EXAMINING INTER-SENTENTIAL INFLUENCES ON PREDICTED VERB SUBCATEGORIZATION

by

Jill Louise Brady

BS, Clarion University of Pennsylvania, 1995

MS, Clarion University of Pennsylvania, 1997

Submitted to the Graduate Faculty

of School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Jill Louise Brady

It was defended on

July 21, 2008

and approved by

Scott Yaruss, PhD, Associate Professor, Department of Communication Science and Disorders

Charles Perfetti, PhD, Professor, Department of Psychology

Christine A. Dollaghan, PhD, Professor, School of Behavioral and Brain Sciences, University of Texas at Dallas

Connie A. Tompkins, PhD, Professor, Department of Communication Science and Disorders

EXAMINING INTER-SENTENTIAL INFLUENCES ON PREDICTED VERB SUBCATEGORIZATION

Jill Louise Brady, PhD

University of Pittsburgh, 2008

This study investigated the influences of prior discourse context and cumulative syntactic priming on readers' predictions for verb subcategorizations. An additional aim was to determine whether cumulative syntactic priming has the same degree of influence following coherent discourse contexts as when following series of unrelated sentences. Participants ($N = 40$) read sentences using a self-paced, sentence-by-sentence procedure. Half of these sentences comprised a coherent discourse context intended to increase the expectation for a sentential complement (S) completion. The other half consisted of scrambled sentences. The trials in both conditions varied according to the proportion of verbs that resolved to an S (either 6S or 2S). Following each condition, participants read temporarily ambiguous sentences that resolved to an S. Reading times across the disambiguating and postdisambiguating regions were measured. No significant main effects or interactions were found for either region. However, the lack of significant findings for these analyses may have been due to low power. In a follow-up analysis, data from each gender were analyzed separately. For the data contributed by males, there were no significant findings. For the data contributed by females, the effect of coherence was significant (by participants but not by items) across the postdisambiguating region, and there was a marginally significant interaction ($p = .05$) between coherence and frequency across this region suggesting that discourse-level information may differentially influence the local sentence processing of female and male participants.

TABLE OF CONTENTS

LIST OF TABLES

PREFACE

I would first like to thank my committee members: Connie Tompkins, PhD, Christine Dollaghan, PhD, Charles Perfetti, PhD, and J. Scott Yaruss, PhD, for their efforts in helping me to shape and refine my ideas for this project, and for providing me with excellent feedback and suggestions. I would especially like to thank my dissertation advisor, Connie Tompkins, PhD, not only for her generous time commitment to this project, but also for the genuine warmth, interest, and enthusiasm that she has expressed over the past two years.

I am also grateful to my parents and my brother, Jason, for their continued support throughout my academic career, and to my friends and colleagues who have supported me in numerous ways. In particular, I would like to thank Stephanie Nixon, PhD, Jackie Bauman-Waengler, PhD, Karen Stein, MA, and David Stein, PhD for their support and encouragement at various times during my doctoral studies. In addition, I owe a debt of gratitude to several graduate students at Indiana University of Pennsylvania who willingly participated in the norming experiments and/or helped me with participant recruitment. They are Amber Pangallo, Jeana Tate, Talayne Suder, Angela Sodano, and Elizabeth Mudrey.

Finally, I dedicate this dissertation to my grandmother, Mrs. Violet Flaig Rieck (1925-2005). She will always be my role model, and her compassion and honesty will continue to inspire me for the rest of my life.

1.0 INTRODUCTION

Sentence comprehension involves not only processing words and structures as they occur, but also predicting upcoming words and structures. Because making correct predictions facilitates language processing, comprehenders consider various sources of information in making their predictions. Their predictions are influenced not only by information from within the sentence (e.g., semantic and syntactic information), but also by information from previously encountered sentences. While there is general agreement that sentence comprehension is influenced by prior sentences, varying accounts have been offered to explain the precise mechanism(s) by which earlier occurring sentences influence the comprehension of later ones.

