitival complements were counted as Other, as were tensed complements headed with a *wh*-complementizer. Neither case is straightforward, but there were so few examples of either in the corpus data that the classification decision had no effect on the subcategorization patterns in the results. However, the question of how they might be classified is certainly open to debate, as is the question of their influence on the processing system. Considering the case of the Wh_S continuations, these are indeed embedded clauses, and raise the interesting question of whether they add to expectations for an SC continuation at the verb, despite the fact that they never contain an ambiguous post-verbal NP. For our verbs, however, only 34 instances were found in the Brown corpus, and 36 in the WSJ (1% of the data in each case). Because there were not sufficient examples to affect the bias counts, these were left in the Other category. The infinitival complements were largely raising structures, and there are good reasons for counting these as either DO or as SC. Although the post-verbal NP has a number of syntactic properties of an embedded subject, it also behaves in many ways as the DO of the main clause (hence the early analysis that these NPs were base-generated in the lower clause and 'raised' to the higher clause in the course of a syntactic derivation). Given the structural indeterminacy, and the fact that again there were too few examples to affect the outcome, these were counted as Other. For other reasons, passive constructions did not add to the DO count for our verbs; although conceptually the verb takes a patient in those cases, there is no overt post-verbal NP that may be ambiguously treated as DO or SC subject.[3] Conversely,

[2] Verb sense distinctions in WordNet were determined by lexicographers, using a variety of methods, which included consulting standard dictionaries and use of linguistic judgments (C. Fellbaum, Personal communication, June 13, 2002).

[3] If passive forms were added to the DO count, this raised the mean DO percentage for these verbs from 45% to 57%. in the Brown corpus. This did not change the direction of bias for any individual verb, although it did increase the transitivity count of the majority.

Table 1
Categorization of fine-grained parse categories

Category	Parse	Example
Direct object		
	NP	Neither **acknowledged** the gift
	NP_NP	When Giffen decided to **charge** him interest...
	NP_PP	The work **added** two beds to the hospital
	NP_That_S	Jack would have **bet** his life that...
	NP_Wh_S	They **accepted** it because...
		He **held** the controls where they were
	Perception complement	I could **feel** the hair stand up...
Sentential complement		
	That_S	The Russian experimenters **claim** that only a small
	that-less-S	He **claimed** this was the favorite refrain
Other		
	NP_infin_S	I **found** it to be March 15th
	Infin_S, Infin_S_PP	The guerrillas **admitted** to being a little bit tired
	Wh_S	He **admits** what he does
		The British government will **decide** whether to let...
	Verb-ing	He **agreed**, acknowledging that...
	Nominal	They admitted to **betting** on the game
	PP	We **add** to their burden...
	0	When he **charged**, Mickey was ready.
	Passive	Three students were **admitted**...
	quote	Coombs has **declared**, "Two strong arms were..."

NP–NP structures were included because the post-verbal NP is an object of the verb and not resolved as subject of an embedded S. Therefore, even if the first NP is interpreted as a benefactive (*found Jeff a cheap plane fare*), it does add to the expectation of an NP continuation following the verb.

Results and discussion

Subcategorization probabilities were computed for each verb in each of the four corpora. The total number of tokens of the 20 verbs varied greatly across corpora, from a maximum of approximately 40,540 in WSJ87/BLLIP, to 3,000 in Brown and the WSJ, and fewer than 500 in SWBD. The frequencies of individual verbs varied widely as well, although the least frequent had a minimum of 40 tokens in the WSJ87/BLLIP corpus.

Subcategorization probabilities for the 20 verbs are given in Appendix A. Overall, these verbs were classified as DO-biased in the Brown corpus. In the other corpora, however, the overall count was roughly equi-biased. This is consistent with previous studies, which have found differences in subcategorization probabilities across corpora (e.g., Roland & Jurafsky, 1998; Roland et al., 2000).

We then computed sense-contingent bias counts to test whether these would differ from the overall probabilities. For each of the 20 verbs, we identified two senses that appear to be sufficiently distinct, that we believe are known to undergraduates, and that allow different subcategorization frames according to WordNet. We searched WordNet's Semantic Concordance for the two senses of each target verb. The Concordance consists of a subset of the Brown corpus, with all content words tagged for sense. This allowed us to extract all sentences containing the relevant verb senses. The sentences in the Concordance are not parsed, so these examples were then classified by hand into the categories presented in Table 1.

Because the Concordance consists of a subset of the Brown Corpus, there were no or extremely few sense-specific examples for eight of the target verbs. We analyzed the remaining 12, which had a minimum of five tokens of each sense ($M = 20$, range = 5–82). Table 2 shows that these exhibited reversing subcategorization biases based on the sense: although both structures were possible in many cases, there were more DO continuations for one sense of the verb, and more SCs for the other. For comparison, Table 2 also includes the overall subcategorization percentages for the 12 verbs, as measured from the Concordance. The number of tokens making up these percentages will be smaller than for the overall Brown analysis, but as in that analysis, the DO structure is dominant for these verbs overall.

The corpus analyses show a probabilistic relationship between a verb's sense and its subcategorization preferences. The critical question then becomes whether comprehenders are sensitive to these probabilistic

Table 2
Overall and sense-contingent subcategorization biases in the WordNet semantic concordance for 12 of the 20 verbs

Condition	Subcategorization type		
	% DO	% SC	% Other
DO-biased sense	68	9	23
SC-biased sense	16	64	20
Overall	61	16	23

relationships between meaning and structure. To test this, we constructed a set of items based on the 20 polysemous verbs. Each item consisted of a context sentence intended to promote either the SC-biased or DO-biased sense (as determined by the corpus analyses), followed by a target sentence that contained the verb and continued with an SC. The items are presented in Appendix B. In the next section, four norming experiments are presented in which we explored the properties of our sentences. In Norming Experiment 1, subjects completed sentence fragments in which a pronoun plus the verb was presented without preceding context (*They found__*). This was conducted to establish baseline measures of the tendency to use either sense of the verb, and to measure the sense-contingent subcategorization preferences out of context. In Norming Experiment 2, the same fragments to the verb were presented following their associated contexts. This study established the degree to which subjects interpret the verb in its intended sense given the context plus the verb, as well as measuring sense-contingent subcategorization preferences in context. In Norming Experiment 3, the NPs were included in the fragment to measure whether they biased off-line productions in a systematic fashion. Finally, Norming Experiment 4 investigated the nature of the expectations generated from the contexts by having subjects complete fragments that were preceded by the contexts and ended at the subject pronoun prior to the polysemous verb. Thus, Norming Experiments 2, 3, and 4 constitute what McRae, Spivey-Knowlton, and Tanenhaus (1998) termed gated norming.

Norming Experiment 1: Completions from the verb out of context

Sentence completion norms exist for many of the candidate verbs (e.g., Garnsey et al., 1997). However, because intra-sentential context may influence completion probabilities, completion norms were collected to ensure that verb bias measurements were accurate for the specific sentence fragments used in this study. More importantly, we calculated the percentage of completions in which each verb was used in each sense, as well as sense-contingent subcategorization preferences, neither of which have been measured previously.

Method

Subjects

Thirty-seven native English-speaking undergraduates from the Psychology and Cognitive Science subject pools of UC San Diego participated for course credit.

Materials and procedure

A single list was created containing sentence fragments using the 20 target verbs. Each fragment consisted of an optional adverbial (in four of the items), a pronominal subject, and the verb in its past tense form (i.e., each target sentence in Appendix B presented without the context and only to the verb; e.g., *They found __*). An additional 40 filler fragments of various syntactic structures were created; these were truncated at varying points. The fillers were added so that test items never followed one other in succession.

Results and discussion

Completions were categorized in two ways. First, based on the overall meaning of the completed sentence, the verb's sense was judged as corresponding to the SC-biased or DO-biased sense (as established in the WordNet Concordance corpus analyses). Second, for each sense, each completion was categorized as SC, DO, or Other based on the classifications in Table 1. Thus we computed the percentage of trials on which our target sense was used and sense-contingent subcategorization preferences, both of which are presented in the first line of Table 3, along with the results of Norming Experiments 2 and 3 for comparison.

The verbs out of context were nonsignificantly biased toward the DO sense (52% DO sense vs. 41% SC sense), $t_2(19) = 0.83$, $p > .2$.[4] The remaining 7% of the completions either used an alternate sense (e.g., the mathematical sense of *add*), or the completion did not allow for sense to be judged accurately. If a minimum of 60% of completions for a given verb used one sense, we counted that sense as dominant. By this criterion, the

[4] For all inferential statistics, $p < .05$ unless otherwise noted. F_1 and t_1 refer to analyses by subjects, whereas F_2 and t_2 refer to analyses by items.

Table 3
Percent use of DO-biased or SC-biased sense of verb in completion norms, by context, as well as percent sense-contingent use of SC- and DO-structures

Norming experiment	SC-biased sense			DO-biased sense		
	% Use of sense	% SC structure	% DO structure	% Use of sense	% SC structure	% DO structure
1. Verb, no context	41	54	32	52	1	85
2. Matching context plus verb	89	71	15	76	3	89
3. Matching context plus verb and NP	84	82	2	84	1	99

DO-biased sense was dominant for eight of the 20 verbs, while the SC-biased sense was dominant for 4.

Table 3 also shows that the DO sense is strongly DO-biased in completions out of context (85% DO vs. 1% SC), $t_2(18) = 14.74$.[5] The SC sense was not as strongly biased (54% SC vs. 32% DO), $t_2(19) = 1.61$, $p < .07$. This occurred because a number of the verbs in their SC sense allow DOs, some of which are rather colloquial (e.g., *grasped the concept*), and subjects were free to use these completions when no context was provided. In part for this reason, and in part because subjects tend to prefer shorter (DO) completions, the DO sense is more strongly biased in these norms than in the sense-contingent analysis of the corpus data, while the bias for the SC sense is comparatively weaker (compare the first line of Table 3 to the corpus results in Table 2).

