Understanding the Constraints on Syntactic Generation: Lexical Bias and Discourse Congruency Effects on Eye Movements

Julie E. Boland

Institute of Cognitive Science, University of Louisiana at Lafayette

and

Allison Blodgett

Ohio State University

We measured eye movements as people read short stories. The target sentences contained noun/verb homographs (e.g., *duck*) and were preceded by a biasing context sentence. The homograph in the target sentence was always disambiguated by a case-marked pronoun, e.g., *She saw his/him duck*. Lexical bias effects (reflecting the relative frequency of the noun and verb forms) were found in the initial fixations on the homograph. In contrast, discourse congruency effects were first observed several words downstream in the probability of a regressive eye movement. Strong discourse congruency effects were also observed in the second pass reading times. We concluded that the lexical bias effects reflect processing difficulty during the initial generation of syntactic structure, while the discourse congruency effects reflect later anomaly detection. Thus, the data challenge syntactic processing models in which all relevant and available constraints are brought to bear uniformly and simultaneously. © 2001 Academic Press

Key Words: sentence comprehension; parsing; lexical bias; frequency effects; syntactic category ambiguities; discourse context.

A central issue in sentence comprehension research is the means by which we structure incoming words into a hierarchical representation according to the grammar of our language. This process, known as syntactic analysis or parsing, can be broken down into at least three components: (1) *generation* of syntactic structure, including the identification of alternative structures at points of ambiguity; (2) *selection* of a single structure; and (3) *reanalysis* if the structure initially selected turns out to be incorrect.

Of course, in a given parsing theory, these components are not necessarily implemented in three distinct stages. For example, because only one structure is initially generated in the traditional garden path model (e.g., Frazier, 1978, 1987), the theory makes no distinction between syntactic generation and syntactic selection during the initial parse. In contrast, one can distinguish between syntactic generation and syntactic selection in any parsing model that activates or generates syntactic alternatives in parallel.

The current paper investigates how lexical frequency and discourse context influence syntactic generation. Our central hypothesis is that these two levels of constraint influence syntactic generation in quite different ways: lexical frequency influences the ease with which a structure is accessed or generated, whereas discourse constraints can only be used to select among grammatical alternatives once they have been generated. Secondarily, we test several hypotheses concerning the maintenance of multiple syntactic alternatives throughout an ambiguous region.

Our central hypothesis derives from the lexically driven processing theory spelled out in Boland (1997a). Syntactic forms are accessed from the lexicon and these syntactic forms constitute essential building blocks of syntactic structure. Much of the time, a word will have several alternative syntactic forms that compete for selection. The accessibility of the competing forms is influenced by their relative frequency in the language. The syntactic forms from the lexicon, together with global syntactic knowledge, determine the set of syntactic alternatives at each word in a sentence. Thus, we envision syntactic generation as involving both the access of lexical structures and the integration of these lexical structures into larger syntactic objects, following the global constraints of the grammar. Discourse constraints, unlike lexical and syntactic constraints, cannot influence the set of syntactic alternatives that is generated, though discourse constraints can guide selection of the most likely syntactic structure. Indeed, a theory that allowed discourse to restrict syntactic generation would make the undesirable prediction that we cannot construct a syntactic analysis for a sentence in an incongruent discourse context. As will become clear, our account of sentence comprehension is a type of constraint-based lexicalist theory, but it differs from other constraint-based accounts in drawing a distinction between syntactic generation and syntactic selection.

All constraint-based lexicalist models initially activate multiple syntactic alternatives at points of ambiguity (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994).[1] The term "constraint" refers to any pattern from our language experience that becomes part of our linguistic knowledge; constraints can be based on frequency, plausibility, grammaticality, and so forth. Most attempts to specify constraint-based models have maintained that all constraints are brought to bear simultaneously during syntactic analysis. However, different models implement constraint usage in different ways. Most often, the constraints guide syntactic selection once the syntactic alternatives have been identified (e.g., Spivey & Tanenhaus, 1998), but in at least one model, the constraints influence which syntactic structures are initially applicable (Tabor, Juliano, & Tanenhaus, 1997). To illustrate, we briefly describe the architecture of two example models.

Spivey and Tanenhaus (1998) used a connectionist framework to model the analysis of a reduced relative clause ambiguity (e.g., *The horse raced* . . .). In their computational model, processing difficulty at the verb was realized as competition between a main clause structure and a relative clause structure, which were entered into the simulation as unweighted syntactic alternatives. Because the syntactic alternatives are equally (un)activated at the start of the simulation, no constraints are allowed to influence syntactic generation. Rather, all classes of constraint are used simultaneously during syntactic selection. The structures compete for activation from the relevant constraints (discourse context, lexical frequency, etc.), and the winning syntactic alternative is selected for the developing syntactic representation.

The Tabor et al. model (1997; see also Tabor and Tanenhaus, 1999) incorporates a gravitational dynamical systems component, with syntactic alternatives represented as attractors in metric space. This model allows multiple constraints to determine the syntactic alternatives in the initial candidate set, because at each word position, the model uses the available constraint information to situate the system in a new starting position in the space. Although there is no actual generation of structure at this point, the starting position determines which attractors (i.e., syntactic alternatives) will affect processing by exerting a gravitational pull. (The amount of pull exerted by each attractor is determined by its frequency.) Intermediate states between two or more attractors map loosely to temporary syntactic ambiguity, but syntactically unambiguous inputs can also result in intermediate states if the grammatical information is inconsistent with other constraints. Syntactic se-

---

[1] This is not literally true in a model like that of Tabor, Juliano, and Tanenhaus (1997), but one could consider the involvement of two or more attractors as equivalent to the activation of syntactic alternatives.

lection occurs after each successive word, as the system gravitates to a single attractor/syntactic structure.

These two models illustrate some common properties of constraint-based lexicalist approaches. Syntactic alternatives are evaluated in parallel, using constraints from any level of representation. Thus one can, in principle, distinguish between the generation of the syntactic alternatives and the selection of the optimal structure whenever there is a local syntactic ambiguity. For example, the fragment *They saw her. . .*, is consistent with either the accusative (acc) or the possessive (poss) form of *her*, and each form can be attached to the matrix verb phrase (VP) as part of either an embedded sentence (S) or a direct object noun phrase (NP), as shown in (1). At *her* at least some of these structural alternatives would serve as competitors in a model like Spivey and Tanenhaus's (1998) or as the attractors in a model like Tabor et al.'s (1997). In either case, the alternative that is most consistent with the available constraints would be selected.

(1)   a. *They saw (her_acc signing)*s.

 b. *They saw ((her_poss father) signing)*_S.

 c. *They saw (her_acc)*_NP.

 d. *They saw (her_poss signpost)*_NP.

(2)   *They saw him sign.*

Now consider the situation in (2) at the word *sign*. There is no syntactic ambiguity, but *sign* is lexically ambiguous because it is a noun/verb homograph. According to multiple access theories of word recognition, both the noun and the verb forms of *sign* would be accessed (e.g., Duffy, Morris, & Rayner, 1988; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979). However, a model of syntactic ambiguity resolution like that of Spivey and Tanenhaus (1998) makes no prediction about processing difficulty in this case. Only one syntactic alternative would be available for evaluation, because only the verb form of *sign* can be attached to the syntactic structure associated with *They saw him. . . .* The situation is somewhat different in the Tabor et al. (1997) model, because the gravitational system can be influenced by ungrammatical attractors if they are strongly supported by some constraints. Therefore, Tabor et al. predict that both lexical and discourse constraints could potentially influence syntactic generation. These examples demonstrate that unambiguous structures provide an important test case for determining how various constraints influence syntactic generation.

Our hypothesis, that lexical frequency plays a special role in syntactic generation, can be illustrated by contrasting *sign* with *play*. *Sign* occurs most frequently as a noun, while *play* is most frequently a verb. The extensive literature on frequency effects in word recognition demonstrates that the dominant (more frequent) form of an ambiguous word is more strongly available during lexical access than weaker forms (Binder & Morris, 1995; Binder & Rayner, 1998; Carpenter & Daneman, 1981; Dopkins, Morris, & Rayner, 1992; Duffy et al., 1988; Hogaboam & Perfetti, 1975; Pacht & Rayner, 1993; Rayner, Pacht & Duffy, 1994; Tabossi, Colombo, & Job, 1987), although some accounts maintain that these frequency effects can be negated if the context strongly biases a particular form (Kellas, Paul, Martin, & Simpson, 1991; Paul, Kellas, Martin, & Clark, 1992; Martin, Vu, Kellas, & Metcalf, 1999; Simpson & Krueger, 1991; Vu, Kellas, Metcalf, & Herman, 2000; responses to some of these claims can be found in Rayner, Binder, & Duffy, 1999, and Binder & Rayner, 1999). Given the neutral context in (3), all constraint-based approaches would predict that *her sign* would likely be analyzed as a possessive NP, while *her play* would likely be analyzed as an embedded subject and verb. All constraint-based approaches would also agree that selection of the syntactic structure could be influenced by biasing discourse context and other relevant constraints. In support of these predictions, Boland (1997b) found that syntactic selection for the ambiguity exemplified in (3) was influenced by both lexical and discourse constraints.