One explanation regarding how sentence comprehension is influenced by the nature of the previously occurring sentences is that information is integrated across sentences to form a global representation or *discourse context* (Cook & Myers, 2004; Hess, Foss, & Carroll, 1995; Schwanenflugel & White, 1991; Sharkey & Sharkey, 1992; Traxler, Foss, Seely, Kaup, & Morris, 2000). The discourse context reflects the accumulated "meaning" of the sentences up to the point at which the current word or sentence is being processed. There is strong evidence that words and syntactic structures that are more closely related to the discourse context (or are more predictable based on the discourse context) are processed more easily than those that are not. A number of research studies testify that information from the prior discourse context can aid listeners and readers in resolving both structural ambiguities (Altmann & Steedman, 1988; Britt,

Perfetti, Garrod, & Rayner, 1992; Grodner, Gibson, & Watson, 2005) and ambiguities involving homonym meaning (Hare, McRae, & Elman, 2003; Vu, Kellas, Metcalf, & Herman, 2000). There is also evidence that discourse context can influence predictions about upcoming words (Cook & Myers, 2004; Hess, Foss, & Carroll, 1995; Schwanenflugel & White, 1991; Sharkey & Sharkey, 1992).

Because it is widely held that listeners and readers use information from the prior discourse context to resolve the temporary ambiguities that are encountered during sentence processing, as well as to make predictions about upcoming words and structures, it seems likely that they would also use discourse context information to predict upcoming verb subcategorizations. A verb's *subcategorization* is its specification for required and obligatory upcoming syntactic structures. Because many verbs have more than one subcategorization, they are ambiguous regarding the following structure at the point at which they are encountered. For example, the verb *believe* can subcategorize for either a noun phrase [NP] (as in *Robert believes the story*) or for a sentential complement [S] (as in *Robert believes that Joan is telling the truth*). Potential subcategorizations are regarded as part of the verb's lexical entry, resulting from an individual's experience with that verb (Boland, 1993; Jennings, Randall & Tyler, 1997; Jurafsky, 2002; Jurafsky, 1996; Mitchell, Cuetos, Corley & Brysbaert, 1995; Trueswell, Tanenhaus, & Kello, 1993).

Because predicting verb subcategorizations involves the resolution of a type of temporary local ambiguity (Trueswell et al., 1993), it seems likely that discourse context could influence this type of syntactic ambiguity resolution in addition to its known influence on other types of structural, semantic and lexical ambiguities. However, there is another potential influence on the prediction of verb subcategorizations known as *syntactic priming*, which occurs when a

previously processed syntactic structure facilitates processing of the same syntactic structure. While most research in syntactic priming has investigated local syntactic priming effects among adjacent sentences (Bock, 1986; Bock & Loebell, 1990; Hartsuiker & Kolk, 1998; Noppeney & Price, 2005; Pickering & Branigan, 1998; Potter & Lombardi, 1998), Kashak, Loney, and Borreggine (2006) report evidence of a *cumulative* syntactic priming effect, in which the influence of syntactic priming accumulates across trials or sentences. In particular, they found that the likelihood that participants would produce a particular structure was related to its frequency across previously occurring trials. According to Kashak et al. (2006), the existence of cumulative syntactic priming suggests that syntactic priming reflects more than short-term activation, and is the result of an adaptation of the language processing system that is sensitive to "the cumulative effect of recent experience" (p. B74).

In addition, there is evidence that syntactic priming involves more than local priming effects and can persist across several sentences. Bock and Griffin (2000) found evidence that the influence of syntactic priming can persist across as many as ten unrelated sentences. Kashak et al. (2006) reported that the influence of syntactic priming on the proportion of produced target structures did not vary significantly between conditions in which the priming sentences were placed either in the first half or second half of a set of twenty sentences that were presented before the target sentence.

Taken together, the results of these studies suggest that an individual's prediction for verb subcategorizations could be influenced by the prior discourse context, cumulative syntactic priming, or both. However, no known study has attempted to determine whether each of these factors has a significant (and independent) influence on predicting verb subcategorizations. Separating these influences requires a well-controlled study, because there is evidence that these

two variables are related. The results of three corpus-based studies (Biber, 1998; Roland, 2001; Roland & Jurafsky, 2002) suggest that the frequency of subcategorizations associated with specific verbs may vary with other characteristics of discourse context (e.g., formality, modality, and purpose).

Moreover, it is unknown whether syntactic priming would exert the same degree of influence on the processing of syntactic structures within a discourse context as has been demonstrated across sets of unrelated sentences. Series of unrelated sentences provide the listener or reader with no global representation to which to attend. Therefore, the syntactic structures of these sentences might be more salient to the individual than would be the case during natural language comprehension.