We also calculated overall verb bias. If the sense of the verb is disregarded, 62% of completions used a DO and 20% used an SC. These results differed from those in the corpus analyses as well—there was a higher percentage of DO completions than in any of the corpora. This difference is consistent with Roland and Jurafsky (1998) and Merlo (1994) who found that completion data tend to inflate the proportion of DO structures relative to written corpora—and in fact the completion percentages in the current study are quite similar to results for the same items in earlier completion norms (Garnsey et al., 1997).

Norming Experiment 2: Completions from verb in context

This study was identical to Norming Experiment 1 except that one-sentence contexts preceded each sentence fragment. These contexts were designed to establish a scenario in which one or the other of the verb's target senses was more plausible, or followed more naturally. Thus, when the verb is read, the context-appropriate target sense should be more strongly activated, and the sense-contingent subcategorization probabilities should then influence completions.

There were two goals in this study: First, to estimate the degree to which the context promotes the intended sense of the verb; and second, to estimate sense-contingent verb subcategorization preferences when the verb was presented with preceding context.

Method

Subjects

Eighty native English-speaking undergraduates from the Psychology and Cognitive Science subject pools of UC San Diego participated for course credit, 40 per list.

Materials and procedure

The same 20 polysemous verbs were used. Pairs of one sentence contexts were constructed for each verb, one biased toward the DO sense, and the other toward the SC sense. There were no SC structures in the context sentences, particularly to avoid structural priming when the context was SC-biased. In addition, SC-biased verbs were avoided in the context sentences, regardless of the structure in which they occurred. The few that did occur were equally distributed across contexts (three in the DO contexts, four in SC). It is impossible to avoid simple transitive structures, but these were equally distributed across the two contexts.

Each context sentence was followed by a sentence fragment ending at the target verb; this initial target fragment was identical in both contexts (and was identical to that used in Norming Study 1). Subjects were asked to complete the fragment. An example sentence pair is presented below for the verb *find*. The complete set of items is presented in Appendix C.

SC biasing context (sense: REALIZE):

The intro psychology students hated having to read the assigned text because it was so boring. They found ____

DO biasing context (sense: LOCATE):

Allison and her friends had been searching for John Grisham's new novel for a week, but yesterday they were finally successful. They found ____

Items were rotated across two lists so that each subject saw each verb in a single context. Each subject saw half of the verbs following a DO-biasing context, and half following an SC-biasing context. An additional

[5] Note that this test involved only 19 verbs because no subject used the librarian sense of *recall*, as in *recall the book because it was overdue*.

20 filler items of varying syntactic structures were created, with the second sentence of each filler (the fragment to be completed) truncated at varying points. These were added to each list so that test items never followed each other in succession.

Results and discussion

An underlying assumption in the norming study was that the combination of the context and a polysemous verb would lead subjects to prefer one or another of the verb's senses, so that the completions would reflect biases associated with that sense. To measure these sense-contingent biases, the norming responses were first scored for use of verb sense based on the overall meaning of the completed sentence. As shown in Table 3, the target sense was used in the majority of the responses in both contexts, though more often following the SC-biasing than the DO-biasing contexts (89% vs. 76%), $t_2(19) = 2.21$. The effect of the contexts in promoting verb sense can also be measured by comparing completions to Norming Experiment 1, in which no contexts were used. In both the SC and DO cases, the intended sense was used more often following the context than with no context; SC: $t_2(19) = 7.50$, DO: $t_2(19) = 4.19$.

Subcategorization probabilities were then calculated for those sentences that used the intended sense. The subcategorization tendencies of the 20 verbs reversed with sense, as they had in the corpus analyses and the norms without context. There was a significantly greater percentage of SC than DO completions when the SC-biased sense of the verb was used (71% SC vs. 15% DO), $t_2(19) = 6.30$, and the opposite pattern obtained with the DO-biased sense (89% DO vs. 3% SC), $t_2(19) = 21.13$. Again, when used in the DO sense, the verbs can rarely take an SC, but when used in the SC sense, a DO is generally possible. Finally, these norms can be used to estimate *that*-preference when an SC is used. In the SC context, when the SC sense was used with an SC, the verbs showed a strong *that*-preference ($M = 82\%$, range = 61–100%, $SE = 3\%$). The individual verbs and their completion percentages are given in Appendix C.

A comparison to the norms without context shows that the SC contexts had a greater influence on sense-contingent structural biases than the DO contexts did. In the DO contexts, percent sense-contingent DO completions increased only 4%, and SC completions increased 2% (both nonsignificant). However, in SC contexts, sense-contingent SC completions increased 17%, $t_2(19) = 2.24$, and DO completions decreased by the same amount, $t_2(19) = 2.65$. The stronger effect of context in the SC case is not surprising. First, out of context there was little to constrain responses except the subjects' imagination, and the responses contained a large number of DOs that were very typical, almost idiomatic for these verbs (for example, *declare war* occurred 10 times in 37 responses). This was true even when the SC sense was used because sense does not entirely constrain the structure. However, when context was provided, the subjects restricted their responses to those that fit the context. As a result, the responses no longer contained typical DOs if these did not fit the context—to continue the example of *declare*, no performatives occurred in the SC context (severely reducing the number of DO responses compared to the out-of-context norms).

This is not to say that the context selected for a specific structure in the response—perhaps a more accurate description is that the context-appropriate sense of the verb fit into a different, or more restricted, set of frames, in the sense of Fillmore and colleagues (e.g., Johnson & Fillmore, 2000). Again, this does not absolutely determine the structure—for example, the occurrence of the verb in one of many conversational frames does not require that the message be expressed (in an SC). On the other hand, the elements that are not consistent with that frame (in that context) are unlikely to occur as DOs. And in fact any DOs that were given in the SC ('state') sense of *declare* were topics referring back to the context sentence, such as *"declared their opposition"* and *"declared a formal protest"* (to the suggestion being made in the context).

Norming Experiment 3: Completions from the NP in context

This study was identical to Norming Experiment 2 except that each sentence fragment ended at the post-verbal NP, rather than at the verb. The goal was to estimate the degree to which the NP altered both the perceived sense of the verb and the sense-contingent verb subcategorization preferences.

Method

Subjects

Thirty-four English-speaking undergraduates from the Psychology and Cognitive Science subject pools of UC San Diego participated for course credit, 17 per list.

Materials and procedure

The same items, lists, and fillers were used as in Norming Experiment 2, except that the target sentence fragments ended at the post-verbal NP.

Results and discussion

The fragments to be completed were actually not fragments at all, but full sentences (e.g., *He admitted the*

students; She acknowledged her son). Somewhat surprisingly, however, subjects always provided a continuation. Even when the post-verbal NP was taken to be the direct object, subjects added an adverb (*claimed the honor joyfully*), added a prepositional phrase (*added their names to the list*), modified the NP (*reported the students that were tardy*), or, most commonly, continued with a conjoined VP (*bet his house and lost*).

As shown in Table 3, the NP strengthened the bias in the DO context (as compared to Norming Experiment 2). There was an 8% increase in the use of the intended DO sense (to 84%), $t_2(19) = 1.77$. There was also a stronger sense-contingent structural bias in that DO completions increased to 99% from 89%, $t_2(19) = 3.27$, and SC completions decreased to 1% from 3%, $t_2(19) = 1.40$, $p < .09$. Therefore, as would be expected, the introduction of a plausible NP increased DO responses when the context and other constraints biased subjects toward a DO interpretation.

The NP resulted in a nonsignificant 5% decrease in use of the SC sense (to 84%) following SC-biasing contexts, $t_2(19) = 0.81$, $p > .2$. In terms of individual items, 7 of 20 verbs decreased in SC sense use when the NP was included in the fragment, with a corresponding increase in DO completions for these items. Apparently in these cases, the NP was a sufficiently good direct object to promote the DO sense of the verb, and the DO structure was used in these completions. On the other hand, five verbs showed at least a 10% increase in SC sense use, and across all 20 items the sense-contingent SC structural bias increased in strength (82% SC with the NP vs. 71% without), $t_2(19) = 1.74$. Overall, therefore, the completions show that the NP did not greatly disrupt the SC-biasing context items, but instead was generally taken as a plausible SC subject.

Norming Experiment 4: Completions from the preverbal pronoun in context

This was identical to Norming Experiment 2 except that each sentence fragment ended at the preverbal pronoun, rather than at the verb. The goal of these norms differed from the previous three norming experiments. Here the norms were used to estimate the extent to which the contexts performed their intended role of establishing a scenario more consistent with one sense of the verb.

Method

Subjects

Eighty English-speaking undergraduates from Bowling Green State University participated for course credit, 40 per list.

Materials and procedure

The same items, lists, and fillers were used as in Norming Experiments 2 and 3, except that the target sentence fragments ended at the preverbal pronoun.

Results and discussion

The completions showed that the contexts do not invoke strong expectations for the specific target verbs. The target verb itself was used in only three of the 40 cases: *found* was used in 47% of completions following the DO-context, *recognized* in 32% of completions following the DO-context, and *admitted* in 29% of completions following the SC-context. Thus, mean cloze probability for our target verbs was low (.05 for DO contexts and .01 for SC). Compare these results to a study such as Federmeier and Kutas (2001) who had subjects provide noun phrase completions for items such as, "*They wanted to make the hotel look more like a tropical resort. So along the driveway, they planted rows of* ____." The expected target for their sentences had a mean cloze probability of .74. The cloze probability of our verbs more closely matches that of their unexpected targets, which were defined as having cloze probabilities less than .05 (as did our verbs in 37 of 40 cases). Part of the reason for this difference is, of course, that the short fragments (to the pronoun) in our items did not provide the constraining intrasentential context of Federmeier and Kutas' relatively long fragments. In essence, there are numerous ideas that subjects can express given the context plus a fragment to the pronoun, as well as numerous ways in which an idea can be realized linguistically.