(3)   a. *They saw her sign.*

 b. *They saw her play.*

As illustrated above, using Spivey and Tanenhaus (1998), most constraint-based models make clearer predictions about constraint influence during syntactic selection (ambiguity resolution) than about how constraints influence syntactic generation in unambiguous structures like (2). We hypothesize that the relative strength of the noun and verb forms of the homograph should still influence processing difficulty in unambiguous sentences, because the less frequent form is less strongly available during word recognition. In contrast, supportive or conflicting discourse context should not influence the ease of syntactic generation in unambiguous structures. This two-part hypothesis is consistent with exhaustive access models of word recognition (e.g., Duffy et al., 1988), but it would not follow straightforwardly from selective access models that allow strongly biasing context (such as unambiguous syntactic information) to overwhelm bottom-up frequency effects (e.g., Paul et al., 1992).

To examine the roles of lexical frequency and discourse context, Boland (1997b) embedded unambiguous and ambiguous sentences in simple stories. The target sentences were preceded by a biasing context sentence; examples for the item constructed around the noun/verb homograph *duck* are given in (4). The matching target sentences are in Table 1, along with a summary of the results. Additional examples are provided in Appendix B. For each set of target sentences, context type and target type were crossed to create four conditions, noun context with noun target, noun context with verb target, verb context with noun target, and verb context with verb target.[2] In Table 1, the disambiguating words in the target sentences are in boldface type, to highlight the differences in target sentence type.

(4)   a. Noun context: *As they walked around, Kate looked at all of Jimmy's pets.*

b. Verb context: *As they walked around, Kate watched everything that Jimmy did.*

As shown in Table 1, the lexical bias of the noun/verb homograph influenced processing difficulty in both unambiguous and (long) ambiguous target sentences, consistent with exhaustive access models of word recognition. In targets with long ambiguous regions, the lexical bias effect was found at the first word after the point of disambiguation. In unambiguous target sentences, the effect was found one to two words after the homograph, depending on whether the discourse was congruent or incongruent. Effects of lexical bias have since been replicated in isolated, unambiguous sentences containing noun/verb homographs, though not using eye tracking (Corley, 1998; Boland et al., 2000, in-published manuscript). Boland (1997b) found discourse congruency effects only in the target sentences with an ambiguous region. When the ambiguous region was short, an effect of discourse congruency occurred on the first word after the pronoun, demonstrating that discourse congruency information guided syntactic selection at the pronoun, as predicted by constraint-based models of sentence comprehension. This finding provided clear evidence that discourse constraints were computed rapidly. Nonetheless, discourse congruency did not influence processing (syntactic generation) in the unambiguous target sentences at any point.

The data pattern for the unambiguous sentences like those in the top of Table 1 is predicted by the hypothesis that both lexical and discourse constraints guide syntactic selection, but only lexical frequency influences syntactic generation. Boland (1997b) found that only the noun/verb bias of the homograph—and not discourse congruency—influenced processing difficulty in unambiguous sentences. Thus, in contrast to Spivey and Tanenhaus (1998) and Tabor et al. (1997), Boland concluded that all constraints are not used simultaneously. If Boland is correct, she has falsified a central claim of constraint-based models. However, there are several concerns about the results that might qualify her conclusions.

---

[2] The point at which the discourse incongruity of mismatched contexts was first apparent differed slightly across items. The incongruity was usually apparent at the pronoun, although congruent continuations were sometimes still possible at that point. For example, in the context of (4b), sentence (5a) below becomes awkward at *his*, although a congruent continuation like *She saw his arm twitch as he. . .* was possible until it was ruled out by the following word, *duck*.

TABLE 1

Word Position Where Effects Were First Observed in Boland (1997b)

Target sentence type		Observed effects by word position	
		Lexical bias of N/V homograph	Discourse congruency
Unambiguous			
Noun target:	*She saw **his** duck and chickens near the barn*	congruent context: *and*, incongruent context: *chickens/stumble*	no effect
Verb target:	*She saw **him** duck and stumble near the barn*		
Short ambiguous region			
Noun target:	*She saw her **ducks** and chickens near the barn.*	not applicable	*ducks/bend*
Verb target:	*She saw her **bend** and stumble near the barn.*		
Long ambiguous region			
Noun target:	*She saw her duck and **chickens** near the barn.*	*near*	*near*
Verb target:	*She saw her duck and **stumble** near the barn.*		

Note. The disambiguating word in each target sentence is highlighted in bold.

The first concern is that the lexical frequency effects were one to two words later than predicted. Effects often emerge a word or two downstream of the predicted location in the self-paced, word-by-word reading paradigm that Boland (1997b) used (e.g., see Taraban & McClelland, 1988), but the timing makes it difficult to argue decisively that the relevant effects impacted processing during the initial generation of syntactic structure. Likewise, the delay effectively prevents this finding from impacting the exhaustive/selective access debate in the word recognition literature. Both Boland et al. (2000, unpublished manuscript) and Corley (1998) found effects of lexical bias on the noun/verb homograph itself, but those studies used a different measure of lexical bias. Boland (1997b) used a contingent frequency measure based on sentence completion norms, whereas Boland et al. and Corley used corpus norms to determine the relative frequency of the noun and verb forms, without considering the local syntactic and semantic context. This distinction is theoretically important. Tabor et al. (1997) predict that contingent frequency should predict processing difficulty because their model incorporates a context-sensitive mechanism of lexical access and syntactic generation, as in Elman (1990). However, Boland (1997a) predicts

that overall bias should predict reading time in unambiguous contexts.

A second concern was the lack of consistent anomaly effects for mismatched discourse conditions in the unambiguous target sentences. Although Boland (1997b) did not predict context effects on syntactic generation, the mismatched target sentences were infelicitous in their discourse context. On any account, this infelicity should have produced an anomaly effect for mismatched conditions.

A third concern is that the semantic and syntactic disambiguation of the homograph was less immediate for the noun targets compared to the verb targets. After a context like (4b)—which is biased toward the verb reading of the homograph—one might initially interpret *duck* in the "unambiguous" noun target as a nominalized verb: *She saw his duck as the barn door nearly hit him in the head*. In our judgment, this is a possible complication in at least 5 of the 16 items (*duck, roll, slip, play, bow*). There is a syntactic complication as well. Suzanne Stevenson pointed out to us that the noun targets *can* actually be analyzed as verbs: e.g., *Jimmy's dog and Margie's dog were playing near the barn while Jimmy and Margie watched. The kids saw HIS duck under a gate while HERS chased a cat.*

This analysis is only felicitous in a spoken discourse with contrastive stress on *his*. It might well be generated under a constraint-based account in which all lexical alternatives are partially activated, but we would not expect the [*his* V. . .] analysis to be selected in the context of our experimental texts. No comparable ambiguities apply to the verb targets. If the homograph in the noun target sentences is not immediately disambiguated by the local context as we expected, lexical bias effects may be delayed for the noun targets relative to the verb targets. Noun disambiguation would be most delayed when the discourse context was incongruent, because the context would provide some support for the unintended form. If correct, this could explain why Boland's (1997b) lexical bias effects were a word later for incongruent conditions, when the target sentences were unambiguous.

The fourth issue is that Boland's (1997b) "unambiguous" targets illustrated in Table 1 were not completely devoid of local syntactic ambiguity. There is a temporary structural ambiguity that is independent of the syntactic class of the homograph. As shown in (5) and (6), the pronoun can either be (part of) the direct object or (part of) the subject of an embedded clause. More importantly, the noun and verb target sentences differ in the point at which these two possibilities were disambiguated. The verb targets were immediately disambiguated as the structure in (6b), e.g., *She saw him duck. . . .* In contrast, the noun targets were not fully disambiguated as the structure in (5a) until the end of the sentence—until the period was encountered, the sentence could have continued in a way consistent with (5b), i.e., *She saw* [*his duck and chickens near the barn* $_{NP}$] *eating*.

