Studying the Grammatical Aspects of Word Recognition: Lexical Priming, Parsing, and Syntactic Ambiguity Resolution

Jared M. Novick,[1,3] Albert Kim,[2] and John C. Trueswell[1]

Two experiments are reported examining the relationship between lexical and syntactic processing during language comprehension, combining techniques common to the on-line study of syntactic ambiguity resolution with priming techniques common to the study of lexical processing. By manipulating grammatical properties of lexical primes, we explore how lexically based knowledge is activated and guides combinatory sentence processing. Particularly, we find that nouns (like verbs, see Trueswell & Kim, 1998) can activate detailed lexically specific syntactic information and that these representations guide the resolution of relevant syntactic ambiguities pertaining to verb argument structure. These findings suggest that certain principles of knowledge representation common to theories of lexical knowledge—such as overlapping and distributed representations—also characterize grammatical knowledge. Additionally, observations from an auditory comprehension study suggest similar conclusions about the lexical nature of parsing in spoken language comprehension. They also suggest that thematic role and syntactic preferences are activated during word recognition and that both influence combinatory processing.

KEY WORDS: On-line language comprehension; grammatical knowledge; lexical priming; word recognition; syntactic processing; parsing.

INTRODUCTION

As part of the process of comprehending speech or text, the language system must achieve two important goals: (1) the recognition of individual

words and the meaning that they convey, and (2) the linking of these words together to form phrases, which convey further combinatory meaning. Most psycholinguistic theories have assumed that these processes are instantiated within separate subsystems that involve diametrically opposed sorts of computations. When comprehending sentences, we *recognize* words but *build* phrase structure.

Recent developments in psycholinguistic theorizing have led many to question this supposed dissimilarity between lexical and phrasal processing. Many have begun to ask: What if we also recognize the presence of phrases? And what if the detection of these phrases triggers the conveyance of combinatory meaning? Such assumptions might provide continuity within theories of language comprehension because the system would become a probabilistic pattern recognition device *through and through,* detecting linguistic events of various sorts at multiple levels (phonemes, words, phrases). These assumptions might also change what is thought to be involved in lexical and phrasal processing. As we discuss below, these assumptions imply a notion of lexical processing that bears considerably more responsibility for the combinatory analysis of language.

Psycholinguistics has not been alone in this focus on the lexical aspects of combinatory processing. Syntactic theory has increasingly moved detailed combinatory information into the lexicon, where individual lexical items are associated directly with their syntactic and semantic combinatory options (e.g., Chomsky's 1993 Minimalist Program, and grammatical formalisms such as Lexicalized Tree Adjoining Grammar (LTAG; Joshi & Schabes, 1996) and Combinatory Categorial Grammar (CCG; Steedman, 2000)). The field of applied parsing in computational linguistics has also seen a shift toward lexicalization (Srinivas & Joshi, 1999). Many have recognized the effectiveness of coding these syntactic options as tendencies (i.e., the probability of each option given a word and its local context). In doing so, statistical natural language processing systems have begun to be able to recover the grammatical structure of novel sentences with astonishing accuracy (Collins, 2000; Marcus *et al.,* 1993; see also Jurafsky & Martin, 2000).

These movements in linguistics and computational linguistics touch on many of the same issues that have given rise to the development of constraint-based lexicalist theories of parsing in psycholinguistics (e.g., Kim *et al.,* 2002; MacDonald *et al.,* 1994; Trueswell & Tanenhaus, 1994). In ways similar to the statistical NLP systems, these theories propose that the recognition of a word involves the relative activation of detailed grammatical options, which are used to guide further combinatory processes. As a result, the frequency-based activation of lexical alternatives becomes the basis for the resolution of many syntactic ambiguities.

experiments differ from most studies of syntactic ambiguity resolution because we have opted to intervene in the ongoing comprehension process, typically covertly, via the brief / attenuated presentation of a lexical prime. Because our primes are rarely identified by readers and listeners, we are free to manipulate the properties of our primes quite severely, often using words that would be completely inappropriate in the target sentence if they were perceptually identified. In spite of this, our priming interventions influence comprehension processes in ways anticipated by lexicalist parsing theories. As we attempt to illustrate here, these priming patterns tell us a great deal about the role of lexical processing in the combinatory analysis of language.