Clearly, then, the contexts do not result in strong expectations for a specific verb. But this was not the role of the context sentences. Instead they were intended to promote a scenario, part of which involved the activation of a semantic space that included the target verb sense. And in this respect they were successful: for 65% of the SC and 80% of the DO contexts, subjects' completions described a scenario consistent with the context and the target sense of the verb (although the verb itself rarely was used). For example, for the DO-biased sense of *acknowledge* (as in *acknowledged her son*), the context plus sentence fragment was, "At the dinner table, Kenny was eager to comment on the plans for the family trip, but his mom paid no attention to him. Finally, though, she." A completion that was counted as describing a consistent scenario was, "turned to him and asked his opinion." Furthermore, semantically similar verbs were used in 24% of the completions for the SC contexts (e.g., *confessed* for *admitted*) and 31% for the DO contexts. Finally, the results of Norming Experiments 2 and 3 indicate that the majority of completions did involve the verb's target sense once the verb was provided (see Table 3).

As was pointed out above, a critical issue for the contextual manipulation is that the contexts activate the appropriate sense of the verb, rather than directly activating a structure. Therefore, we were careful to control for verbs and constructions that might structurally prime a specific subcategorization frame. In addition, there was also the possibility that SCs might be primed directly by the context alone, independent of verb sense. A number of the contexts were designed to produce an expectation for a conversation or exchange of ideas (as opposed to a physical action, for example), and this alone might have primed an SC structure. We note that it is impossible to promote the various senses of some of our verbs without indirectly probabilistically promoting corresponding structures. And, in fact, there is no reason to assume that this sort of constraint would not be a factor in comprehension and production. However, the completions suggest that even if the contexts did promote SCs, verb sense did so more strongly.

There are two reasons for claiming this. First, not surprisingly, there was a significantly greater percentage of SC completions following SC contexts (23%) than following DO contexts (5%), $t_2(19) = 4.43$. The opposite is also true: There were significantly fewer DO completions following SC contexts (23%) versus DO contexts (40%), $t_2(19) = 5.12$. Note, however, that for all but two items, these were the cases in which the target verb or a semantically similar verb was used. More importantly, comparing completions following SC contexts in Norming Experiments 2 versus 4 shows that the verbs' sense and sense-contingent subcategorization preferences influenced completions over and above the influence of the contexts. In Norming Experiment 4, when the target verb was not available, completions in the SC context were equi-biased; 23% DO and 23% SC (so that completions using Other structures were actually most frequent, 64%). However, providing the target verb (Norming Experiment 2) resulted in a significant increase in SC completions (71% vs. 23%), $t_2(19) = 9.04$, and a marginally significant decrease in DO completions (15% vs. 23%), $t_2(19) = 1.45$, $p < .09$. Therefore, although the contexts presumably did directly promote the SC structure to at least some extent, the influence of the sense-contingent structural information associated with the verb was stronger.

Experiment: Moving-window self-paced reading

Two main observations can be drawn from the corpus analyses and off-line norming studies. First, for at least the verbs studied here, there are reliable associations between verb sense and subcategorization probabilities. These results are consistent with the corpus analyses of Roland and colleagues. Second, in off-line sentence completion tasks, context can bias subjects toward a specific sense of a verb and thus toward syntactic structures that are more highly associated with that sense. What remains unknown is whether people's knowledge of verbs' sense-contingent subcategorization probabilities is used during on-line processing. In this experiment, we used self-paced reading to assess whether this information influences how readers resolve temporary DO/SC ambiguities on-line using target sentences that include an SC.

Although we are focusing on the question of whether a verb's sense-contingent structural tendencies are important for ambiguity resolution, this is only one of a number of probabilistic constraints that may interact in complex ways in our items. We list these constraints here, but defer a description of how they may interact to the Discussion. First, independent of our particular items, there is a global structural preference favoring a DO following *noun verbed* in English (Bever, 1970). This transitivity bias can be taken to be a probabilistic version of Minimal Attachment (Frazier, 1987b). Second, the contexts tend to promote a certain type of event, and this type of event probabilistically correlates with both verb sense and the possible ensuing structure.

Third, when the verb is read, it acts as a further constraint: The context-appropriate sense should become more highly activated, and sense-contingent subcategorization biases should then become available. Fourth, in the unambiguous versions of the target sentences, the complementizer is a strong cue for both the SC structure and the SC-correlated sense of the verb—and conversely, its absence in the ambiguous versions cues the DO structure and potentially cues the verb's DO sense as well (Juliano & Tanenhaus, 1994). In the SC-biasing contexts of Norming Experiment 2, all 20 verbs had a *that*-preference of at least 61%, so the complementizer is expected and difficulties may result when it is absent.

Fifth, a comparison of Norming Experiments 2 and 3 shows that the NP provides a bias toward the DO sense and structure for target sentences following DO contexts, but is relatively neutral following SC contexts. Finally, the remainder of the target sentence, the disambiguating region and beyond, constrains readers both structurally and semantically toward an SC because it continued the SC, and included the main verb of the SC (see the example in the Materials below).

The implications for the reading time study are as follows. For ambiguous DO-context items, the global transitivity bias, context, sense-contingent subcategorization preferences, lack of a *that*, and the post-verbal NP all constrain readers toward a DO interpretation. Therefore, we predict that readers should read through the NP region interpreting the sentence in this way. However, the disambiguating region (and beyond) provides constraints supporting an SC interpretation, so readers should show a large ambiguity effect as this

information accumulates. For ambiguous SC-context items, the global transitivity bias and lack of a *that* constrain readers toward a DO interpretation. These constraints conflict with the information from the context and the sense-contingent subcategorization preferences, which bias readers toward an SC interpretation. The NP appears to be a relatively neutral cue when averaged across items. At the disambiguating region and beyond, the semantic and structural information again favor an SC. Therefore, particularly in comparison to the unambiguous version containing *that*, we expect some influence of the conflicting constraints but readers should not show a great deal of difficulty when the sentence disambiguates toward an SC. As a result, it is reasonable to expect a context by ambiguity interaction in the disambiguating region, as has been obtained in previous studies (e.g., Garnsey et al., 1997; Trueswell et al., 1993).

Method

Participants

Forty-five native English-speaking undergraduates from the University of Western Ontario were paid for their participation. One subject was dropped for having an excessively high error rate (30%) on the comprehension questions, and was replaced, leaving 11 subjects per list.

Materials. The stimuli consisted of 20 verbs with their associated SC- and DO-biasing contexts. SC-continuations were constructed that were sensible for both contexts. Unambiguous control sentences were created by adding the complementizer *that* after the main verb, as shown in the examples designated (ii). An example stimulus set is shown below for the verb *find*.

(a) *SC biasing context* (sense: REALIZED):

(i) The intro psychology students hated having to read the assigned text because it was so boring.

(ii) They found (that) the book was written poorly and difficult to understand.

(b) *DO biasing context* (sense: LOCATED):

(i) Allison and her friends had been searching for John Grisham's new novel for a week, but yesterday they were finally successful.

(ii) They found (that) the book was written poorly and were annoyed that they had spent so much time trying to get it.

The properties of context sentences (i) were presented in Norming Experiment 1. The second sentence (ii) contained the critical verb and took the form shown in the examples above: (1) an optional adverbial linking the sentence with the discourse context; (2) a pronominal subject coreferential with a context participant; (3) the past tense form of the verb; (4) a full NP that is plausible as both a direct object of the DO-biased sense and a subject of the SC; and (5) a two-word disambiguating region. In 16 of the 20 cases, the first word of the disambiguating region consisted of an auxiliary or copula (*was* in the example above), and in 15 of the 20 the second word of this region was a content word such as the main verb (*written*, in the example), predicate adjective, predicate nominal, or adverb. Context and target sentences were created so that the target sentences were equally coherent following both biasing contexts, and the pattern of results presented below indicate that this was successful. Context-target sentence pairs also were constructed to equate for the recoverability of the antecedent of the target sentences' post-verbal NP from both context sentences. Of the 20 target-context pairs in each biasing condition, 12 involved direct reference, with a recoverable antecedent synonymous with the NP (e.g., *his mom's sister*/*his aunt*). Indirect reference (a non-synonymous recoverable antecedent) occurred three times in the SC-biasing context, and five times in the DO-biasing context (e.g., *Ted*/*her friend*). In the remaining items (five SC-biasing contexts, three DO) the referent was implied by the context (e.g., see items 16 and 20 in Appendix B). Sentences were identical through the second word following the disambiguation point, so that critical regions were always identical across conditions. After this point, sentences sometimes varied to allow both versions to make sense in context.

Items were rotated across four lists, so that each subject encountered each test verb only once, and encountered six items from each condition. No test verbs were included in the context for another verb, nor in any filler or practice item. Forty two-sentence filler items of varying syntactic structures were created. These were added to each list so that test trials never occurred in sequence. In addition to the test and filler items, there were 20 two-sentence practice trials before the main experiment. There were no SC structures in any filler or practice item. Each trial was followed by a yes-no comprehension question that required the reader to understand the two sentences in combination.

Procedure

The experiment was conducted on a Macintosh computer using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). A CMU button box was used to collect decision latencies, providing records that were accurate to the nearest millisecond. Each trial began with the context sentence presented in its entirety on the computer screen, and this remained until the subject pressed a response button to indicate that she had read it. At this point, the context sentence was replaced by the target sentence, which was left-justified and presented with each letter replaced by a dash (–). All critical regions of the target sentence plus a minimum of one word were included on the first line. Subjects read each target sentence in a one-word-at-a-time moving-window fashion in which they made their way through the text by pressing the response button to reveal the next word,

Table 4
Reading times (ms) for the self-paced moving window experiment

Condition	Region						
	that	the	book	was	written	poorly	and
SC-biased, no *that*		342 (10)	356 (18)	369 (17)	359 (14)	366 (17)	373 (25)
SC- biased, *that*	333 (13)	332 (8)	349 (18)	350 (14)	342 (14)	350 (14)	360 (15)
Ambiguity Effect		10	7	19	17	16	13
DO-biased, no *that*		339 (7)	350 (16)	372 (13)	422 (24)	369 (13)	374 (16)
DO-biased, *that*	357 (15)	345 (16)	362 (17)	364 (12)	356 (16)	370 (14)	374 (15)
Ambiguity effect		−6	−12	8	66*	−1	0

Note: Standard errors in parentheses.
* Significant by subjects and items.

with the previous word reverting to dashes. A yes/no comprehension question followed each item.