The purpose of this investigation was to determine whether discourse context and cumulative syntactic priming exert independent influences on predicting verb subcategorizations. In addition, because the influence of syntactic priming has been studied almost exclusively across sets of unrelated sentences, another aim of this study was to determine whether the influence of cumulative syntactic priming is more pronounced following a series of unrelated sentences than following a coherent discourse context.

Examining various potential sources of inter-sentential influence on predicted verb subcategorizations has implications for the development of augmentative and alternative communication (AAC) devices for individuals with severe communication disorders. Recently, this technology has incorporated *linguistic prediction*, in which the system is equipped with algorithms which enable it to predict the next word or phrase based on the portion of the sentence that has been produced up to that point. This technology produces keystroke savings and reduces fatigue in individuals who use these systems. However, the algorithms that are used

by these systems to predict upcoming words and phrases are based only on sentence-level information. Perhaps additional keystroke savings could be achieved if intersentential sources of information were also taken into account.

The results of this investigation also might have implications for further research in the area of language disability. If the outcome of this experiment suggests that the syntactic predictions of individuals without a history of language impairment are influenced by either discourse context information or syntactic priming, then it would be important to determine whether these factors influence the predictions made by individuals with language impairments. If such information does not have a similar influence on the predictions made by these individuals, training these individuals to make more efficient use of intersentential information might improve the efficiency of their language processing.

1.1 THE EFFECT OF VERB SUBCATEGORIZATION BIAS ON PREDICTING UPCOMING STRUCTURES

As stated earlier, a verb's specification for optional and obligatory upcoming syntactic structures is known as its *subcategorization*. In addition, there is evidence that individual verbs are associated with a preferred subcategorization (Ford, Bresnan & Kaplan, 1982; Jurafsky, 1996), which is held to be the one with which it has most frequently co-occurred in the comprehender's experience. Mitchell, Cuetos, Coley, and Brysbaert (1995) proposed a *linguistic tuning hypothesis*, which claims that encountering a verb associated with a particular subcategorization increases the expectation for that subcategorization on subsequent encounters

with that verb. Several theorists have proposed that processing is facilitated for syntactic structures that are associated with a verb's preferred subcategorization relative to those associated with a less preferred structure (Boland, 1993; Jennings, Randall, & Tyler, 1997; Jurafsky, 1996; Mitchell, et al., 1995; Trueswell, Tanenhaus, & Kello, 1993).

Both representational (e.g., Gibson, 1998, 2000) and connectionist (e.g., Elman, 1991, 1992; MacDonald & Christiansen, 2002) models of language processing can account for the increased ease of processing associated with preferred verb subcategorizations. According to one representational model, Gibson's Discourse Locality Theory, (DLT; 1998, 2000), predictions for upcoming syntactic structures are based on phrase structure rules (e.g., S → NP VP; VP →V (NP)). The results of Chen, Gibson, and Wolf (2005) and Warren and McConnell (2006) suggest that this rule-based system can be extended to predictions for verb subcategorizations. The assumption of this rule-based system is that categories and subcategories of words (such as a verb's associated subcategorizations) are inherent properties of the words themselves, and the ways in which they can be combined is stored as part of the word's lexical entry. A verb's association with a particular subcategorization is the result of an individual's experience with that particular verb across a variety of contexts, and its preferred subcategorization is the one with which it most frequently co-occurs.

On the other hand, connectionist models (Elman, 1991, 1992; MacDonald & Christiansen, 2002) posit that the subcategorizations associated with verbs are not inherent properties of the verbs themselves, and are learned by the co-occurrence between the verb and certain types of subcategorizations. Connections are established between the verb and the subcategorizations with which it co-occurs, with stronger connections (greater connection weights) associated with the subcategorizations that occur most frequently with the verb.

As stated earlier, both classes of models would predict that the processing of preferred subcategorizations would be facilitated. For ease of exposition, the discussion to follow is framed within the context of representational models of language processing. However, it is important to note that many of the findings to be described can also be explained in terms of connectionist models.

Evidence for the use of preference information on the processing of subcategorizations comes from several studies (Boland, 1993; Ferreira & Henderson, 1990; Jennings et al., 1997; Trueswell et al., 1993). Table 1 (Appendix A) provides an overview of the results of these studies. While these studies are consistent in their finding of the influence of preference information at some point during sentence processing, they are less consistent regarding whether preference information influences initial (first-pass) sentence parsing, or only has an influence at a later reanalysis stage of processing.