(5) Noun targets

    a. *She saw* [*his* NOUN$_{NP}$]

    b. *She saw* [*his* NOUN VP$_{CLAUSE}$]

(6) Verb targets

    a. *She saw* [*him*$_{NP}$]

    b. *She saw* [*him* VERB$_{CLAUSE}$]

Consider the hypothesis that direct object and clausal structures, corresponding to those in (5)

and (6), are maintained (in parallel) as long as each possibility is syntactically possible, with discourse congruency influencing the amount of activation for each.[3] If so, parallel structures would be sustained over a much longer period for the noun targets than for the verb targets. The increased processing load associated with maintaining parallel structures should lead to longer reading times or more re-reading in the noun targets compared to the verb targets. This account would explain a (nonsignificant) tendency toward slower reading times for noun targets near the end of the sentence in Boland (1997b). We expected to find stronger evidence of the increased processing load for noun targets in the current experiment.

It must be acknowledged, however, that serial accounts of syntactic ambiguity resolution predominate in the literature. The best known of these is the garden path model, which asserts that the simplest analysis is pursued first (e.g., Frazier, 1987), with simplicity being a function of the number of new nodes required in the phrase structure tree. Our target sentences used the structures in (5a) and (6b). No garden path is predicted for noun targets because the structure in (5a) is simpler than the structure in (5b). In contrast, the structure in (6b) is more complex than the structure in (6a). Thus, a heuristic that constructs the minimal structure at the pronoun predicts processing difficulty for verb targets *at the homograph*—the point at which the (6a) analysis is ruled out.

EXPERIMENT

The current experiment used Boland's (1997b, Experiment 2) materials in an eye movement paradigm, which provides several dependent measures reflecting varying amounts of process-

---

[3] Although an unlimited parallel account of syntactic ambiguity resolution is inconsistent with the abundant evidence for garden path effects, various forms of ranked/limited parallelism have been considered (e.g., Gibson, 1991; Just & Carpenter, 1992; Pearlmutter & Mendelsohn, 1999, unpublished manuscript; Stevenson, 1998). In fact, an account in which alternative syntactic representations are ranked according to the available constraint information, but maintained to some degree throughout the ambiguous region, received support from Boland (1997b). Experiments 3 and 4.

TABLE 2

Example Target Sentences, Delineated by the Regions Used in the Data Analysis

	Initial	Verb + pronoun	Homograph	Conjoined phrase	Final
Noun target	*She*	*saw his*	*duck and*	*chickens near*	*the barn.*
Verb target	*She*	*saw him*	*duck and*	*stumble near*	*the barn.*

ing, first fixation duration, summed first pass reading times, probability of a first pass regression, total reading times, etc. The various dependent measures provide a degree of granularity not available with tasks that rely on button pressing to provide the sole dependent measure. Thus, we hoped to better track the influence of lexical bias and discourse congruency, as well as to examine some predictions of serial and parallel accounts of ambiguity resolution.

The most important predictions can be summarized as follows. First, the hypothesis that lexical bias influences syntactic generation predicts lexical bias effects at the earliest possible point (i.e., the initial fixations on the homograph). Because the disambiguation of the homograph is less straightforward for the noun targets, the lexical bias effects might be delayed for the noun target condition when it is in an incongruent discourse context. Second, discourse congruency should not influence syntactic generation, so discourse congruency effects should *not* be found at the earliest possible point (i.e., initial fixations on the pronoun). Instead, the anomaly effects for mismatched context conditions should arise relatively late. And finally, if structural alternatives are generated/evaluated in parallel, and maintained (to some degree) as long as they are syntactically possible, the processing load will be heavier for noun targets compared to verb targets. Alternatively, if a single structural analysis is chosen at the pronoun, using Minimal Attachment (Frazier, 1978), there should be a garden path for verb targets at the homograph, when the simplest analysis is ruled out.

Method

Participants. Twenty-four undergraduates at the Ohio State University participated in the experiment, for course credit in an introductory

psychology class. All were native speakers of English and had normal, uncorrected vision.

Materials. The four stories used in the experiment were taken intact from Boland (1997b), Experiment 2. Thus, target sentences containing a noun/verb homograph were preceded by a context sentence designed to bias either the noun reading or the verb reading. Example context sentences were given above in (4); example target sentences are repeated in Table 2, which delineates the regions to be used in the analysis. The 16 pairs of target sentences were constructed around 16 different noun/verb homographs (see Appendix A for the list of homographs and a summary of the normative measures described below).

In one member of each target pair, the homograph was the head noun in a possessive NP that began with either *his* or, in one case, *their*. In the other member of each pair, the homograph was the first verb in an embedded clause that began with either *him* ($N = 15$) or *them* ($N = 1$). As illustrated in Table 2, the initial region included all words prior to the main verb: the subject NP plus any preceding phrase (e.g., *meanwhile*), preverbal adverb, or auxiliary verb. The words in this region were always identical across the two target conditions, but they varied in length across items from 1 to 7 words, with a mean length of 2.4 words. The second region was always two words long, including just the main verb and the disambiguating pronoun. The third region contained the homograph and the word *and*. The fourth region contained a noun or verb that was conjoined with the homograph, plus the following word. The conjoined word was always different across the two target conditions; the following word and the two words in the last region were identical in 9 out of the 16 items.

The target sentences were designed to fit naturally into one of four short stories of varying lengths. The first story contained five target sentences, the second had three, the third had seven, and the fourth story had only one target sentence. Target sentences were always separated by at least two other sentences, and no target sentence occurred in the first six sentences of a story. A context sentence preceding the target sentence was designed to bias the reader toward either an upcoming possessive NP or an upcoming embedded clause. Across the set of experimental items, the context sentence operated either by shifting discourse focus between objects and events (as in the *duck* example in (4)) or by manipulating real-world plausibility, often based on spatial relationships within the discourse model. An example of the latter is the item using *steer*. The male participant is in a truck with a dirty rear window. The female participant is either inside the truck where she can watch him steer or behind the truck next to his steer. See Appendix B for additional examples. Noun-biased and verb-biased contexts were crossed with noun and verb targets in a two-by-two factorial design to create four conditions, noun context with noun target, noun context with verb target, verb context with noun target, and verb context with verb target.

Each of the 16 critical items was assigned to one of the four conditions in a pseudo-random fashion on the first stimulus list. The condition of each item was then rotated to form three other experimental lists, such that each item occurred once in each experimental condition and there were an equal number of items in each condition on each list. Each story was followed by a comprehension question, to ensure that participants understood the stories.

Lexical bias norms. The lexical bias of each homograph was estimated in two ways: corpus analysis and sentence completions. According to the Francis and Kučera (1982) norms, which are calculated over the widely used Brown Corpus, the homographs had a mean noun usage of 51 per million and a mean verb usage of 18 per million. The log frequencies are provided in Appendix A. This measure provides a rough estimate of the overall frequency of the noun and verb forms of the homographs.[4] As our second measure of lexical bias, we estimated the contingent frequency of the noun and verb forms in ambiguous contexts like *She saw her*. . . [Agent NP, perception verb, ambiguous pronoun]. For this, we relied upon completion norms collected by Boland (1997b) on the local population that was sampled in the current experiment.[5]

There was a generalized shift toward the verb form in the sentence completions compared to the corpus frequency data. However, upon examining the completion scores and the log frequencies for individual items (both can be found in Appendix A), it was clear that the differences in the two measures of lexical bias cannot be fully explained by a generalized shift. In fact, the completions and log frequencies were not correlated [$r < .10$]. The local syntactic context not only influenced resolution of the homograph; it seems to have impacted resolution in different ways for different items.

Apparatus and procedure. A dental impression was made for each participant. This was attached to an adjustable "bite bar" in order to hold the participant's head in a fixed position. A stationary forehead-rest and an adjustable chin-

---

[4] Of course, the corpus may not provide an accurate estimate of how frequently our experimental participants have actually encountered the noun and verb forms, because (a) it is limited to edited text, (b) all of the texts predate 1965, and (c) one million words is a fairly small sample. Nevertheless, these disadvantages are balanced by two significant advantages. Most importantly, the corpus was compiled by conscientiously sampling from a broad range of genres, including newspapers, novels, nonfiction books and periodicals, government documents, and professional journals. To the best of our knowledge, no other available English corpus has this property. Second, the corpus is widely used and broadly accessible among psycholinguists. Therefore, we opted to use the Brown Corpus, despite its disadvantages.