The main experiment was preceded by the 20 practice trials, after which the subject took a short break. Sentence pairs during the experiment were presented in blocks of 20, with a break following each block. The session lasted approximately 30 min.

Design

Analyses of variance were conducted on the reading latencies for the one-word regions described below. The factors of interest were context (DO-biasing vs. SC-biasing) and ambiguity (*that* present vs. absent). Both context and ambiguity were within subjects (F_1) and items (F_2). List was included as a between-subjects dummy variable and item rotation group as a between-items dummy variable to stabilize variance that may result from rotating subjects and items over lists (Pollatsek & Well, 1995). Effects involving these dummy variables are not reported.

The critical regions are shown below.

Separate ANOVAs were conducted for each region: (1) the verb (*found*); (2) the complementizer (*that*); (3) the determiner of the post-verbal NP (*the* in the example, otherwise *his, her,* or *their*); (4) the post-verbal noun (*book*); (5) the first word of the disambiguating region (*was*); (6) the second word of the disambiguating region (*written*); and (7, and 8) each of the first two words following the disambiguation (*poorly*, *and*).

Results

Reading times are presented in Table 4 (with standard errors), as well as graphically in Fig. 1. As is clear, the appropriate sense of the verb was promoted by context, and this modulated the expectation for the related subcategorization frame, resulting in a context by ambiguity interaction at the disambiguating region. At the second word of the disambiguating region, there was no difference between ambiguous and unambiguous reading times for sentences preceded by SC-biasing contexts, but reading times were significantly longer for ambiguous than for unambiguous sentences when the preceding context was DO-biasing. In addition, there was an early influence of context. Reading times for *that* were shorter when an unambiguous SC was preceded by an SC-biasing versus a DO-biasing context.

Verb region (found)

As would be expected at the verb given that the ambiguous and unambiguous versions did not differ at this point, context and ambiguity did not interact, $F < 1$ in both analyses. Planned comparisons showed that there was no ambiguity effect for SC-biasing (no *that:* $M = 351$ ms, $SE = 18$ ms; *that*: $M = 349$ ms, $SE = 18$ ms) or DO-biasing contexts (no *that*: $M = 325$ ms, $SE = 14$ ms; *that*: $M = 332$ ms, $SE = 14$ ms), all F's < 1. There was a main effect of context in that reading times were 31 ms shorter for verbs following DO-biasing ($M = 329$ ms, $SE = 10$ ms) versus SC-biasing contexts ($M = 350$ ms, $SE = 12$ ms), $F_1(1,40) = 6.28$; $F_2(1,16) = 5.85$. Finally, there was no main effect of ambiguity (ambiguous: $M = 338$ ms, $SE = 11$ ms; unambiguous: $M = 341$ ms, $SE = 11$ ms), $F < 1$ in both analyses.

Complementizer (that). Reading times for *that* were 24 ms shorter in sentences preceded by an SC-biasing context than for those preceded by a DO-biasing context, $F_1(1,40) = 4.88$; $F_2(1,16) = 5.70$. This is predicted if readers expect a SC following the SC-biasing contexts plus the polysemous verb, and if the SC-biased sense of the verbs exhibit a *that*-preference, as Norming Experiment 2 shows they do.

Determiner (the). Context did not interact with ambiguity, $F_1(1,40) = 1.24$, $p > .2$; $F_2(1,16) = 1.06$, $p > .3$. Planned comparisons showed a nonsignificant 10 ms ambiguity effect for SC-biasing contexts and a

They	found	(that)	the	book	was	written	poorly	and
	1	2	3	4	5	6	7	8

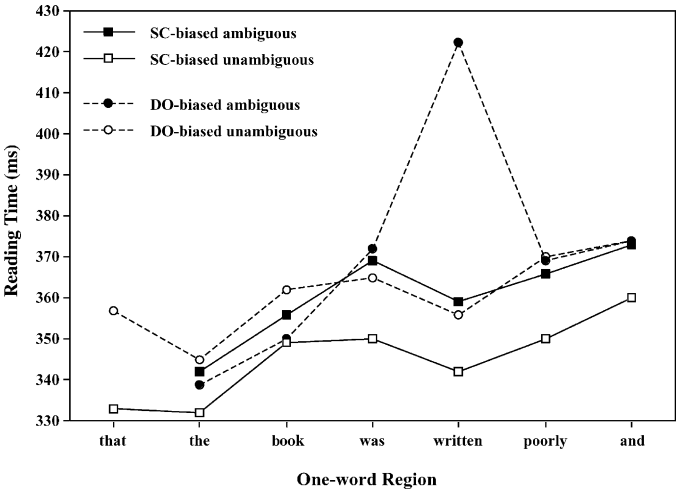

Fig. 1. Self-paced reading times from *that* through the post-disambiguating region.

nonsignificant −6 ms effect for DO-biasing contexts, all F's < 1. There was no main effect of context (SC-biasing: $M = 337$ ms, $SE = 10$ ms; DO-biasing: $M = 342$ ms, $SE = 10$ ms), $F < 1$ in both analyses. Finally, there was no main effect of ambiguity, (ambiguous: $M = 340$ ms, $SE = 9$ ms; unambiguous: $M = 338$ ms, $SE = 11$ ms), $F < 1$ in both analyses.

Noun (book)

Context did not interact with ambiguity, $F_1 < 1$; $F_2(1,16) = 1.31$, $p > .2$. Planned comparisons showed a nonsignificant 7 ms ambiguity effect for SC-biasing contexts and a nonsignificant −12 ms effect for DO-biasing contexts, all F's < 1. There was no main effect of context (SC-biasing: $M = 353$ ms, $SE = 12$ ms; DO-biasing: $M = 356$ ms, $SE = 12$ ms), $F < 1$ in both analyses. There was also no main effect of ambiguity, (ambiguous: $M = 353$ ms, $SE = 12$ ms; unambiguous: $M = 356$ ms, $SE = 12$ ms), $F < 1$ in both analyses.

First word of the disambiguating region (was). Context and ambiguity did not interact (both Fs < 1). Planned comparisons showed a nonsignificant 19 ms ambiguity effect for SC-biasing contexts, $F_1(1,40) = 2.72$, $p > .1$; $F_2(1,16) = 2.48$, $p > .1$, and a nonsignificant 7 ms effect for DO-biasing contexts, $F < 1$ in both analyses. There was no main effect of context (SC-biasing: $M = 360$ ms, $SE = 11$ ms; DO-biasing: $M = 368$ ms, $SE = 12$ ms), $F_1 < 1$; $F_2(1,16) = 1.69$, $p > .2$. Finally, there was no main effect of ambiguity (ambiguous: $M = 371$ ms, $SE = 12$ ms; unambiguous: $M = 357$ ms, $SE = 10$ ms), $F_1(1,40) = 2.37$, $p > .1$; $F_2(1,16) = 2.52$, $p > .1$.

Second word of the disambiguating region (written)

As predicted, context and ambiguity interacted, $F_1(1,40) = 5.77$; $F_2(1,16) = 9.50$. Planned comparisons revealed that the interaction occurred because reading times were a significant 66 ms longer for ambiguous than for unambiguous sentences when the context was DO-biasing, $F_1(1,40) = 20.63$; $F_2(1,16) = 34.02$, but a nonsignificant 17 ms longer when it was SC-biasing, $F_1(1,40) = 1.31$, $p > .2$; $F_2(1,16) = 2.17$, $p > .1$. Reading times for sentences preceded by an SC-biasing context ($M = 350$ ms, $SE = 10$ ms) were shorter than for those preceded by a DO-biasing context (DO-biasing: $M = 389$ ms, $SE = 15$ ms), $F_1(1,40) = 11.67$; $F_2(1,16) = 21.94$. Finally, there was a main effect of ambiguity in that reading times for ambiguous sentences ($M = 390$ ms, $SE = 14$ ms) were longer than for unambiguous sentences ($M = 349$ ms, $SE = 10$ ms), $F_1(1,40) = 16.36$; $F_2(1,16) = 7.97$.

Post-disambiguation region (poorly, and)

At the first word after the disambiguation (*poorly*), context no longer interacted with ambiguity, $F_1(1,40) = 1.55$, $p > .2$; $F_2(1,16) = 1.19$, $p > .2$. Planned comparisons showed a nonsignificant 16 ms ambiguity effect for SC-biasing contexts, $F_1(1,40) = 2.72$, $p > .1$; $F_2(1,16) = 2.09$, $p > .1$, and a nonsignificant −1 ms effect for DO-biasing contexts, $F < 1$ in both analyses. There was no main effect of context (SC-biasing: $M = 358$ ms, $SE = 11$ ms; DO-biasing: $M = 369$ ms, $SE = 10$ ms), $F_1(1,40) = 2.03$, $p > .1$; $F_2(1,16) = 3.11$, $p > .09$. Finally, there was no main effect of ambiguity (ambiguous: $M = 367$ ms, $SE = 11$ ms; unambiguous: $M = 360$ ms, $SE = 10$ ms), $F_1 < 1$; $F_2(1,16) = 1.94$, $p > .1$.

At the second word after the disambiguation region (*and*), there was no interaction between context and ambiguity, and the planned comparisons and both main effects were nonsignificant, all Fs < 1.

Error rates on comprehension questions

Error rates were averaged across fillers and targets to verify that subjects were paying attention to the task. Mean error rate was low ($M = 6\%$, $SE = 1\%$), and no individual subject had an error rate greater than 15%.

Discussion

Readers were influenced by structural expectations contingent on verb sense, where the verbs' sense was promoted by one-sentence contexts. The influence of the sense manipulation was shown early, at the complementizer, where reading times were significantly longer following the DO-biasing versus SC-biasing contexts. In addition, context interacted with ambiguity in the disambiguating region; there was a large ambiguity effect for sentences following DO-biasing contexts (i.e., the contexts that led readers to the verbs' incorrect interpretation and structural expectation) that contrasted with the lack of an ambiguity effect following SC-biasing contexts. This interaction is consistent with results of earlier studies contrasting DO-biased with SC-biased verbs (Garnsey et al., 1997; Trueswell et al., 1993).