The question of whether this information influences first pass sentence parsing has *implications for refuting one of two general positions on sentence processing. Proponents of the autonomous viewpoint* (e.g., Connine, C.M., Ferreira, F., Jones, C., Clifton, C., & Frazier, L., 1984; Frazier, 1987) propose that various types of linguistic knowledge operate independently. This position holds that, while syntactic information is used during the initial (first pass) parsing of sentences, lexical, discourse context, and other information is used only during a later reanalysis or integration (second pass) processes. On the other hand, proponents of an *interactive viewpoint* (e.g., Trueswell, et al., 1993) posit that various sources of information interact during the initial parsing of sentences. There is, however, a wide variety of views within each of these theoretical positions.

Proponents of the autonomous viewpoint would predict that verb subcategorization preference would not influence first-pass sentence parsing, and would only be used during a later reanalysis stage of processing. Support for this viewpoint comes from Ferreira and Henderson (1990). They used eye tracking (Experiment 1) and self-paced reading (Experiments 2 and 3) to investigate the influence of verb subcategorization preference (i.e. *verb bias*) on the processing of reduced complement sentences that varied according to whether the verb was a transitive preference verb or a complement clause preference verb. Examples of these sentences are shown below.

Transitive preference verb: He wrote [that] Sara fired her sister again.

Complement clause preference verb: He hoped [that] Sara fired her sister again (p. 567).

They hypothesized that if verb bias information were used during sentence processing, then reading times and eye fixation times should be shorter when the sentence was disambiguated toward the verb's preferred subcategorization frame.

For the eye tracking experiment, no significant effect of verb bias was found for either first pass or total reading times across either the ambiguous or the disambiguating regions. For the self-paced reading experiments, Ferreira and Henderson (1990) found no evidence for an influence of verb bias on first-pass reading times. For total reading time, they found a significant interaction between verb bias and complementizer presence, suggesting that verb subcategorization preference information does not influence the initial parsing of sentences, but does have an impact on a later reanalysis stage of processing.

Boland (1993; Experiment 1) also investigated the potential influence of verb subcategorization preference information on the processing of upcoming syntactic structures, and the time course of this influence. Sentence fragments were completed with verbs that were

biased either toward a simple transitive or sentential complement (S) continuation. Nominative or accusative pronouns were used to continue the fragment. The nominative pronoun suggested that an S would follow, whereas the accusative pronoun suggested a simple transitive continuation. Examples of these conditions are shown below.

Simple transitive: The waitress hit THEY/THEM.

Sentential complement: The liberal politician insisted THEY/THEM (p.138).

The participants completed either a cross-modal naming task or lexical decision task. In the naming task, the participants listened to the sentence fragment and then named the visually-presented target pronoun. For the lexical decision task, participants responded to a lexical decision target that was presented just before the last word in the sequence. Boland (1993) used these two tasks because of evidence from previous work that lexical decision times are influenced by both first pass and later integrative processes, whereas naming times are more reflective of initial processing (Shapiro, Zurif, & Grimshaw, 1989). Boland (1993) found that both naming and lexical decision times were significantly faster in the congruent than in the incongruent condition. She concluded that verb subcategorization preference information influences first pass sentence parsing.

In addition, the difference in lexical decision times between the congruent and incongruent conditions was greater in the accusative (simple transitive) condition than in the nominative (sentential complement) condition. Boland (1993) interpreted the difference in the effect of congruity across the two structures as being due to the greater structural complexity in the sentential complement condition. She also suggested that the absence of a complementizer *that* in the sentential complement condition might have seemed unusual to some participants, and therefore increased their reaction times.

The work of Trueswell et al. (1993; Experiment 1) was nearly identical in purpose and methods to Boland's study (1993; Experiment 1). Verbs with either an NP-bias or an S-bias were inserted into sentence fragments that began with an NP and ended either with a verb or with a complementizer (e.g., *the old man insisted* or *the old man insisted that*). Each fragment was then paired with either the nominative pronoun *he* or the accusative pronoun *him*.

The participants listened to sentence fragments and then named the visually presented target pronoun as quickly as possible. As in Boland's study (1993; Experiment 1), Trueswell et al. (1993) found a significant interaction between verb bias and pronoun case. However, the interaction was only significant for one of the verb types. Naming times for *him* were significantly faster following NP-bias verbs. The difference in naming times for *he* failed to reach statistical significance, although reaction times were numerically faster following S-biased verbs.