[5] Boland first determined that the accusative and nominative forms of *her* are about equally frequent in this type of syntactic context, in both corpus norms and sentence completion norms. Then she collected sentence completions on fragments like these to assess the lexical bias of the homograph; she found that 54% of the completions were consistent with the verb form and 33% were consistent with the noun form (the remainder could not be disambiguated).

rest also helped secure the head in a comfortable fashion. Eye movements were measured using a Dr. Bouis monoculor oculometer. The apparatus continuously outputs two voltages, corresponding to eye position along the *X* and *Y* axes. During the experiment, eye position was sampled every millisecond (ms) and converted to screen coordinates.

The sensor of the apparatus was first roughly aligned by mechanical means. It was further adjusted to give zero-output voltages when the participant looked straight ahead, and balanced positive and negative voltages when the participant looked at equidistant points along the *X* and *Y* axes. The experimenter then ran a calibration routine during which the participant was asked to fixate on nine disparate points on the computer screen in order to establish the relationship between *X/Y* voltages and screen position. If this could not be done with an error rate of less than 10 pixels in each dimension, the experiment was aborted. This calibration criterion was used throughout the experiment.

Participants who were successfully calibrated read a practice story, complete with a comprehension question. A short break followed, then the setup was recalibrated before the experiment proper began. Each sentence was presented on the screen in its entirety, on a single line, in a fixed order. The target sentences always appeared alone, but some sentences appeared in pairs, separated by a blank line. For each target sentence, the screen position and duration of each fixation were computed and stored. Most participants completed the experiment in about 15 min. The equipment was recalibrated after each comprehension question to ensure accurate tracking throughout the experiment. Aside from the 24 participants whose data are reported here, four participants' data were omitted because the calibration error was greater than our accuracy criterion during part of the experiment.

Results

We analyzed four dependent measures, expecting to find lexical bias effects in the early measures and discourse congruency effects in the later measures. We assume that an early ef-

fect will be reflected in the first pass dependent measures, including first fixations and first pass reading times, in the region of interest. In contrast, a late effect might be reflected in the second pass reading times or in the first pass measures—but downstream of the critical region. Lexical bias was varied continuously in our stimuli, whereas discourse context was manipulated categorically. Thus, lexical bias effects were evaluated using linear regression, whereas the discourse congruency effects were evaluated using analysis of variance (ANOVA). An effect of discourse congruency would be revealed as the interaction of target type with context type. We also were interested in contrasting the predictions of our syntactic parallelism hypothesis against the predictions of the minimal attachment hypothesis. Both predict an effect of target type in the ANOVAs, but syntactic parallelism predicts greater difficulty for noun targets, while the minimal attachment hypothesis predicts greater difficulty for the verb targets. We present all of the ANOVAs first, and then present the regression analyses. In each of the principal ANOVAs, we included three critical regions, the Verb + Pronoun region, in which the discourse incongruity is first evident, the Homograph region, in which lexical bias effects were predicted, and the Conjoined Phrase region.

ANOVAs. The duration of the *first fixation in a region* provides the earliest available index of processing difficulty in the region. For each participant and item, we computed the mean duration for the first fixation in each region, for each condition. These data are summarized in Fig. 1. To determine whether there was a discourse congruency effect or an effect of target type in the critical region, participant and item means for the three critical regions were submitted to 4(list) $\times$ 3(region) $\times$ 2(target) $\times$ 2(context) ANOVAs. Neither a two-way interaction of target and context [$F1(1,20) = 1.68$, $p > .10$; $F2(1,12) = 1.58$, $p > .10$] nor a three-way interaction of target, context, and region was found [$F1(2,40) = 1.17$, $p > .10$; $F2(2,24) = 1.64$, $p > .10$]. Thus, discourse congruency did not affect the duration of the initial fixations. There
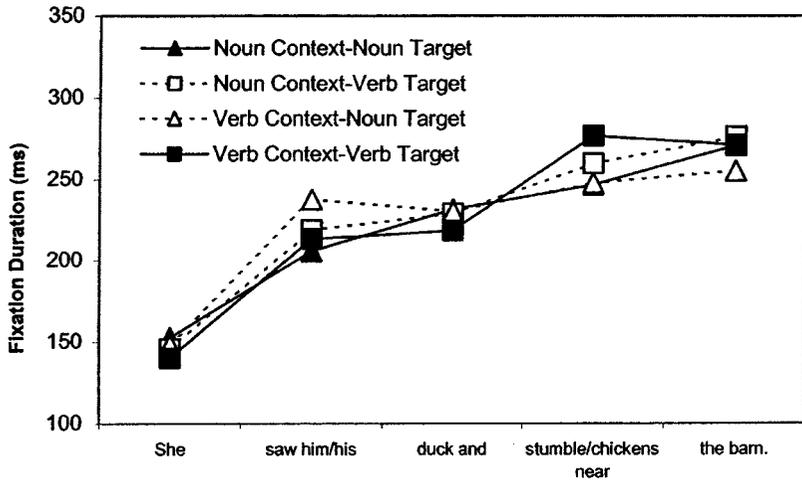
FIG. 1. Mean duration (in milliseconds) of the first fixation in each region, for each of the four experimental conditions.

was no main effect of target type [F's < 1.0], but target interacted with region in the items analysis [$F1(2,40) = 2.31$, $p > .10$; $F2(2,24) = 4.46$, $p < .05$]. Upon examining Fig. 1, the interaction seems to be driven by long fixations in the verb target conditions during the Conjoined Phrase. Such an effect would be roughly consistent with the minimal attachment hypothesis, which predicted a garden path for verb targets at the Homograph, in the prior region. However, a 4(list) $\times$ 2(context) $\times$ 2(target) ANOVA over the Conjoined Phrase region found no effect of

target type [$F1(1,20) = 2.30$, $p > .10$; $F2(1,12) = 2.86$, $p > .10$].

The first fixation duration is an incomplete measure of first pass processing difficulty, because the initial fixation in a region is often followed by another fixation before the reader moves on to the next region. This is especially likely when a region includes two or more words, as our regions did. Therefore, for each participant and item, we computed the mean of the *summed first pass fixation durations* in each region, for each condition. These data are sum-
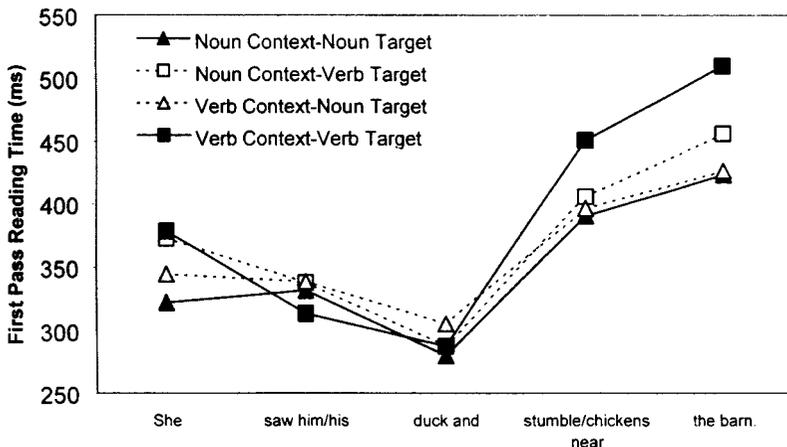

FIG. 2. Condition means (in ms) for the summed first pass fixations in each region.
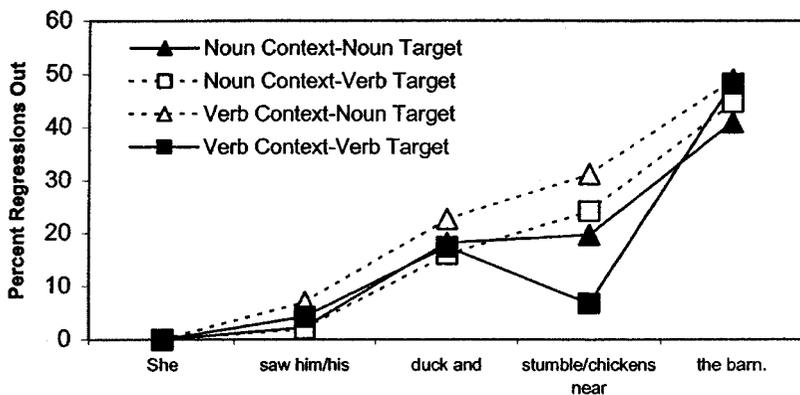
FIG. 3. The probability of a first pass regression out of each region, for each condition.

marized in Fig. 2. Participant and item means from the three critical regions were submitted to 4(list) × 3(region) × 2(target) × 2(context) ANOVAs. As predicted, neither a two-way interaction of target and context nor a three-way interaction of target, context, and region was found [F's < 1.0]. There was no main effect of target type [F's < 1.10], nor was there an interaction between target and region [$F1(2,40) = 1.12, p > .10; F2(1,12) = 1.61, p > .10$].