We describe the main results using a constraint-satisfaction account based on a competition-integration framework, as developed by Spivey and colleagues (McRae et al., 1998; Spivey & Tanenhaus, 1998). We chose this framework because it is a useful vehicle for describing how the constraints may interact in our items. In this framework, available constraints combine in a nonlinear fashion to produce competition among alternative interpretations. A difference in reading times between ambiguous and unambiguous versions of sentences is considered to result from competition among constraints that support alternative interpretations. When the constraints strongly support a single interpretation, this interpretation is activated highly and there is little competition among alternatives, corresponding to the prediction of little or no ambiguity effects. At the other extreme, when the constraints are balanced among different alternatives, the activation levels of those interpretations are more equal and there is a great deal of competition, corresponding to a prediction of large ambiguity effects. Here we apply this model to our data, indicating what sources of constraint are assumed to be active at each word, and what their interaction would predict. We consider sentences in the DO-biasing and SC-biasing contexts in turn.

At the point that the verb (e.g., *found*) is read following the DO-biasing context, we assume at least four sources of information are available to combine and influence the interpretation: (1) a global bias towards transitive structures, which is arguably true of English in general and is evidenced in our verbs by the overall DO bias in corpora, (2) the verb sense, given the context, (3) the sense-contingent subcategorization bias, and (4) the fact that the verbs tend towards the DO-biased sense (as shown by the out-of-context norms). All four constraints strongly support a DO interpretation. The next two words, *the* and *book*, also are evidence for this structure, since they form a plausible NP object for the verb. Since all constraints support the same interpretation, there is no competition, and as stated above, this predicts little or no effect of ambiguity. In fact, the data at this point show a small negative ambiguity effect, consistent with that prediction. In terms of the model, at this point the DO interpretation should be highly active relative to the SC interpretation, and the overt NP is strong evidence against other potential interpretations, such as a prepositional phrase.

At the next point, contradictory evidence begins to become available. The first word of the disambiguation (*was*) is an auxiliary verb, and therefore a strong structural cue for an SC interpretation. This inconsistent constraint will lead to competition, which can be expected to begin reducing the activation of the DO interpretation. However, multiple active constraints (now including the absence of a complementizer) continue to support the DO. The second word of the disambiguation (*written*) is also a verb, and therefore a further strong constraint favoring an SC interpretation. Now the competing constraints are more balanced, leading to greater competition between alternative interpretations. As explained earlier, greater competition corresponds to a prediction of large ambiguity effects, and this is what is found in the data at this point.

Finally the post-disambiguation region (the continuation of the sentence complement) provides further strong cues for both the SC sense of the verb and the SC structure. At this point the ambiguity effect disappears quickly (i.e., there is little evidence of competition at *poorly*). Two facts would conspire to produce this effect: First, the first word of the post-disambiguation region (*poorly*) was the third consecutive input to provide a strong structural cue for an SC. Second, Williams (1992) has shown that the context-inappropriate sense of a polysemous word tends to remain active for some time (unlike context-inappropriate meanings of homonyms). As the constraints activated the SC structural interpretation, the activation of the SC sense would increase as well, and this in turn should combine with the structural information to facilitate the resolution of the structural ambiguity.

For the SC-biasing context items, the constraints suggest a different pattern. At the verb (*found*), the global transitivity bias supported the DO interpretation, but verb sense and sense-contingent subcategorization preferences (and, for example, any contextually induced expectations for statement-making) supported the SC. At this point there is some conflict among constraints, but the majority favor the SC interpretation, making it the more strongly activated alternative. In the data, this

interpretation is reflected in the shorter reading times at *that* for the SC compared to the DO-context items.[6]

The two words of the NP (*the* and *book*) add a new set of constraints. Norming Experiment 3 showed that these are relatively neutral cues, and collapsed across items would not change the interpretation, other than providing evidence against incompatible possibilities such as a post-verbal prepositional phrase. However, the presence of the NP also indicates the absence of *that*. The average *that*-preference for these verbs was 82%, and combined with the context effect at *that* shows that it is highly expected in this context. Its absence is therefore a constraint in favor of the DO interpretation. Although this constraint conflicts with the information available to this point, the majority of the constraints favor the SC, leading to little competition. And indeed, the ambiguity effects found at both words in this region were small and nonsignificant.

As in the DO-biasing contexts, the first word of the disambiguation region (*was*) provided a structural cue for both the SC sense of the verb and the SC interpretation. In this case, this cue adds to the majority of constraints supporting the same interpretation, and should not increase the level of competition. The data reflect this: although the ambiguity effect was numerically slightly larger at this point, it was again nonsignificant. Small ambiguity effects continued through the disambiguation and post-disambiguation regions, though at no point did they approach significance.

Thus the strongest constraints favored an SC interpretation, but these were countered by the global transitivity bias, the absence of a complementizer, and the fact that the contexts probabilistically cued the SC sense. Although these conspired to produce sufficient competition to result in small ambiguity effects throughout the SC-context sentences, they were never sufficient to override the dominant SC interpretation. Further evidence for this interpretation comes from correlations between *that*-preference and SC context ambiguity effects. Ambiguity effects correlated with *that*-preference at *the* ($r = .49$), *was* ($r = .47$), and *written* ($r = .47$), although not at *book* ($r = .27$) nor at either word of the post-disambiguation region. These results are consistent with the interpretation that the absence of *that* supported the DO structure and the verbs' DO-biased sense, leading to increased competition.

Finally, the context effect at the verb deserves comment. Subjects were significantly faster to read the verb following the DO context, collapsed across ambiguous and unambiguous conditions (which are identical to this point). Although this effect was unexpected and it is not entirely clear why it occurred, there are some indications in the norming data. First, Norming Experiment 4 (with context but without the verb) indicates that the verbs may be somewhat more highly expected or easily integrated following the DO contexts—at least 35% of subjects completed the fragment with the target verb or a semantically similar verb for 9 of 20 DO contexts versus 5 of 20 SC contexts. Second, the overall number of target or semantically similar verbs nonsignificantly favored the DO contexts. Third, the out-of-context norms (Norming Experiment 1) showed that the target verbs are nonsignificantly DO-biased in meaning. Finally, numerous studies have shown that the meaning of concrete words is activated more quickly than the meaning of abstract words (e.g., Schwanenflugel, Harnishfeger, & Stowe, 1988), although these studies have not focused on verbs. Given that in most cases the DO sense of our target verbs was more concrete than the SC sense, it is possible that this difference may have contributed to reading times at the verb. In summary, all of these small differences favor the DO sense of the verb, and in combination may have been sufficient to produce the effect.

General discussion

Several recent theories have emphasized the probabilistic nature of language knowledge, and there have been a number of demonstrations that processing phenomena can insightfully be described in terms of probabilistic constraints. This assumption figures prominently in constraint-based theories (e.g., Altmann, 1998, 1999; MacDonald et al., 1994; MacWhinney & Bates, 1989; Spivey & Tanenhaus, 1998) and is a central and explicit assumption in many computational accounts of lexical access and syntactic disambiguation (e.g., Jurafsky, 1996; Manning & Schütze, 1999; Narayanan & Jurafsky, 1998; Resnik, 1996). Knowledge of verb subcategorization biases is one example of a phenomenon that has been described probabilistically. Much recent work has argued successfully that knowledge of the relative probability of a verb occurring with a given subcategorization is used to guide syntactic analysis. To a first approximation, this is a reasonable account, and one with which we are very sympathetic.

One question that arises, however, is where these probabilities come from. One possibility is that for verbs that permit multiple subcategorizations, the subcategorization frames appear in free variation with each other, subject only to their individual probabilities. In this case, there would be no determinants of usage above and beyond the probability distribution. Another possibility

[6] Note that the central claim of the model is that all available constraints combine (in a non-linear fashion) to competitively influence alternative interpretations. Thus although it is generally used to predict ambiguity effects, it does so by modeling which interpretation is more active at each point. The strength of the interpretation also has implications for cases like the present one, where no ambiguity is present.

is that the probabilities reflect patterns of usage for which there are (perhaps only partially) determining factors. Thus, the choice of subcategorization frame is not determined directly by a probability distribution, but rather by these other factors, which might in turn be described probabilistically, and might or might not be controlled by yet other causes.

The present research moves toward an explanation of the source of subcategorization preferences by appealing to a semantic variable (cf., Argaman & Pearlmutter, 2002). The literature shows that verb bias reflects comprehenders' awareness of meaning-structure relationships, leading us to suggest that subcategorization biases may be influenced by variations in verb sense.

Our test for sense-specific subcategorization biases proceeded in several steps. We first calculated subcategorization biases across corpora for a set of verbs that allow multiple subcategorization frames, then showed, using WordNet's semantic concordance, that the subcategorization probabilities for these verbs differed by sense. Off-line sentence completion studies showed that a context could be used to promote a specific sense of a verb, and that our verbs showed clear sense-contingent subcategorization probabilities. This pattern demonstrates that subjects were indeed guided by sense-based knowledge of subcategorization preferences in the completion task. Finally, using the self-paced reading paradigm, we demonstrated that this knowledge rapidly influences readers' analysis of temporarily ambiguous sentences. The fact that a similar pattern can be found for identical verbs in differentially priming contexts strongly indicates that structural frame probabilities are based on specific verb senses.