Background: Lexical Priming and Syntactic Ambiguity Resolution

The use of lexical priming techniques in psycholinguistics has arguably been a highly effective tool for studying both the content of lexical representations and the time-course with which such information is activated (e.g., Foss, 1988). For example, priming has been used to map the time-course of activation of phonological and orthographic information during word recognition (e.g., Forster & Taft, 1994; Kouider & Dupoux, 2001) as well as the activation of the alternative meanings of ambiguous words (e.g., Simpson & Burgess, 1985). In most cases, however, the tasks used in these studies to measure participants' reaction to target words (i.e., naming and lexical decision) are unlikely to be influenced by any hypothesized coactivation of sub-

e1

for the parallel activation of orthographic, phonological, and semantic information associated with a letter string.

A clear prediction of lexicalist parsing theories is that word recognition also includes the activation of detailed combinatory information, in the form of possible complements that a word may take. It is this activation process that ought to influence the relative availability of alternative syntactic analyses. If this is the case, the syntactic preferences of a briefly presented prime word ought to have a direct impact on a reader's parsing preferences of a syntactically ambiguous phrase. Trueswell and Kim (1998) tested these predictions in a series of experiments using a self-paced reading version of fast priming. In the study, they examined how the fast priming of verbs can influence the way in which readers process sentences containing the Direct Object/Sentence Complement ambiguity in (1) above. Like Garnsey *et al.* (1997), Trueswell and Kim (1998) compared the reading times of temporarily ambiguous sentences (e.g., Example 1) to unambiguous versions that included a "that" (e.g., . . . *accepted that the* . . .). Increased reading times at the disambiguating phrase (e.g., . . . *would go* . . .) were taken as signs of misanalysis of the ambiguous NP (e.g., *the fire*) as the direct object of the verb.

Prior to reading each sentence, the words of the sentence were masked with each character in a word covered by an equal sign ('='). Each press of the button uncovered a word and replaced the previous word with equal signs. On critical trials however, when the participant reached the matrix verb (e.g., *accepted*), a prime word was displayed in its place for 39 ms. The prime verb was then replaced by the target verb, which remained on the screen until the next press of the button. This event was typically perceived as a flicker on the screen, with participants reporting in a detailed post-experiment questionnaire that they rarely identified any prime words. Two different types of prime words were compared, which had been selected based on their argument-taking properties as measured from a separate sentence completion study. DO-Primes (e.g., "obtained") were verbs that strongly prefer a direct object and do not permit a sentence complement. SC-Primes (e.g., "realized") were verbs that strongly prefer a sentence complement and rarely use a direct object. If the recognition of a verb includes the activation of its possible argument structures, we might expect that the argument preferences of the "flicker" (the prime) would influence the size of the garden-path effect; DO-Primes should induce a large garden-path effect, whereas SC-Primes should reduce the garden-path effect.

Indeed, the experiments showed the expected pattern of priming. In particular, Trueswell and Kim (1998) observed that the processing difficulty usually found in the disambiguating region of ambiguous (no-that) sentences was significantly decreased when the matrix verb had been primed with a SC-Prime as compared to when it had been primed with a DO-Prime. That is, dif-

ficulty that stems from readers erroneously committing to a direct object analysis, due to their detecting the implausibility of the noun as the direct object and having to revise this syntactic commitment, was far less likely to occur when the prime provided argument structure information that could help the reader avoid this misinterpretation. Crucially, the priming had its influence only on the ambiguous conditions and not the unambiguous conditions, implicating the primes' influence on avoiding the garden-path and not the general fit of the prime into the sentence. One other observation from Trueswell and Kim (1998) is also worth mentioning, because we will be comparing it to the findings of Experiment 1 reported here. In both of their experiments, they observed a 30 ms effect of priming at the verb, with SC-Primes being faster than DO-Primes. This effect may suggest a partial role for semantic priming in this process, a point we return to later in this paper.

In sum, the Trueswell and Kim (1998) findings show that garden-path effects can be considerably mitigated by a briefly displayed prime verb, even though the primes were rarely identifiable to the readers. Prime verbs that prefer to take a sentence complement reliably reduced the garden-path typically associated with ambiguous sentence complement constructions. Moreover, the timing of these primes, as well as their subjective perception by participants, strongly suggests that verb combinatory information of this sort is automatically activated during word recognition.

EXPERIMENT 1: FAST PRIMING OF NOUN ARGUMENT STRUCTURE

The Trueswell and Kim (1998) results contribute to a debate on the lexicon's role in comprehension that has focused heavily on the combinatory properties of verbs. This focus on verbs makes sense, given the wide range of research assuming an anchorlike status of verbs in the syntactic and thematic organization of language. Although verbs may be the obvious starting point for the study of lexically based combinatory knowledge, lexicalist proposals clearly hypothesize that other types of words share the burden of combinatory knowledge representation (e.g., Kim *et al.,* 2002).