In Experiments 2 and 3, Trueswell et al. (1993) investigated further whether verb subcategorization information predicts upcoming structure by measuring reading times for sentential complements that either followed S-biased or NP-biased verbs using word-by-word self-paced reading (Experiment 2) and monitoring eye movements (Experiment 3). In Experiment 2, total reading times for sentential complements following S-biased verbs were significantly shorter than for those following NP-biased verbs, across the region of the first noun following the matrix verb and across the final region of the sentence. However, in Experiment 3, a significant difference in *first pass* reading times between the two conditions was only found in the post-disambiguating region of the sentence.

Jennings et al. (1997) investigated not only whether verb subcategorization influences sentence processing, but also whether this influence is graded according to the strength of the

preference. They used a word naming task that was virtually identical to the one used by Trueswell et al. (1993; Experiment 1) and Boland (1993). They found a significant interaction between verb bias and pronoun case, which occurred for both the NP-biased and S-biased verbs. They also examined the effect of the strength of verb bias on participants' reaction times. They found that, the stronger the preference for one subcategorization frame, the greater the advantage when the preferred structure was used.

In general, the results of these studies suggest that a verb's preference for a particular subcategorization does influence the ease with which upcoming syntactic structures are processed. However, the findings are less clear regarding whether this influence occurs during first-pass sentence parsing or during a later reanalysis stage of processing. There are several methodological issues that may account for this inconsistency. The first relates to the eye tracking experiments in Ferreira and Henderson (1990; Experiment 1) and Trueswell et al. (1993; Experiment 3). In Ferreira and Henderson (1990; Experiment 1), the ambiguous, disambiguating, and post-disambiguating regions each consisted of only a single word. Similarly, Trueswell et al. (1993; Experiment 3) examined differences in reading time for very small (one or two word) regions, rather than across the entire ambiguous or disambiguating regions. Because reading times for a single word are typically quite short, it seems plausible that the influences of verb bias that would have occurred in the disambiguating region spilled over into the postdisambiguating and final regions. Examining only these smaller sections may have masked real differences between verb bias conditions that would have occurred if the entire ambiguous or disambiguating regions were examined.

In addition, not all of the verbs selected by Ferreira and Henderson (1990), Trueswell et al. (1993) and Boland (1993; Experiment 1) actually subcategorized for both an NP and an S.

While including these verbs might allow the investigators to study the effects of verb subcategorization information on sentence processing, it does not allow for an accurate account of how and when listeners use *preference* information.

Finally, Trueswell et al. (1993; Experiment 1) and Boland (1993; Experiment 1) found a significant interaction between verb bias and pronoun case. However, for both experiments, reliable simple effects of verb bias were larger for accusative than nominative pronouns. In fact, Trueswell et al. (1993; Experiment 1) failed to find a significant effect for accusative pronouns. The investigators in both studies attributed this pattern of results to a reduced complement clause condition in which the complementizer 'that' was omitted from half of the complement clauses. The authors of both studies explained that omitting the complementizer in these cases might have elevated the processing difficulty for S-bias verbs.

Evidence consistent with this explanation came from a follow-up study in which Trueswell et al. (1993) found a significant correlation between the difficulty of the S-bias preference verbs in the reduced complement sentences and the percentage of complement clause completions using *that* in their verb preference pre-tests. That is, the more unusual it was for a particular verb to occur without a complementizer, the greater the processing difficulty for reduced complement clause structures. Trueswell et al. (1993) stated that this might have masked effects that were due to verb subcategory preference.

The results of these studies can be contrasted with those of Jennings et al. (1997), in which two of the three aforementioned shortcomings did not occur. First, all of the verbs used in this study did in fact subcategorize for both a NP and an S. Second, there was no reduced complement condition. It is likely that this is why Jennings et al. (1997) found an effect of verb bias on the first pass sentence processing of upcoming structures. In fact, Jennings et al. (1997)

showed evidence for a graded effect, such that a stronger bias was related to a stronger advantage for the preferred structure.

The studies reviewed in this section provide evidence that verb subcategorization information enables the prediction of upcoming structures. However, they do not address the issue of whether the expectation for upcoming verb subcategorization is influenced by other information within the sentence or in previously occurring sentences. Evidence of either type of influence would suggest that this prediction is not solely determined by the preceding verb, but is also affected by other sources of information.