The research that is most similar to the present study is that of Boland (1997, Experiments 3 and 4), who used intrasentential semantic information to attempt to bias comprehenders toward various senses of a verb. Subjects listened to sentence fragments such as, "Which salad did Jenny toss" or "Which baseball did Jenny toss" and either named or made a lexical decision to a proper name presented on a computer screen following the auditory fragment. Off-line norms showed a greater percentage of three-argument completions for the *baseball* fragment, and Boland reasoned that if people are sensitive to sense-contingent subcategorization preferences, and the sentence fragments are sufficient to activate the correct sense of the verb, then subjects should be faster to respond to the proper name following the *baseball* than the *salad* fragment. The hypothesized effect was found with the longest ISI (where the target proper name was presented 300 ms after the offset of the verb) of Experiment 4 (naming task), but was not found with −150, 0, nor 150 ms ISIs of Experiment 4, with a −150 ms ISI in Experiment 3a (naming task), nor with the same ISI in Experiment 3b (lexical decision task). It is difficult to ascertain why Boland failed to find the hypothesized difference, although there are a number of possibilities. First, the minimal contexts may not have sufficiently promoted the various verb senses in the absence of further supporting constraints. Second, comprehenders may have expected a preposition following the verb in the three-argument cases (e.g., *Which baseball did Jenny toss—to Bill?*), rather than the double-object construction, and there may have been a cost when there was none. Finally, with only ten items, the experiment may have suffered from insufficient power.

Lexical representations

We now turn to the question of how sense differences might be represented in order to account for the reading-time data. There are two major ways in which the lexicon tends to be construed. The first is a lexical entry view, which typically involves localist representations and relatively static processing (e.g., activation of those localist nodes). One way in which a static localist model incorporates polysemy is to include separate, linked entries for each sense (see Klein & Murphy, 2001; for a discussion). On this view, each sense of our verbs would correspond to a node, with each node including a subcategorization profile associated with that sense. That profile might be realized as frequency counts derived from experience with language, or as weights on links from a verb's sense-specific node to subcategorization frames (MacDonald et al., 1994). Because subcategorization profiles differ for a verb's DO-biased and SC-biased senses, accounting for the self-paced reading effects in a model such as this requires the preferential activation of the intended sense of the verb, and thus the preferential influence of the intended sense-contingent subcategorization profile. That is, when subjects encountered the verb in our target sentences, the context-appropriate sense should have become more highly activated than the inappropriate sense (as the results of Norming Study 2 show) and the associated subcategorization preferences then guided comprehension. The shorter reading times at *that*, which is the word directly following the verb, for SC versus DO contexts suggests that the time course of this process is short, as the intended sense was activated and used quite quickly. One issue that complicates the story concerns whether the less preferred sense is activated to some degree and thus influences structural expectations, or whether only one sense is considered at any single point in time. Boland (1997), for example, has argued in favor of a propose and select model, so that only one subcategorization profile would influence comprehension at any one time. In contrast, it is also possible that there may be some influence of the less activated sense node (and its subcategorization profile). However, we see no way that our data distinguish between these two possibilities.

Other researchers have argued that polysemous words, such as those exemplified by our verbs, do not have separate nodes for each sense (Caramazza & Grober, 1976). Instead, there is a single core meaning that can be shaded by context. In this case, under the assumption that a subcategorization profile is stored with a lexical entry, there would be a single amalgamated subcategorization profile for each verb, and this profile would capture structural statistics collapsed across multiple senses. The question then becomes whether an account of our results could be constructed that does not involve sense-contingent subcategorization preferences. The basic results in the DO-biased contexts might be explained without reference to sense differences: There is a global transitivity bias in the language, and out of context, our verbs are somewhat DO-biased. In addition, in the norms with the context but without the verb (Norming Experiment 4), without taking sense into account, 40% of the completions included a DO, whereas only 5% included an SC. Hence a tendency to expect a DO continuation could be explained on the grounds of frequency alone. The absence of ambiguity effects following the SC-biased contexts presents a more difficult challenge. The best candidate for an account in this case is one in which the contexts directly primed structure, irrespective of the verb. However, the same norms (Norming Experiment 4) do not show an SC bias following the SC-biased contexts without the verb. Instead the responses were equi-biased, with 23% DO and 23% SC completions. It is unlikely that the equivalent biases resulting from context would result in the lack of significant ambiguity effects following the SC-biased contexts given the other constraints that favor a DO interpretation. These results appear to be better accommodated by the bias toward the SC-structure found when subjects completed fragments including not just the context, but the verb as well.

A third view, and the one that we prefer, revolves around distributed attractor network accounts of semantic memory and dynamic context-dependent approaches to language processing (Elman, 1991, 1995; Tabor, Juliano, & Tanenhaus, 1997). On this view, there are separate but nearby basins of attraction for each sense (Kawamoto, 1993; Klein & Murphy, 2001). The distance in a high-dimensional semantic/syntactic space is determined by a number of factors, including semantic overlap and structural considerations, where structural knowledge on this view corresponds to trajectories into and out of attractor basins. There are pressures from orthography and phonology as well: The senses will be pulled closer in the high-dimensional space because their spelling and pronunciation are identical. Semantically, our verbs' senses tend to be highly related. For seventeen of the verbs, the contrast is between action, on the one hand, and mental events or statement making on the other, where the abstract meaning is often an extension of the concrete one. For the final three, the difference involved the extension from one abstract meaning to another. Thus on semantic grounds, the verbs will be structured closely together in space as well.

Note that on this view, frequency plays a different role in lexical representation than in some other accounts (e.g., Jurafsky, 1996; Narayanan & Jurafsky, 1998). In those accounts, frequency of usage (of a verb, of its possible subcategorization frames, of its various senses, etc.) are incorporated more or less explicitly into the lexical representation. Lexical entries thus record usage in a verbatim and literal manner. In the view we favor, usage is also important, but helps to shape representation through its effects on learning. The distinction might be thought of as the difference between "learning statistics" and "statistically driven learning." In the latter, representations are the outcome of inductive processes. Those processes permit generalizations that may go beyond the tabulated instances of usage experienced by a language learner. Note however, that there is nothing in the present data that is inconsistent with Narayanan and Jurafsky's (1998) view.

If one interprets the relationship between sense and structure in terms of the processing metaphor of moving through a semantic/syntactic space, then knowledge of subcategorization preferences is realized in terms of the probability of following various trajectories out of the attractor basin corresponding to a specific sense of the verb. The norming and self-paced reading results show that comprehenders tend to follow a trajectory that can be realized as an SC when coming out of the attractor generally residing in the part of space representing mental events or statement making. The role of the contexts, on this view, is to move the interpretation into the intended attractor basin for the verb.

Other determinants of subcategorization choice

Does sense always determine the choice of subcategorization? Undoubtedly not. We suggest rather that sense is only one of several factors that determine, for verbs that allow multiple subcategorization frames, which of several subcategorizations is actually used in a given utterance. The great majority of our verbs allowed DOs in our SC-biased sense, with no difference in verb meaning. Clearly, factors other than verb sense determine the use of the DO structure in what we have been referring to as the SC-biased sense. Similarly, a number of verbs like *believe* can occur in different frames and yet retain essentially the same sense in all of them: for example, there is no sense difference between *Norma believed Arthur* or *Norma believed what Arthur said*. Here the manner in which the message is conveyed linguistically might result from discourse considerations.

To this point, we have focused on a verb's subcategorization frame, and restricted our interest to the

grammatical structure of the verb's internal arguments. This is due partially to the practical limitations imposed by the currently available tools: Parsed corpora make it possible to identify only subcategorization frames. A more complete account will consider argument structure more generally, because information about the syntactic and semantic properties of all arguments, including the subject, undoubtedly plays a role in the expectations that are generated during reading. The current work attempts to go beyond what has been done to this point by considering the role of meaning in subcategorization bias, but we have nonetheless collapsed distinctions we feel are important, and should be considered in future research.

Insofar as context can affect argument structure choice, an open question is what aspects of context might be relevant. In this study, we focused on extrasentential context, which is known from prior work to influence on-going comprehension (Pickering & Traxler, 1998; Vu, Kellas, Metcalf, & Herman, 2000). It is also likely that intrasentential context can cue probable argument structures. There are two reasons for believing this.

First, in corpus data, there are strong correlations between argument structure and features of either the agent or patient, or presence of certain modifiers. Animacy is one high-level feature that plays an obvious role. If the subject of a verb like *worry* is inanimate, the normal completion in corpora is DO (*the book worried him*). For other verbs the opposite would be expected (*the book claimed that*...). Similarly, in the Brown Corpus, when *announce* is used with a DO structure, the agent is almost always making a formal announcement of an event, often with predictable adverbs. Thus, the sentence that begins *Mr. and Mrs. J. Reginald Smith, III, proudly announce* strongly predicts an event (in this case, a wedding or birth). Often these patterns reflect pragmatic factors or even social convention. It remains to be seen whether comprehenders exploit knowledge of these associations in resolving temporary ambiguities.

Second, it has been shown, with studies using single-word priming and cross-modal sentence fragment priming, that verbs immediately activate relatively detailed information about typical fillers of their agent and patient roles (Ferretti, McRae, & Hatherell, 2001). This suggests that verbs provide access to the generalized situation structure corresponding to the event to which it refers, and thus facilitate processing of participants associated with that event. Conversely, the participants themselves (agents, patients, instruments, locations) can serve as access points to the event. If a noun is a good agent for a verb, it activates an event structure to which that verb refers; empirically, this is realized as the ability of such typical participants to prime their associated verbs (McRae, Hare, Ferretti, & Elman, 2001). We suggest that, just as thematic role concepts are verb-specific (McRae et al., 1997), they may be sense-specific as well. The noun *criminal*, for example, may be a good agent for *admit* in its "confess" sense, but not for its "let in" sense. This leads to the prediction that one might find an agent-by-ambiguity interaction in sentences such as *The criminal admitted his friend* ... (with *friend* interpreted as the subject of an SC, e.g., ... *helped him rob the bank*) versus *The doorman admitted his friend*... (*friend* interpreted as a DO, e.g., ...*into the theatre through the back door*).

Another open issue concerns the grain-size over which probabilities are computed (or to which comprehenders are sensitive). It has been pointed out that disambiguation preferences do not always exactly reflect corpus frequencies (Gibson & Schütze, 1999; Gibson, Schütze, & Salomon, 1996). Some deviations may arise from pragmatic or discourse considerations (which might not be readily measurable in any corpus), or because of a failure to consider the relevant grain-size(s).