Perhaps the simplest reason for an account that extends beyond verbs alone is that grammatical phenomena suggest it. For instance, nouns like *opinion* can, like the verb *accepted* in (1), combine with a finite clause (e.g., *we should go*), as in (2):

(2) (She expressed) the **opinion** that we should go.

If the knowledge that drives combinatory processing is lexically generated, then nouns like *opinion* ought to encode knowledge about the complements

they can occur with. This same construction illustrates another reason to extend the inquiry beyond verbs: The syntactic similarity between nouns like *opinion* and verbs like *accepted* (in (1)) suggests an underlying connection between their lexical representations. Specifically, the representations of words like *opinion* and *accepted* may contain shared components, even though the words belong to different basic grammatical categories. That is, knowledge of sentential complements is encoded by overlapping distributed lexical representations.

As reported here, we conducted an experiment to investigate the hypothesis that the comprehension of sentential complement constructions is driven by lexical knowledge representations that are shared by both nouns and verbs. We used an experimental approach that was similar to that used by Trueswell and Kim (1998), except that the prime words were nouns rather than verbs.

Methods

Participants

Thirty-six adults from the University of Pennsylvania community volunteered for the experiment. Participants received course credit or were paid for their participation. All were native speakers of English.

Materials and Procedure

Participants read sentences like (3), which were structurally identical to the sentences of Trueswell and Kim (1998).

> (3) The ice skater doubted (that) the judges would keep her from competing.
>   a. opinion (SC-bias prime)
>   b. freedom (Abstract prime)
>   c. machine (Concrete prime)

In each sentence, the target verb (e.g., *doubted*), could occur with either a direct object or a sentential complement. Unlike Trueswell and Kim (1998), however, the potentially ambiguous noun (e.g., *judges*) was always a good direct object of the target verb. This was done to eliminate a potential confound that existed in the previous experiments, in which increased reading times in the ambiguous conditions could be attributed either to the semantic anomaly of the noun (and its spillover) or to effects of garden-pathing.

Syntactic ambiguity was again manipulated by including or excluding the complementizer *that*. Sentences were read in a self-paced moving window

presentation. When readers encountered the position of the target verb, a prime noun was first presented for 39 ms and was then replaced by the target verb. Three levels of Prime-Type were crossed factorially with the Ambiguity factor: (a) "SC-bias" primes, nouns that frequently take a sentential complement (e.g., *opinion*); (b) "Abstract" primes, nouns that are

Results

Mean reading times are presented in Table II. Reading times for Abstract and Concrete primes are collapsed, because whenever effects of Prime Type occurred, planned comparisons showed that these two levels of Prime Type were not distinguishable, as discussed later. Figure 1 plots Ambiguity Effects (as a function of word position), i.e., the difference in reading times between the ambiguous and the unambiguous conditions.

In the disambiguating region (*would keep*), a main effect of Ambiguity was evident ($F1(1,35) = 13.83, p < .001; F2(1,23) = 14.66, p < 0.001$), indicating processing difficulty connected with syntactic misanalysis. Ambiguity effects were also seen in the determiner that followed the main verb ($F1 = 26.16, p < .001; F2 = 47.45, p < .001$). A similar effect of ambiguity was observed in Trueswell and Kim (1998) and may in part be the result of slight disruptions in processing due to the priming event.[4] Main effects of Prime Type were also seen, both in the disambiguating region ($F1(2,70) = 6.63, p < .005; F2(2,46) = 4.42, p < .05$) and also earlier, at the ambiguous verb (e.g., *doubted*) ($F1(2,70) = 5.11, p < .01; F2(2,46), p < .05$).

Planned comparisons revealed that the effect of Prime Type was driven by the difference between the SC-bias condition and the other two Prime Types. In the disambiguating region, reading times for the SC-bias condition were faster than both the Abstract ($F1(1,35) = 8.42, p < .01; F2(1,23) < 4.06, p = .056$) and the Concrete conditions ($F1(1,35) = 12.88, p < .01; F2(1,23) = 10.14, p < .005$), while Abstract and Concrete were indistinguishable ($Fs < 1$). The same pattern occurred at the target verb, where SC-

Discussion and Summary

We found that fast priming effects generalize beyond verbs to another part of the lexicon, nouns. The short-lived processing of SC-bias noun primes influenced the syntactic processing of the host sentences in a manner that supports the sentential-complement interpretation. We claim that (even partial) processing of SC-bias nouns activates predictive knowledge about sentential complements. The representations that encode this knowledge are shared with sentential-complement verbs, and their preactivation influences the recognition of the target verb. By supporting the sentential-complement sense of the ambiguous target verb, the priming pushes the processing system toward the correct grammatical analysis of the sentence.