1.2 OTHER SOURCES OF INFLUENCE ON PREDICTED VERB SUBCATEGORIZATIONS

As stated earlier, Gibson's DLT (1998, 2000) posits that sentence processing involves continually selecting among alternative syntactic structures, with the choice of a particular structure being based on the relative activation levels of the various alternative structures that are available at that point. The activation levels of these structures reflect the extent to which lexical, plausibility, and discourse-context information supports the choice of a particular structure. The original version of Gibson's DLT (1998, 2000) suggests that the choice among structural alternatives involves selecting the particular phrase structure rule that represents the way in which a particular sentence or constituent (e.g., VP, NP, PP) is likely to be completed at a particular point. For example, there are various alternatives for constructing a verb phrase (VP → V NP PP; VP → V NP NP, etc.). Additionally, the findings of Chen, Gibson and Wolf (2005)

and Warren and McConnell (2006) suggest that parsing might also involve selecting among alternative verb subcategorizations, and that this information is part of the verb's lexical entry.

While Gibson's DLT (1998, 2000) does not specify the precise types of information that influence the activation level of candidate structures, there is evidence that the relative ease with which alternative verb subcategorizations are processed can be influenced by other lexical and syntactic information presented within in the sentence. For example, Garnsey, Pearlmutter, Myers, and Lotocky (1997) found that first pass reading times for the disambiguating region following equi-biased verbs were faster for sentential complements when the NP was implausible. However, these results were not obtained for verbs that have an NP- or S-bias.

Less direct evidence comes from Roland, Elman, and Ferreira (2006; Experiment 1). They used corpus data to determine factors that were related to the likelihood that a noun phrase would follow a verb that could subcategorize for NP or S. They found a relationship between the proportion of NP or S subcategorizations and the following variables: NP plausibility, the frequency and animacy of the postverbal head noun, the presence of the determiner *the*, and to a lesser extent, the identity of the main subject. As in the study by Garnsey et al. (1997), the influence of these factors was greater for equi-biased than for strongly biased verbs. However, because Roland et al. (2006) examined factors related to sentence production, the results do not necessarily imply that all of these sources of information are used by listeners and readers to predict upcoming structures.

Evidence that intrasentential information can be used to influence the prediction of upcoming verb subcategorizations suggests that information from the preceding sentences could also have such an influence. Because predicting verb subcategorizations involves utilizing syntactic and lexical information, it seems likely that both structural information and the

discourse context (which represents the cumulative meaning of the preceding sentences) might influence this type of prediction.

The results of many of the studies that will be reviewed in this paper suggest that sentence processing is influenced by either the discourse context or the frequency of various syntactic structures within previously occurring sentences. However, none of the reviewed studies of either of these potential sources of influence controlled for the influence of the other. Because the expected influences on the processing of a particular structure are proposed to occur over a similar set of sentences, separately the influences of these two factors is important to determine whether each has an independent effect on current sentence processing.

A study to determine whether either or both of these two types of information exert an independent influence on listeners' and readers' predictions for verb subcategorizations would need to account for a number of variables, because there are at least two sources of evidence which suggest that verb subcategorization frequency is related to the nature of the prior discourse context. First, there is evidence that verb subcategorization frequencies vary between those obtained from isolated norming experiments and those found in spontaneous corpora. Second, there is reason to believe that verb subcategorization frequencies vary among corpora that vary in modality, purpose, and register. These sources of variability in verb subcategorization frequency will be discussed in the next section of this paper.

1.2.1 Differences in verb subcategorization frequency between experimental norming data and spontaneous corpora

Verb subcategorization preferences are typically determined through experimental norming studies, in which participants are required to either generate a sentence that includes a particular verb, or to complete a sentence fragment that ends with a verb. The subcategorizations that are generated for each verb during these tasks are tallied, and the subcategorization that co-occurs most frequently with a verb is considered to be the verb's *preferred* subcategorization. This is the subcategorization preference information that has been used in the vast majority of verb processing studies.

However, this method raises the question of how predominant a particular verb subcategorization must be in order to be considered the *preferred* structure. Because many of the verbs used in these experiments have more than two subcategorization possibilities, some investigators (e.g., Trueswell et al., 1993; Jennings et al., 1997) have not required that a preferred subcategorization follow the verb the majority of the time (> 50% of cases), and that it simply occur more often than other subcategorizations that are the investigated in that study. In some cases, a third subcategorization that is not the being investigated is actually the most prevalent one associated with a particular verb. This raises question of whether the most frequent subcategorization is actually preferred (or predicted), because it may not occur in the majority of cases. In addition, it is unclear whether the number of subcategorization options influences that degree to which the processing of a preferred option is facilitated.