Earlier studies have found significant effects of verb bias, even when computing bias over all senses of a verb. Nothing in the current studies rules out the possibility that overall bias does play a role in guiding processing. First, and most obviously, if only one sense exists, then computing verb-specific and sense-specific biases is equivalent. In studies such as Garnsey et al. (1997) and Trueswell et al. (1993), the majority of the verbs fell into this category (by our estimates, in Garnsey et al., 11 of 16 DO-biased verbs, 10 of 16 equi-biased verbs, and 7 of 16 SC-biased; in Trueswell et al., 7 of 10 SC-biased verbs, and 5 of 10 DO-biased verbs). Second, the lexical ambiguity literature shows that the dominant sense or meaning of a word is more highly activated unless there is a context that is strongly biased toward the subordinate meaning. Therefore, for the verbs that do have both senses, as long as the verb was used in the condition that corresponds to its dominant sense and structure, and no intrasentential elements existed to strongly bias readers toward the subordinate sense, then the intended effects should be found. The existing literature appears to support this. For example, by our estimates, 9 of the 16 SC-biased verbs in Garnsey et al. (1997) possess both SC and DO-biased senses. However, although we cannot state definitively which sense is dominant, our intuitions suggest that the SC sense is dominant for all nine. For two of the nine verbs (*claim* and *decide*), it might be argued that intrasentential context favors a DO-biased sense, but for the majority this is not the case. The fact that Garnsey et al. found no ambiguity effect for their SC-biased verbs is consistent with the claim that their categorization matched that of the dominant sense.

Conclusion

The contribution of the present work is to establish that at least one of the relevant levels of consideration in

processing temporarily ambiguous subcategorization lies in verb meaning. All else held constant, when context is sufficiently strong to cue one sense of a verb, and when that sense is highly correlated with a specific subcategorization, comprehenders interpret a temporary ambiguity in a manner consistent with it. We view this result as an important step in understanding the more complex process by which this constraint interacts with the many other factors—including dominant verb bias, global word order patterns in the language, expectations generated by specific choice of thematic role fillers, and discourse considerations—that ultimately determine the syntactic structure of an utterance.

Acknowledgments

This work was supported by NIH Grant MH6051701A2 to all three authors, NSERC Grant OGP0155704 to the second author, and NFS DBSS92-09432 to the third author. We would like to thank Doug Roland, Dan Jurafsky, Maryellen MacDonald, and Susanne Gahl for their helpful discussions and comments.

Appendix A

Results of corpus analyses: Percentage of SC, DO, and Other (Oth) structures.

VERB	Brown			WSJ			WSJ87/BLLIP			Switchboard		
	SC	DO	Oth	SC	DO	Oth	SC	DO	Oth	SC	DO	Oth
Acknowledge	16	60	24	70	18	12	67	18	15	0	0	0
Add	10	32	58	41	29	29	31	17	52	0	61	39
Admit	29	24	47	42	30	28	39	30	31	64	36	0
Anticipate	0	55	45	13	63	24	14	45	41	33	33	33
Bet	44	11	44	54	13	33	47	16	37	61	9	30
Claim	33	34	32	69	22	8	55	20	25	71	29	0
Confirm	3	68	29	35	48	17	37	38	25	0	0	0
Declare	33	16	51	26	29	45	11	49	40	0	100	0
Feel	21	31	48	33	15	52	37	15	48	21	8	71
Find	9	60	31	23	47	30	18	46	37	19	45	36
Grasp	5	57	38	0	60	40	7	44	49	0	100	0
Indicate	39	33	28	58	27	16	66	19	16	100	0	0
Insert	0	59	41	0	43	57	0	40	60	0	0	0
Observe	16	30	54	46	17	38	21	26	53	0	0	0
Project	4	61	36	29	48	24	17	48	36	0	0	0
Recall	19	54	27	23	28	49	10	30	60	0	33	67
Recognize	9	65	26	20	52	28	27	48	25	14	86	0
Reflect	1	53	46	0	92	8	1	88	11	0	0	0
Report	16	18	66	22	61	16	16	41	44	0	67	33
Reveal	11	70	19	23	58	19	23	56	21	0	0	0
M	16	45	39	31	40	29	27	37	36	21	33	46

Note. Because individual items are rounded, percentages may not sum to 100%.

Appendix B

Stimuli used in moving-window self-paced reading experiment

1. observe

(SC) Trevor's teacher asked him to explain why there had been riots following the election in Bosnia.

(Target) He observed (that) the election had probably been rigged/the previous year and that is what caused the problems.

(DO) A United Nations official was sent to Bosnia to keep an eye on the election.

(Target) He observed (that) the election had probably been rigged/the previous year, so the UN wanted to make sure it wouldn't happen again.

2. admit

(SC) For over a week, the trail guide had been denying any problems with the two high school kids walking the entire Appalachian Trail.

(Target) Finally, though, he admitted (that) the students had little chance of/succeeding.

(DO) The two freshmen on the waiting list refused to leave the professor's office until he let them into his class.

(Target) Finally, he admitted (that) the students had little chance of/getting into the course.

3. recall

(SC) Mary Anne was happy to have time to sit down and read, but she couldn't locate the book she had started three days earlier.

(Target) She recalled (that) the novel was sitting underneath the/magazines on her coffee table, so she got it and sat down on her favorite rocking chair.

(DO) Two people had requested the overdue book, so the librarian agreed to get it for them right away.

(Target) She recalled (that) the novel was sitting underneath the/front counter, and that it had actually been returned earlier that day.

4. grasp

(SC) As Rhonda lay alone in her bed, she began to understand why Ted had become such a good student lately.

(Target) She grasped (that) her friend wanted to make a/good impression on the new teacher.

(DO) Rhonda saw Ted trip at the head of the stairs.

(Target) She grasped (that) her friend wanted to make a/fool of himself and had done it on purpose.

5. recognize

(SC) Joe had taken his mom's ailing sister into his home, and he wanted to keep her with him even though she wanted to move to a nursing home.

(Target) Finally though, he recognized (that) his aunt was sick and her/care would be better at the home.

(DO) When Joe opened the door, he did not immediately know his mom's sister.

(Target) Finally though, he recognized (that) his aunt was sick and her/appearance had changed dramatically.

6. indicate

(SC) Ken had finally allowed his landlady to rent out his garage space while he was in Europe for a year.

(Target) He indicated (that) the car was gone because he/had lent it to his nephew.

(DO) The day care worker asked the little boy to show her which toy he wanted.

(Target) He indicated (that) the car was gone because he/had let another kid have it.

7. add

(SC) Matthew was complaining to his wife about their kids' ridiculously busy schedule when he thought of one last thing to tell her.

(Target) He added (that) their kids were fine just playing/in the local soccer league with their friends and he didn't want them trying out for the traveling team.

(DO) Matthew asked his wife for a pen as the two of them stood in front of the sign-up list for the kids' traveling team.

(Target) He added (that) their kids were fine just playing/in the local league last year, so why not let them try out for the traveling team this year.

8. anticipate

(SC) Liz and George were reassured by their broker's projections after stock prices fell badly in August.

(Target) He anticipated (that) the market was going to fluctuate/but then prices would rise rapidly.

(DO) Unlike many people, George didn't lose any money when stock prices fell badly in August.

(Target) He anticipated (that) the market was going to fluctuate/and moved his money into bonds.

9. reflect

(SC) Maureen debated whether to park near the visitor's center and hike in from there to Mt. Shasta, or to try to get a spot in the more crowded lot nearer the base of the peak.

(Target) She reflected (that) the mountain might be too far/away if she parked in the first lot, so she took a chance and kept going.

(DO) When Maureen moved into her new apartment, she set up mirrors in her living room to try to get a good view of Mt. Shasta out her front window.

(Target) She reflected (that) the mountain might be too far/away to see clearly, but it was worth trying all the same.

10. acknowledge

(SC) For the past hour and a half, Susan had been bragging to her friend about her endless patience when it came to dealing with her curious 4-year-old son, even though it was something of a lie.

(Target) Finally though, she acknowledged (that) her son bugged her very often/because of his boundless energy and endless questions.

(DO) At the dinner table, Kenny was eager to comment on the plans for the family trip, but his mom paid no attention to him.

(Target) Finally though, she acknowledged (that) her son bugged her very often/because of his ridiculous off-topic comments, and that she had been ignoring him on purpose.

11. find

(SC) The intro psychology students hated having to read the assigned text because it was so boring.

(Target) They found (that) the book was written poorly and/difficult to understand.

(DO) Allison and her friends had been searching for John Grisham's new novel for a week, but yesterday they were finally successful.

(Target) They found (that) the book was written poorly and/were annoyed that they had spent so much time trying to get it.

12. bet

(SC) Anthony was deeply depressed about the damage to his property caused by the earthquake.

(Target) He bet (that) his house was going to be/worth much less than it used to be.

(DO) Anthony had experienced a string of bad luck in the high stakes poker game, but because he was holding such a great hand he decided to stay in, using his property as collateral.

(Target) He bet (that) his house was going to be/worth enough to let him stay in the game and to win back his money besides.

13. confirm

(SC) Roger's secretary asked him if he really did want to have tomorrow's meeting in the small conference room that was completely lacking any decent audiovisual equipment.

(Target) He confirmed (that) the room was precisely the right/one because there were only going to be five people at the meeting.

(DO) Roger called university classroom reservations to finalize the location of his course for this semester.

(Target) He confirmed (that) the room was precisely the right/one because there were only going to be five students in his course.

14. declare

(SC) At the meeting, the parents objected strongly to the principal's decision to have yet another long weekend in May.

(Target) They declared (that) a holiday was inappropriate because there/were important exams coming up soon.

(DO) Congress was looking for a way to honor the slain civil rights leader.

(Target) They declared (that) a holiday was inappropriate because there/were better ways to honor him.

15. reveal

(SC) Luke's wife finally asked him why he wasn't concerned about the kids stealing the package he had left on the seat of his car.

(Target) He revealed (that) the box had actually been empty/and her pearls were safe in the upstairs closet.

(DO) Bob finally agreed to show Cindy the package that he had hidden under the bed.