When a reader experiences difficulty during the first pass reading, the difficulty can be reflected either in longer first pass reading times in the region or by regressive eye movements to earlier regions. (Information about the landing sites for regressions is provided in Appendix C.) The *probability of a first pass regression* out of a region was computed by participant and by item for each condition in each region. These probabilities are summarized in Fig. 3, which suggests that discourse congruity effects began to appear several words downstream of the first incongruous word in the target sentence (the pronoun). Note that the probability of a regression in the Initial region is zero, by definition. The probabilities for the three critical regions were submitted to 4(list) × 3(region) × 2(target) × 2(context) ANOVAs. The main effect of target was reliable by participants and marginal by items [$F1(1,20) = 7.25, p < .05; F2(1,12) = 3.19, p < .10$]. Note that the direction of the effect is reversed from the trend observed in the first fixations: Verb targets exhibited less processing difficulty, as predicted by the syntactic

parallelism hypothesis. There was also an interaction of context and target by participants [$F1(1,20) = 6.13, p < .05; F2(1,12) = 4.71, p < .10$] and a fully reliable three-way interaction of region, context, and target [$F1(2,40) = 6.73, p < .05; F2(2,24) = 7.38, p < .05$]. These interactions are consistent with an effect of discourse congruity. In ANOVAs at the individual word positions, the interaction of context and target was not reliable at the Homograph region [F's < 1.0], but it was reliable at the Conjoined Phrase [$F1(1,20) = 19.17, p < .01; F2(1,12) = 12.35, p < .01$]. Likewise, the main effect of target was absent at the Homograph but present at the Conjoined Phrase [$F1(1,20) = 5.22, p < .05; F2(1,12) = 4.17, p < .10$].

Next, we examined the *second pass reading times*. This dependent measure sums fixations on secondary passes through each region. As shown in Fig. 4, there was a large effect of discourse congruency, beginning in the first or second region. This is to be expected, given our finding that regressive eye movements were more likely in the incongruent conditions. The mean summed duration of second pass fixations in a region was computed by participants and by items for each condition × region cell. These means for the three critical regions were submitted to 4(list) × 3(region) × 2(target) × 2(context) ANOVAs. The expected interaction between target and context was found, confirming that second pass reading times were longer in incongruent conditions [$F1(1,20) = 30.14, p < .01; F2(1,12) =$
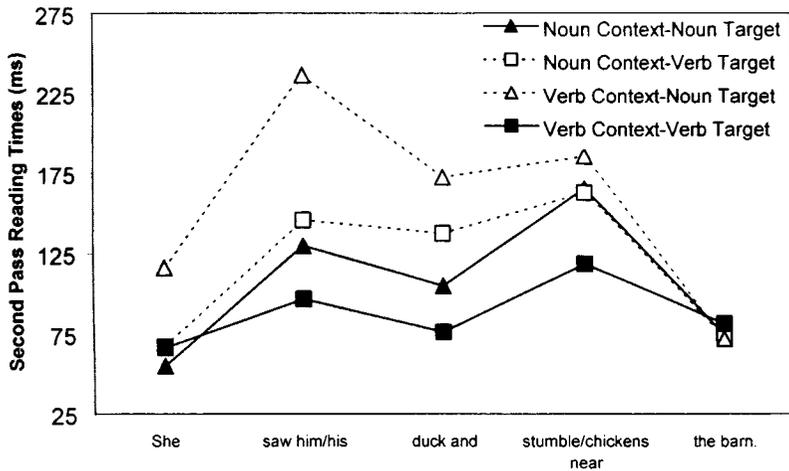
FIG. 4.   Condition means (in ms) for the second pass reading times in each region.

6.86, $p < .05$]. In addition, longer times were found for noun targets compared to verb targets [$F1(1,20) = 15.81$, $p < .01$], but this effect did not generalize across items [$F2(1,12) = 1.99$; $p > .10$].

In sum, discourse congruency effects were clearly present, although they were not found in the most immediate measures of online processing. The discourse congruency effect first emerged at the Conjoined Phrase region in the percentage of regressive eye movements during the first pass through that region. Because there were more regressive eye movements for target sentences following incongruent contexts, second pass reading times for the incongruent conditions were much longer than those for the congruent conditions. In contrast, the target type effects were only partially reliable, and they were much more difficult to interpret. The interaction of target and region that was observed in the item analysis of the first fixation data is roughly consistent with the minimal attachment hypothesis, which predicted a garden path for the verb targets at the Homograph. However, there was no evidence of a garden path for verb targets in the other dependent measures. Quite the contrary, support for syntactic parallelism was found in the participant analyses of the regression data and the second pass reading times. While we do not take these effects as unequivocal support for syntactic parallelism, they are certainly suggestive.

Linear regressions. Next, we evaluated the effects of lexical bias on each of the first pass dependent measures, in both the Homograph region and the Conjoined Phrase region. We began by examining the *initial fixations*. The Homograph region consisted of the homograph and the word *and*, so the first fixation was very likely to have been on the homograph. Thus, the fixation duration may reflect access of the homograph and the addition of the homograph constituent to the developing syntactic representation for the sentence. Of course, we expected syntactic integration to be easier when the lexical bias of the homograph was consistent with the syntactic context. To provide a unidimensional measure of overall lexical bias, we subtracted the log verb frequency from the log noun frequency, as measured in the Francis and Kučera (1982) norms derived from the Brown Corpus. A high value on this overall lexical bias score indicates that the noun form was much more frequent than the verb form. The predicted lexical bias effect would result in a negative correlation between bias and fixation duration for noun targets and a positive correlation for verb targets.

Figures 5a–5d illustrate the relationships between the initial fixation duration in the Homograph region and overall lexical bias, for each context by target condition. Upon inspecting the scatterplots, we removed two problematic

TABLE 3

Correlation Coefficients (*r*) for Initial Fixation Durations, Predicted by Overall Lexical Bias, with the Predicted Direction of the Effect for Each Condition

Condition	Homograph	Conjoined phrase	Predicted direction
Noun context-Noun target	−.55*	−.30	−
Verb context-Verb target	.49 mg	.30	+
Verb context-Noun target	.28	−.72*	−
Noun context-Verb target	.49 mg	.17	+

Note. In all tables, the correlations that are reliable (α = .05) are indicated by asterisks. Marginal correlations (*p* < .10) are indicated by "mg."
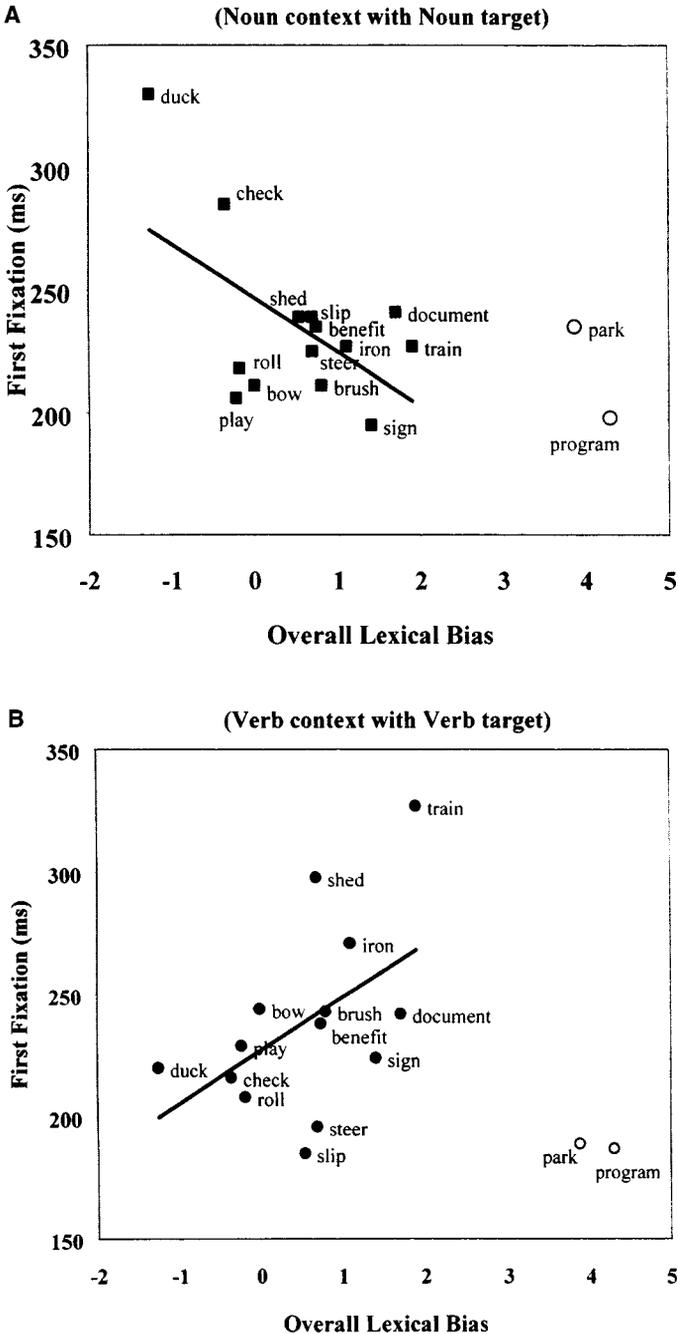
items, *park* and *program*, which are identified by open circles in Fig. 5.[6] The remaining lexical bias values were used to predict initial fixation time in the Homograph region and the Conjoined Phrase region, yielding the correlation coefficients shown in Table 3. Correlations that reliably differ from the null hypothesis are starred (α = .05) and those marginally different are annotated "mg" (.05 < *p* < .10). All correlations are in the predicted direction, except for the verb context with noun target condition in the Homograph region. As noted above, Boland (1997b) found lexical bias effects earlier for the congruent discourse conditions compared to the incongruent conditions. We suspected that the difference in timing was due to difficulties in resolving the lexical form of the noun target in in-