In addition, the results of a number of other studies suggest that the verb subcategorization frequencies generated in these norming experiments are not always strongly

correlated with those derived from spontaneous corpora (Lapata, Keller, & Schulte im Walde, 2001; Merlo, 1994). For example, Merlo (1994) compared the subcategorization frequencies for a set of 105 verbs that were found in four experimental norming studies (Connine, Ferreira, Jones, Clifton, & Frazier, 1984; Garnsey, 1994; Holmes, Stowe, & Cupples, 1989; Trueswell, et al., 1993) to those obtained from spontaneous corpora (The Penn Treebank corpus, the Wall Street Journal, transcripts from radio broadcasts, and spontaneous sentences collected from the DARPA Air Travel Information System program). For some comparisons, he found that the verb subcategorization frequency information obtained in the four norming experiments was not strongly correlated with that obtained for the same verbs found in the spontaneous corpora. The correlations between each of the norming studies that he investigated and the corpus data are shown in Table 2.

Table 2: Comparison of verb subcategorization data obtained from three experimental norming studies to combined data from spontaneous corpora. The correlations indicate the proportion of verbs assigned a certain preference in one data set that received the same preference in the other data set.

Merlo et al. (1994)	Comparison made	Structure studied	r
	Trueswell et al. (1993) vs. corpora	NP	.739
	Holmes et al. (1989) vs. corpora	NP	.594
	Garnsey et al. (1994) vs. corpora	NP	.727
	Trueswell et al. (1993) vs. corpora	S	.444
	Holmes et al. (1989) vs. corpora	S	.667
	Garnsey et al. (1994) vs. corpora	S	.585

However, these results should be interpreted in light of the fact that Merlo (1994) did not separately consider the various homonym meanings for a particular verb when determining the relative frequency of their associated verb subcategorizations. According to Roland (2001), verb subcategorizations preferences differ significantly across different homonym meanings of a particular verb. Additionally, Hare, McRae, and Elman (2004) found that verb subcategorization

frequencies differ meaningfully even across different senses within a particular homonym meaning. Therefore, the results of Merlo (1994) may be at least partially due to the various homonym meanings and senses of a verb occurring with different frequencies in the norming experiments as in the corpus data.

Lapata, Keller, and Schulte im Walde (2001) compared the subcategorization preferences for sets of verbs found in the British National Corpus to those obtained for the same verbs in four different experimental studies through sentence completion and sentence elicitation (Connine et al., 1984; Garnsey et al., 1997; Pickering, Traxler, & Crocker, 2000; Trueswell, et al., 1993). Lapata and colleagues examined the subcategorization frequencies for one set of verbs that subcategorize for either an NP or an S, and for another set of verbs that subcategorize for either an NP or a null complement. For both sets of verbs, the verb subcategorization frequencies in the two types of datasets were not consistently well-correlated. The correlations found for both sets of verbs are found in Table 3 below.

Table 3: Comparison of the subcategorization frequencies obtained from the British National Corpus (BNC) to four different norming studies. The results for Pickering et al.'s (2000) sentence completion and sentence generation studies are reported separately.

Lapata et al. (2001)	Comparison made	Structure studied	r
	Garnsey et al. (1997) vs. BNC	NP/S ambiguity	.81
	Trueswell et al. (1993) vs. BNC	NP/S ambiguity	.69
	Connine et al. (1984) vs. BNC	NP/S ambiguity	.74
	Connine et al. (1984) vs. BNC	NP/Ø ambiguity	.61
	Pickering et al. production (2000) vs. BNC	NP/Ø ambiguity	.66
	Pickering et al. completion (2000) vs. BNC	NP/Ø ambiguity	.42

While not all of the correlations found in these studies are particularly low, they should be interpreted in light of the fact that, for any given verb, there are only a few subcategorizations. This is particular true for the Lapata et al. (2001) study in which verbs with two possible subcategorizations were studied.