(Target) He revealed (that) the box had actually been empty/all along and her present was actually in the upstairs closet.

16. claim

(SC) After his promotion, John sent a letter to the selection committee thanking them for choosing him.

(Target) He claimed (that) the honor made him very happy/and was the most exciting thing that had ever happened to him.

(DO) After he won the competition, John went down to the awards center.

(Target) He claimed (that) the honor made him very happy/and was the most exciting thing that had ever happened to him.

17. project

(SC) Because the historian expressed concern about the delays that were piling up, the studio executives asked her when she might be delivering the finished product to them.

(Target) She projected (that) the documentary would take about two/months longer than originally planned.

(DO) As she began her presentation in the viewing room, the producer asked her assistant to dim the lights.

(Target) She projected (that) the documentary would take about two/hours and then she would answer questions.

18. insert

(SC) The newspaper editors were arguing intensely and the reporter was having a hard time getting a word in edgewise.

(Target) Finally though, she inserted (that) the paper seemed to be falling/apart and radical change was needed.

(DO) While Bob was sweeping the attic, June was getting frustrated at how hard it was to put the musty old documents back into their boxes.

(Target) Finally though, she inserted (that) the paper seemed to be falling/apart and that she couldn't put it away without ripping it.

19. feel

(SC) Rick was snug inside the cabin, but his horses were outside for the night and that worried him.

(Target) He felt (that) the weather might become a problem/as the night wore on.

(DO) Rick was beginning to get a little cold as he climbed the icy mountain.

(Target) He felt (that) the weather might become a problem/as time wore on.

20. report

(SC) The newscaster had to take a deep breath before he could give details of the deaths at the high school.

(Target) He reported (that) the students were caught by surprise/when the gunman appeared out of nowhere.

(DO) Danny loved being hall monitor at his high school because it gave him such a sense of power.

(Target) He reported (that) the students were caught by surprise/when he walked into the bathroom and caught them smoking.

Appendix C

Results of norming studies by verb: Percentage use of each sense, sense-contingent use of SC, DO, and Other (Oth) structures, and *that*-preference.

VERB	SC context SC sense				DO context DO sense			No context SC sense			DO sense		
	Sense	SC	DO	%that	Sense	SC	DO	Sense	SC	DO	Sense	SC	DO
Acknowledge	85	68	28	93	80	3	88	59	59	36	41	0	87
Add	100	90	0	92	63	0	100	30	64	36	35	0	100
Admit	100	93	0	78	63	0	96	97	58	19	3	0	100
Anticipate	48	42	45	86	95	0	100	30	0	91	46	0	94
Bet	88	100	0	61	95	0	95	14	100	0	84	0	55
Claim	98	95	0	70	68	0	100	76	79	0	22	0	100
Confirm	73	81	16	85	33	8	85	35	54	31	54	0	100
Declare	93	89	3	76	75	6	54	27	70	10	65	0	88
Feel	95	36	0	71	83	0	59	5	50	0	95	0	9
Find	98	47	3	67	98	0	100	3	100	0	95	0	100
Grasp	90	69	31	94	88	0	71	27	0	90	70	0	92
Indicate	90	90	5	89	50	58	77	46	88	12	35	8	69
Insert	90	25	74	100	100	0	100	5	50	50	95	0	100
Observe	58	96	0	100	70	28	73	5	100	0	86	3	97

Appendix C. (*continued*)

VERB	SC context SC sense				DO context DO sense			No context SC sense			DO sense		
	Sense	SC	DO	%that	Sense	SC	DO	Sense	SC	DO	Sense	SC	DO
Project	93	54	35	70	93	0	97	49	67	22	51	0	100
Recall	98	73	5	69	20	0	100	100	0	76	0	0	0
Recognize	100	63	35	76	95	0	97	22	75	25	78	0	97
Reflect	100	40	5	88	63	0	100	78	7	10	14	0	60
Report	90	72	19	92	95	0	97	57	43	52	35	15	69
Reveal	100	100	0	80	93	5	95	54	20	80	38	0	100
M	89	71	15	82	76	3	89	41	54	32	52	1	85

References

Altmann, G. T. M. (1998). Ambiguity in sentence processing. *Trends in Cognitive Sciences, 2*, 146–152.

Altmann, G. T. M. (1999). Thematic role assignment in context. *Journal of Memory and Language, 41*, 124–145.

Argaman, V., & Pearlmutter, N. J. (2002). Lexical semantics as a basis for argument structure frequency biases. In P. Merlo & S. Stevenson (Eds.), *The lexical basis of sentence processing: Formal, computational and experimental issues* (pp. 303–324). Amsterdam: John Benjamins.

Bever, T. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language*. New York: Wiley.

Boland, J. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes, 12*, 423–484.

Caramazza, A., & Grober, E. (1976). Polysemy and the structure of the subjective lexicon. Semantics: Theory and application. In C. Rameh (Ed.), *Georgetown University roundtable on languages and linguistics* (pp. 181–206). Washington, DC: Georgetown University Press.

Clark, H., & Clark, E. (1977). *Psychology and language*. New York: Harcourt Brace Jovanovich.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257–271.

Connine, C., Ferreira, F., Jones, C., Clifton, C., Jr., & Frazier, L. (1984). Verb Frame preferences: Descriptive norms. *Journal of Psycholinguistic Research, 13*, 307–319.

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

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

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 202–224.

Ferreira, F., & McClure, K. (1997). Parsing of garden-path sentences with reciprocal verbs. *Language and Cognitive Processes, 12*, 273–306.

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., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*, 516–547.

Fodor, J. A., & Garret, M. (1967). Some syntactic determinants of sentence complexity. *Perception & Psychophysics, 2*, 289–296.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry, 9*, 427–474.

Ford, M., Bresnan, J., & Kaplan, R. M. (1982). *A competence-based theory of syntactic closure*. Cambridge, MA: MIT Press.

Frazier, L. (1987a). *Sentence processing: A tutorial review*. Hillsdale, NJ: Erlbaum.

Frazier, L. (1987b). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory, 5*, 519–560.

Garnsey, S. M., Pearlmutter, N. J., 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.

Gibson, E., & Schütze, C. (1999). Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. *Journal of Memory and Language, 40*, 263–279.

Gibson, E., Schütze, C. T., & Salomon, A. (1996). The relationship between the frequency and complexity of linguistic structures. *Journal of Psycholinguistic Research, 25*, 59–92.

Godfrey, J., Holliman, E., & McDaniel, J. (1992). SWTICHBOARD: Telephone speech corpus for research and development. Paper presented at the Proceedings of the ICASSP-92, San Francisco.

Grimshaw, J. (1979). Complement structure and the lexicon. *Linguistic Inquiry, 10*, 279–326.

Johnson, C. R., & Fillmore, C. J. (2000). The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure. In *Proceedings of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics (ANLP-NAACL 2000), April 29–May 4, 2000, Seattle WA* (pp. 56–62).

Juliano, C., & Tanenhaus, M. K. (1994). A constraint-based lexicalist account of the subject/object attachment preference. *Journal of Psycholinguistic Research, 23*, 459–471.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science, 10*, 137–194.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32*, 474–516.

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

Klein, D. E., & Murphy, G. (2001). The representation of polysemous words. *Journal of Memory and Language, 45*, 259–282.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

Levin, B. (1993). *English verb classes and alternations, a preliminary investigation*. Chicago: University of Chicago Press.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language, 32*, 692–715.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

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

Manning, C., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Marcus, M., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics, 19*, 313–330.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes, 12*, 137–176.

McRae, K., Hare, M., Ferretti, T. R., & Elman, J. L. (2001). Activating events from typical agents, patients, instruments, and locations via event schemas. In *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society*. Princeton, NJ: Erlbaum.

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.

Merlo, P. (1994). A corpus-based analysis of verb continuation frequencies for syntactic processing. *Journal of Psycholinguistic Research, 23*, 435–457.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography, 3*, 235–244.

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.

Narayanan, S., & Jurafsky, D. (1998). *Bayesian models of human sentence processing, proceedings of the 20th annual conference of the Cognitive Science Society* (pp. 752–757). Hillsdale, NJ: Erlbaum.

Pesetsky, D. (1995). *Zero syntax*. Cambridge, MA: MIT Press.

Pickering, M., & Traxler, M. J. (1998). Plausibility and recovery from garden-paths: An eyetracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 940–961.

Pinker, S. (1989). *Learnability and cognition*. Cambridge, MA: MIT Press.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 785–794.

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition, 61*, 127–159.

Rice, S. A. (1992). Polysemy and lexical representation: The case of three English prepositions. In *Proceedings of the fourteenth annual conference of the cognitive Science Society* (pp. 89–94). Hillsdale, NJ: Erlbaum.

Rodd, J., Gaskell, M. G., & Marslen-Wilson, W. D. (1999). Semantic competition and ambiguity disadvantage. In *Proceedings of the twenty-first annual conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Roland, D. (2002). *Verb sense and verb subcategorization probabilities*. Ph.D. Thesis, University of Colorado, Boulder.

Roland, D., & Jurafsky, D. (1998). How verb subcategorization frequencies are affected by corpus choice. In *Proceedings of COLING/ACL* (pp. 1122–1128).

Roland, D., & Jurafsky, D. (2002). Verb sense and verb subcategorization probabilities. In P. Merlo & S. Stevenson (Eds.), *The lexical basis of sentence processing: Formal, computational and experimental issues* (pp. 303–324). Amsterdam: John Benjamins.

Roland, D., Jurafsky, D., Menn, L., Gahl, S., Elder, E., & Riddoch, C. (2000). Verb subcategorization frequency differences between business-news and balanced corpora: The role of verb sense. In: *Proceedings of the workshop on comparing corpora, Hong Kong* (pp. 28–34).

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27*, 499–520.

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.

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*, 211–272.

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.

Vu, H., Kellas, G., Metcalf, K., & Herman, R. (2000). The influence of global discourse on lexical ambiguity resolution. *Memory and Cognition, 28*(2), 236–252.

Williams, J. N. (1992). Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research, 21*, 193–218.