congruent discourse contexts. These suspicions were confirmed by the data pattern reported in Table 3. At the Homograph, the correlation in the noun context with noun target condition was reliably different from that in the verb context with noun target condition [*z* = 2.11, *p* < .05], while there was no difference in the correlations for the two conditions with verb targets. Unfortunately, the predicted correlations between lexical bias and the congruent conditions at the Homograph were not fully reliable. To reduce itemwise noise, we followed the strategy used in Boland (1997b). We subtracted the mean fixation duration for the verb target from the mean fixation duration for noun target, for the congruent conditions of each item. Analyzed in this manner, the lexical bias values were a reliable predictor of initial fixation time for congruent conditions, as expected [*F*(1,12) = 7.32, *p* < .05; *r* = −.62].

We then regressed the initial fixation durations against contingent lexical bias, which was computed as the percentage of verb completions minus the percentage of noun completions, from Boland's (1997b) sentence completion data. In contrast to the overall bias data, the predicted pattern of positive and negative correlations was observed only for the incongruent conditions, as shown in Table 4. None of the regressions were statistically reliable.

Next, we examined the *summed first pass fixation times*. In contrast to the initial fixations, the predicted pattern of correlations was not found in either the Homograph region or the Conjoined Phrase region, using either overall lexical bias or contingent frequency as the predictor variable. As shown in Table 5, the data

---

[6] These items were strongly noun-biased, according to the log frequencies from the Brown Corpus (see Appendix A, columns 1 and 2), but they were verb-biased in Boland's (1997) sentence completions (columns 3 and 4). At first, we suspected that the Brown Corpus was outdated; the increased importance of computers has no doubt increased the use of *program* as a verb quite dramatically since the early 1960s, when the Brown Corpus was compiled. However, these items were even more strongly noun-biased in a current textual corpus, the web-based Chicago News Network archives. (Details are available from the first author. In brief, we analyzed the 100 most recent articles containing the word *program* or *park*, using only the first token from each article.) Thus, there is a clear discrepancy between the textual corpora and the completion norms. Unfortunately, the completion data are problematic because they may be biased by the particular sentence fragments that were used in the completion task. (This topic is explored further in the discussion section.) To be consistent, the *program* and *park* values were excluded from all regressions using the Brown Corpus measure of overall lexical bias.

FIG. 5. The relationship between overall lexical bias (log noun frequency minus log verb frequency) and the duration of the first fixation in the Homograph region. Outliers are shown as open circles and were omitted from the analysis. Figure 5a illustrates the correlation for the noun context with noun target condition. Figure 5b illustrates the marginally reliable correlation for the verb context with verb target condition.

C

(Verb context with Noun Target)

D
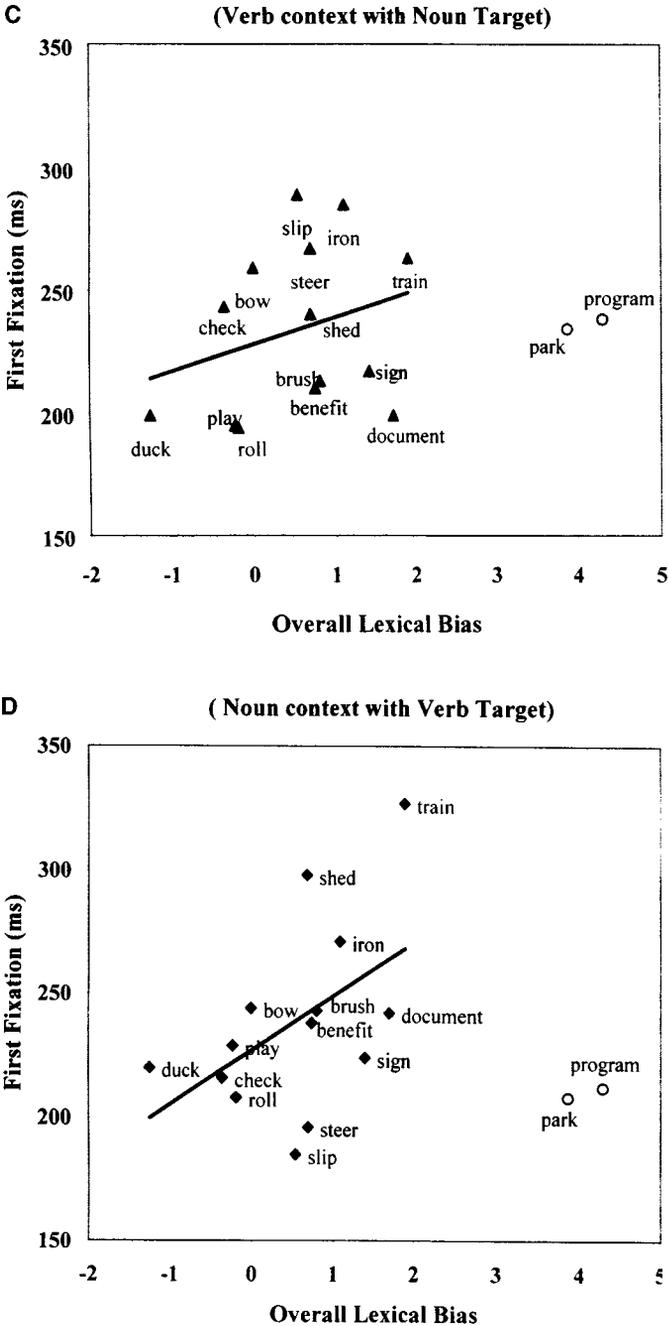
(Noun context with Verb Target)

FIG. 5—*Continued.* Figure 5c reveals the lack of correlation in the verb context with noun target condition. Figure 5d shows the marginal correlation for the noun context with verb target condition.

TABLE 4

Correlation Coefficients (*r*) for Initial Fixation Durations, Predicted by Contingent Lexical Bias, with Predicted Direction of the Effect for Each Condition

Condition	Homograph	Conjoined phrase	Predicted direction
Noun context-Noun target	.22	−.12	−
Verb context-Verb target	.00	−.15	+
Verb context-Noun target	−.22	−.49 mg	−
Noun context-Verb target	.41	.43	+

looked a little more coherent when we examined the *probability of a first pass regression out of a region*. Overall lexical bias did not yield the predicted pattern in the Homograph region, but it did at the Conjoined Phrase. Unfortunately, only one regression was reliable (α = .05), and none of the others approached reliability. The predicted pattern was not obtained in either region, using contingent lexical bias as the predictor variable.

In sum, the overall lexical bias of the homograph was a much better predictor of processing difficulty than contingent lexical bias, but even so, the lexical bias effects were almost completely limited to the initial fixation data. Interestingly, Boland (1997b) found that contingent frequency was a better predictor of word-by-word reading times in her button-pressing task.