The similarity of Abstract and Concrete prime conditions constrains our interpretations. We mentioned the possibility that priming effects might arise from semantic abstractness in the primes. However, Abstract primes behaved differently from SC-bias primes and indistinguishably from Concrete primes. Thus, specifically combinatory knowledge is implicated over simple abstractness. It is intriguing to note that many of the Abstract primes were even capable of taking complements of some sorts (e.g., *the freedom to leave*). Thus, it appears that something quite specific is at work here.

These results demonstrate that nouns can project syntactic structure in a way that plays an active role in the guidance of sentence interpretation (see also Schütze & Gibson, 1999). This contribution of noun argument structure is expected under lexicalist proposals, given the need to specify combinatory information specific to these items. Furthermore, the facts from other languages support the need for a system that hypothesizes detailed structural information based on words other than verbs. For instance, if pro-ustructuy6198 s, (e.g.Kamidn.,

Methods

Participants

Sixteen participants from the University of Pennsylvania volunteered for the experiment. They received course credit or were paid for their participation. All participants were native speakers of English and had normal or corrected-to-normal vision.

Materials and Procedure

Like our fast-priming reading experiment, this experiment crucially combined a priming technique with an on-line measure of ongoing sentence processing commitments, so as to uncover how the combinatory preferences of these primes might influence the concurrent syntactic processing of the sentence. In particular, participants' eye movements were recorded as they heard spoken instructions, such as Example (5).

(5) Now I'd like you to turn the bear with the stick.

Target verbs (e.g., *turn*) were selected based on an earlier sentence completion study and dubbed Equi-bias verbs; that is, participants had been equally likely to complete a fragment like *Turn the doll with . . .* with an NP modifier or a VP instrument. Referential scenes in the on-line study contained a potential instrument object (e.g., a full-scale stick), a target animal (e.g., a toy bear holding a miniature replica of the instrument object), a competitor animal (e.g., a toy pig holding a knife), and a distractor item (e.g., a diskette).

One of two types of primes was digitally mixed and aligned with the target onset: Modifier-bias verbs (e.g., *hug*), which, according to previous norming studies, strongly prefer *with the X* as an NP modifier; or Instrument-bias verbs (e.g., *clean*), which strongly prefer to take *with the X* as a VP instrument. Postexperimental questionnaires revealed that listeners were largely unaware of the presence of the primes, and when aware, they were unable to identify what was being said. The consensus among participants was that some nonintrusive background noise was heard, which never disrupted their understanding of the target sentence.

Predictions

It was anticipated from earlier experiments using nonpriming versions of these same stimuli (Snedeker *et al.,* 2001) that listeners' actions and eye movements would reveal the kind of interpretation they assigned the ambiguous phrase. In particular, subjects who took *with the stick* as an instrument ought to look to the potential instrument upon hearing *stick* and use this object to perform the task; participants who took a modifier interpretation ought not to look

served as categories. For instance, a verb was highly regarded to take an instrument role if there were a high number of prepositional phrase completions containing instruments, for example, *Clean the baby doll with a sponge.* Sentences that contained mere VP-attachments like *Clean the baby doll with care* or NP-attachments like *Clean the baby doll with green eyes* were categorized as Noninstruments. This categorization scheme resulted in a large semantic difference between how likely certain verbs took instruments (e.g., *clean,* which approached 100% Instrument completions) or how likely they were not to take instruments (e.g., *hug,* which never took an instrument role), and this is how the prime types were chosen. Contrastingly, inspection of gross *syntactic* differences in continuations for these verbs based on whether they were likely to take VP-attachments or not resulted in very little difference in syntactic preferences. In particular, both Instrument-bias primes and Modifier-bias primes had high VP-attachment preferences (98% and 76%, respectively). Although it is possible that this small difference in VP-attachment rates could be influencing the on-line priming effects, it seems more likely that there is a strong availability of semantic roles rather than syntactic structure during auditory verb recognition. The most plausible conclusion, and the one consistent with most views of argument structure, is that recognition of a verb includes activation of permissible syntactic complements, permissible thematic roles, and the possible mappings between the two (e.g., see Carlson & Tanenhaus, 1988).

CONCLUSIONS

Several conclusions about the nature of sentence comprehension arise from these results:

Steedman, M. (2000). *The Syntactic Process.* Cambridge, MA: MIT Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12,* 242–248.

Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language, 39,* 102–123.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73,* 89–134.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. J. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on Sentence Processing* (pp. 155–179). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*(3), 285–318.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *s70rtfromuHoA20,ew6denc,.424eedn*