Discussion

Noun/verb homographs represent an important class of lexical ambiguity that is tightly entwined with phrase structure ambiguities. Constraint-based models of sentence processing predict that the relative frequency of the noun and verb forms should predict the accessibility of the two syntactic forms, and confirming evidence has been reported in self-paced reading tasks (Boland, 1997b; Boland & Lewis, 1998;

Corley, 1998; MacDonald, 1993). However, the current study is the first report of such effects using an eye movement paradigm. We analyzed four dependent measures from the eye movement record to obtain detailed information about the timing with which lexical bias and discourse congruency influenced processing. While it is always optimal to gather converging evidence from multiple paradigms, the eye tracking evidence reported here is especially important, because it demonstrates that the lexical bias effects emerged in the very first fixation on the noun/verb homograph. In contrast, anomaly effects from incongruent discourse context did not begin to emerge until several words downstream of the incongruity, in the percentage of regressive eye movements. Robust discourse congruency effects were also observed in the second pass reading times.

It is important to remember that the lexical bias effects reported here arose during syntactic generation, not syntactic selection. The noun form of the homograph cannot be grammatically attached in the context of *She saw him*. . ., so we assume that no such attachment was generated. Even so, the relative frequency of the noun and verb forms influenced the processing difficulty of generating the appropriate structure—as evidenced by a lexical bias effect during the first

TABLE 5

Correlation Coefficients (*r*) for Percentage of Regressions Out of the Region, Predicted by Overall Lexical Bias, with Predicted Direction of the Effect for Each Condition

Condition	Homograph	Conjoined phrase	Predicted direction
Noun context-Noun target	.08	−.57*	−
Verb context-Verb target	.22	.03	+
Verb context-Noun target	.23	−.28	−
Noun context-Verb target	.36	.37	+

Note. Asterisk indicates reliable correlation (α = .05).

fixations on the homograph in three out of four conditions. The lexical bias effects were delayed in the noun target with verb context condition, presumably because the lexical ambiguity was resolved more slowly.

This pattern of lexical bias effects is similar to Boland's (1997b) findings using self-paced reading, except that she found contingent lexical bias to be the best predictor of reading time, whereas we found overall lexical bias to be a better predictor. Contingent frequency combines local contextual information with lexical information whereas overall frequency is a purely lexical constraint. Because the nature of the frequency information is an important issue in constraint-based models, let us consider the implications of models that rely upon contingent frequency versus overall frequency.

Tabor et al. (1997) investigated several different types of frequency effects on the determiner/complementizer syntactic category ambiguity of *that*. Like Boland (1997b), they found that syntactically contingent frequencies were better than the overall category frequencies in predicting self-paced reading times. For example, *that* is most often a determiner when it is sentence initial but is most often a complementizer when it follows a verb. These contingent frequencies seemed to guide syntactic ambiguity resolution; garden path effects were found when the syntactic category of *that* was disambiguated as different from its contingent bias. In the dynamical systems model of Tabor et al., the contingent frequencies are expected to predict reading time in both ambiguous and unambiguous syntactic contexts, because the different syntactic contexts correspond to different locations in the representational space of the system (as in Elman, 1990). However, Gibson and Tunstall (1999) have suggested that the Tabor et al. results can be explained using overall lexical frequency, without recourse to contingent frequencies.

In contrast to Tabor et al. (1997), the model assumed in Boland (1997b) predicts that overall lexical biases should be most relevant for unambiguous syntactic contexts. In the Boland account, effects of lexical bias on syntactic generation arise because competing syntactic forms are accessed during word recognition, and a word's more frequent forms are more strongly

available. Thus, Boland (1997b) had to assume that her contingent frequency data were simply better estimates of overall frequency than her corpus data. Such an explanation was plausible, especially given our claim that the Brown Corpus underestimated the verb frequencies of *park* and *program*. However, the current data suggest that the Brown Corpus does actually provide a reasonable estimate of overall frequency after filtering out those two problematic items. (There is a growing literature comparing corpus and completion measures as estimates of the strength of lexical representations. See Roland and Jurafsky, in press, as an example.)

The discrepancy between the Boland (1997b) self-paced reading data and the current eye movement data may be due, at least in part, to the nature of the dependent measures. In the current experiment, the lexical bias effects were quite transient, so the lexical bias effects in the self-paced study may have reflected a different phenomenon (i.e., some combination of overall bias effects on syntactic generation and other factors that affect sentence interpretation). If the contingent frequencies incorporated the impact of local contextual information (such as *She saw her. . .*), rather than reflecting pure lexical accessibility, it is not surprising that more global measures of processing, such as self-paced reading, reflected local context effects to a larger degree.

On the other hand, discourse congruency effects were much more apparent in the eye movement data reported here than in Boland (1997b). This might raise the concern that discourse-level integration was hampered by the self-paced reading paradigm. However, recall from Table 1 that Boland demonstrated that discourse information from prior sentences was immediately used to guide syntactic ambiguity resolution. Thus, even in the self-paced reading paradigm, discourse information was immediately and efficiently used to guide syntactic selection. It is more likely that the self-paced reading paradigm, with its rigid button-pressing task, was simply insensitive to the relatively subtle anomaly effects that arose when the discourse context was incongruent. Eye tracking, with its multiple dependent measures, offers more opportunities to observe subtle effects.

Despite limitations in our measures of lexical bias, our data clearly support the view that syntactic generation involves the access and integration of lexical structure, with the dominant syntactic form being accessed more strongly than subordinate forms. In contrast, discourse congruency was reflected in the second pass times and in the probability of regressive eye movements. The size of the second pass effects is striking, given that Boland (1997b) failed to find consistent discourse congruency effects with these sentences using self-paced reading. As predicted, the timing of the discourse congruity effects in the current experiment contrasts sharply with the timing of the lexical bias effects, and this suggests that expectations generated from the discourse did not influence syntactic generation. Rather, the effects appear to be anomaly detection effects, in response to the incongruity of the target sentence with the discourse context. Thus, these data provide additional evidence for Boland's distinction between syntactic generation and syntactic selection, with lexical bias—but not discourse congruency—influencing syntactic generation.

Our data obviously challenge syntactic processing models in which all relevant and available constraints are brought to bear uniformly and simultaneously (e.g., Spivey & Tanenhaus, 1998; Tabor et al., 1997). Rather, the results of the experiment presented here, in combination with the results of Boland (1997b), support a constraint-based model with at least two categories of constraints. The first category includes the relative strength of competing lexical forms (i.e., lexical bias) and affects the generation of syntactic alternatives. The second category of constraints, which subsumes the first and also includes discourse bias, affects the selection of one structure from among generated alternatives. Granted, our results might be consistent with the Spivey and Tanenhaus model if they were to add a syntactic generation component that is dominated by syntactic and lexical constraints. However, such a move would change the character of their model dramatically.

Our comparison of lexical and discourse constraints may bring to mind other investigations of unambiguous sentences in which investigators have contrasted different processing levels.

For example, McElree and Griffith (1995) found that readers were faster to detect syntactic anomalies (e.g., *Some people agree books*) compared to thematic role anomalies (e.g., *Some people alarm books*) in a paradigm using a speed–accuracy trade-off. Their primary claim was that syntactic representations are established prior to richer thematic representations. On the other hand, Ni et al. (1998) found that similar anomalies (e.g., *cats won't usually bake*) were detected as rapidly as syntactic anomalies, but were reflected in different aspects of the eye movement record. Syntactic anomalies induced immediate regressive eye movements while semantic anomalies resulted in longer first pass reading times as well as regressive eye movements. Unfortunately, our current data cannot help to resolve the discrepancy between McElree and Griffith and Ni et al., because we did not include similar thematic/semantic anomalies in our stimuli.

The effects of target sentence type found here are also quite interesting, because they provide data that help distinguish two models of syntactic ambiguity resolution. The relevant syntactic ambiguity is in whether the constituent beginning at the pronoun is attached as a direct object NP or as part of a clause. The minimal attachment heuristic predicted processing difficulty at the homograph for verb targets when the minimal structure was ruled out. No such difficulty was predicted for the noun targets, because they used a minimal structure. Although there was a trend toward a garden path for verb targets at the Conjoined Phrase in the initial fixation data, no reliable effects were found. In contrast, we found more regressive eye movements and longer second pass reading times for noun targets in the critical region. Thus, there is little support for the view that the minimal attachment analysis was initially adopted.

On the other hand, our data are consistent with a ranked parallel account. This account predicted a heavy processing load for the noun targets compared to verb targets late in the sentence, because the direct object/clause ambiguity was disambiguated much later for the noun targets. It may or may not be meaningful that the target type effects occurred in the percent-

age of regressive eye movements and second pass reading rather than longer first pass fixations. Both behavioral responses have the effect of spending more time in a difficult region before moving on. Although some researchers link regressive eye movements to syntactic re-analysis (e.g., Frazier & Rayner, 1982; Ferreira & Henderson, 1990; but see Rayner et al., 1989, for a different view), we believe that our understanding of eye movements is still too limited to infer the underlying cognitive process on the basis of an eye movement pattern. We must acknowledge that our data would be consistent with a serial model in which garden paths were predicted for the noun targets. However, such a model would have to predict that readers initially attached *his duck* as the subject of an embedded clause and then reanalyzed it as a direct object. We know of no parsing models that would make such a prediction.

The ranked parallel account also explains an otherwise curious trend found by Boland (1997b), in which verb targets tended to be faster than noun targets near the end of the sentence. However, the itemwise variation observed in our target type effects suggests that parallel structures were not always maintained until definitive syntactic information was encountered. This is not surprising, given numerous other results—including some from our own laboratory—demonstrating that syntactic commitments are often made at the point of ambiguity (e.g., Boland & Boehm-Jernigan, 1998; Mac-Donald, 1994; Tabor et al., 1997; Trueswell et al., 1994). We suspect that the likelihood of maintaining parallel representations is probabilistically determined by the amount of support for a given analysis, assessed over a wide range of constraints. Such an account is consistent with the noncompetitive, ranked parallel model proposed by Pearlmutter and Mendolsohn (1999, unpublished manuscript).

APPENDIX A

This table provides the normative data for our experimental items. The first two columns list the log of the Kučera and Francis (1982) noun and verb frequencies. The last two columns list the percentage of noun and verb completions for ambiguous fragments like *She saw her duck. . . .* Means for all measures are provided in the bottom row.

	Log(KF frequencies)		Sentence completions	
	Noun	Verb	% Noun	% Verb
duck	0.69	1.95	70	20
steer	1.79	1.10	30	70
shed	1.39	0.69	20	70
roll	2.71	2.89	10	90
brush	3.37	2.56	50	50
sign	4.29	2.89	65	15
park	3.87	0	0	90
slip	2.48	1.95	0	90
play	4.48	4.70	0	70
train	4.20	2.30	80	10
check	3.58	3.93	40	50
program	5.91	1.61	20	30
benefit	3.74	3.00	40	40
document	2.40	0.69	60	40
bow	2.56	2.56	30	60
iron	1.10	0	20	70
Mean	**3.04**	**2.05**	**33.44**	**54.06**

APPENDIX B

An excerpt from an example story is presented below, with both versions of the context sentences and both versions of the target sentences. In the experiment there were no paragraph breaks, because each sentence was presented individually.

This is a story about a man named Wilbur. Wilbur is the mayor of a small town called White Bluffs. One of his best friends is his neighbor, Ann. Ann worked as Wilbur's campaign manager. Before Wilbur became a politician, he felt strongly about the environment.

Noun context: *And when a clean water petition came around he decided to run for mayor.*

Verb context: *But when a clean water petition came around he didn't know what to do.*

Noun target: *Ann made his sign and posters before the election.*

Verb target: *Ann made him sign and talk with polluters about the problem.*

As the mayor, Wilbur convinced the town council to develop a nature preserve. He found a wooded area in the foothills that would be perfect. They were able to purchase it within a month. Wilbur was very proud of it, and drove his two daughters out to the site.

Noun context: *The girls liked to ride in the car and they were also anxious to see the spot.*

Verb context: *The girls didn't like to ride in the car, but they agreed to go along.*

Noun target: *They seemed relieved when they saw his park and campgrounds in the mountains.*

Verb target: *They seemed relieved when they saw him park and point near the mountains.*

APPENDIX C

The table below presents the probability of a regressive eye movement landing in a given region. Although none of our predictions hinged upon these data, they are consistent with the patterns found in the regression and second pass analyses. The Initial region was more likely to be refixated for noun targets, compared to verb targets. This is consistent with the increased processing load attributed to the noun targets. The Verb + Pronoun region was most likely to be the goal of a regressive eye movement in the incongruent conditions. Thus, when participants made regressive eye movements from the Conjoined Phrase and Final regions because of the discourse incongruity, they were often regressing to the region of the sentence where the incongruity originated.

	Initial	Verb + pronoun	Homograph	Conjoined Phrase	Final
Noun context-Noun target	.39	.36	.29	.28	.08
Noun context-Verb target	.28	.42	.27	.26	.08
Verb context-Noun target	.46	.54	.43	.20	.05
Verb context-Verb target	.29	.32	.21	.28	.10

REFERENCES

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

Binder, K. S., & Morris, R. K. (1995). Eye movements and lexical ambiguity resolution: Effects of prior encounter and discourse topic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21,** 1186–1196.

Binder, K. S., & Rayner, K. (1998). Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review*, **5,** 271–276.

Binder, K. S., & Rayner, K. (1999). Does contextual strength modulate the subordinate bias effect? A reply to Kellas and Vu. *Psychonomic Bulletin & Review*, **6,** 518–522.

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

Boland, J. E. (1997b). Resolving syntactic category ambiguities in discourse context: Probabilistic and discourse constraints. *Journal of Memory and Language*, **36,** 588–615.

Boland, J. E., & Boehm-Jernigan, H. (1998). Lexical constraints and prepositional phrase attachment. *Journal of Memory and Language*, **39,** 684–719.

Boland, J. E., & Lewis, R. (1998). The lexical generation of syntax and its implications. The 39th Annual Meeting of the Psychonomics Society, Dallas.

Carpenter, P. A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixation. *Journal of Verbal Learning and Verbal Behavior*, **20,** 137–160.

Corley, S. (1998). *A statistical model of human lexical category disambiguation*. Unpublished doctoral dissertation, University of Edinburgh.

Dopkins, S., Morris, R. K., & Rayner, K. (1992). Lexical ambiguity and eye fixation in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language*, **31,** 461–476.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, **27,** 429–446.

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

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.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation, University of Connecticut.

Frazier, L. (1987). Theories of sentence processing. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural language understanding*. Cambridge, MA: MIT Press.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, **14,** 178–210.

Gibson, E. A. F. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. Unpublished doctoral dissertation, Carnegie-Mellon University.

Hogaboam, T. W., & Perfetti, C. A. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, **14,** 265–274.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99,** 122–149.

Kellas, G., Paul, S. T., Martin, M., & Simpson, G. B. (1991). Contextual feature activation and meaning access. In

G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 47–71). Amsterdam: Elsevier.

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

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, **9,** 157–201.

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

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of context as a determinant of the subordinate bias effect. *Quarterly Journal of Experimental Psychology*, **52A,** 813--839.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21,** 134–157.

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. *Journal of Psycholinguistic Research*, **27,** 515–539.

Pacht, J. M., & Rayner, K. (1993). The processing of homophonic homographs during reading: Evidence from eye movement studies. *Journal of Psycholinguistic Research*, **22,** 251–271.

Paul, S. T., Kellas, G., Martin, M., & Clark, M. B. (1992). Influence of contextual features on the activation of ambiguous word meanings. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18,** 703–717.

Rayner, K., Binder, K. S., & Duffy, S. A. (1999). Contextual strength and the subordinate bias effect. *Quarterly Journal of Experimental Psychology*, **42,** 841–852.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language*, **33,** 527–544.

Rayner, K., Sereno, S. C., Morris, R. K., & Schmauder, A. R. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, **4,** 121–149.

Roland, D., & Jurafsky, D. (in press). Verb sense and verb subcategorization probabilities. In P. Merlo & S. Stevenson (Eds.), *Sentence processing and the lexicon: Formal, computational, and experimental perspectives,* John Benjamins: Amsterdam.

Simpson, G. B., & Krueger, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language*, **30,** 627–643.

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, & Cognition*, **24,** 1521–1543.

Stevenson, S. (1998). Parsing as incremental restructuring. In J. D. Fodor & F. Ferreira (Eds.), Reanalysis in sentence processing, pp. 327–363. Kluwer: Boston.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, **18,** 645–659.

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.

Tabor, W., & Tanenhaus, M. K. (1999). Dynamical models of sentence processing. *Cognitive Science*, **23,** 491–550.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, **18,** 427–440.

Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, **27,** 597–632.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research*, **49,** 161–167.

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,** 285–318.

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