One potential shortcoming of Lapata et al. (2001) is that the correlations for NP and S, as well as for NP and Ø were considered together. According to Roland et al. (2006), the frequency of NP and S continuations differs significantly. In a large corpus, they found that, of verbs that subcategorized for NP or S and which did not resolve to Ø, 72.8% resolved to a NP continuation. Therefore, a more accurate analysis of the data would have involved determining correlations separately for each of the continuation types, and considering each in light of its relative frequency within the language.

1.2.2 Differences in verb subcategorization among spontaneous corpora

There is also reason to believe that verb subcategorization frequencies differ between natural corpora of different modalities (written versus spoken), registers (formal versus informal), and purposes. Unfortunately, direct evidence of variability in verb subcategorization frequency across spontaneous corpora is sparse. Various indirect sources of evidence will also be considered.

Before discussing differences among spontaneous corpora, it is important to mention that there is some evidence that verb subcategorization frequencies tend to be stable across corpora that are similar in terms of formality and modality. Roland and Jurafsky (2002) examined subcategorization frequencies for six verbs (for which variation in verb sense had been

controlled) across two written corpora that were of similar register: The Brown Corpus and the Wall Street Journal Corpus. They found no significant differences in subcategorization frequencies across the two corpora. However, it is unknown whether the same results would have been obtained with a larger number of verbs.

Roland (2001) provides the strongest evidence for variability in verb subcategorization frequency across spontaneous corpora of different modalities, although that evidence is indirect. He investigated the extent to which modality influences transitivity. He examined the transitivity preferences of sixty-four single-sense verbs from two written corpora (the Wall Street Journal Corpus and the Brown Corpus) and one spoken corpus (the British National Corpus). Nine of the sixty-four verbs differed meaningfully in terms of their transitivity across the data sets. While these results for verb transitivity do not provide direct evidence of the variability (or lack of variability) in verb subcategorization frequencies across discourse contexts, they do provide the strongest evidence that is currently available.

In addition, Roland (2001) might have introduced an unwanted source of variability by selecting corpora from two different English dialects. The British National Corpus is comprised of spoken British English, whereas the Wall Street Journal Corpus and the Brown Corpus are both written American English corpora. The extent to which this might influence subcategorization frequencies is unknown, but a comparison across corpora of more similar dialects would have provided more convincing results. The same criticism applies to Lapata et al. (2001) who compared the results of experimental norming studies that were conducted in the United States to data obtained from the British National Corpus.

1.2.3 Differences among corpora that could influence subcategorization frequencies

While there is little direct evidence that verb subcategorization varies significantly among spontaneous corpora, several differences among corpora of various registers and modalities suggest that verb subcategorization frequency might vary by register and modality. Some of these are the same factors that Roland and Jurafsky (2002) suggested were related to differences in subcategorization preference between experimental norming studies and spontaneous corpora: zero anaphora, default reference, and variation in verb sense. Biber (1998) suggested three other differences among spontaneous corpora that might be related to differences in verb subcategorization frequencies: the use of anaphoric and exophoric reference, the frequency with which referents refer to new (versus given) information, and the distance between a pronoun (or a repetition of a noun) and its antecedent.

Biber (1998) compared a variety of discourse characteristics that have been claimed to influence verb subcategorization frequency across several different types of discourse contexts. He examined two different registers of spoken corpora (conversation and public speeches) from the London-Lund Corpus and two written registers (news reportage and academic prose) from the LOB corpus. He examined each referring expression in terms of whether it referred to new or given information. For given information, he examined the types of referents that were used, determining whether they could be classified as anaphoric or exophoric expression. *Anaphoric referents* are those whose identity is known from preceding referents in the discourse. *Exophoric referents* are those whose identity is known from the physical situation. For anaphoric referents, he examined the distance between anaphoric referents and their antecedents.

Across the four registers, Biber (1998) found a significant difference in the frequency of referents that referred to new (versus given) information. Academic prose showed the greatest proportion of new referents (65%), while conversation showed the least (30%). There were also substantial differences across the four registers in the types of referring expressions that were used to refer to given information. Exophoric reference was most common in conversation, while it was rarely used in either of the written registers. Anaphoric reference was used much more frequently in the written corpora than in the spoken corpora.

Biber (1998) also examined the average distance (in number of intervening referring expression) between the use of a full-noun referent and the pronoun and repeated full-nouns referring back to them. For all registers, the distance between a pronoun and its antecedent was shorter than the distance between a noun and a full repetition of the noun. While the average distance between a pronoun and its antecedent was similar across the four registers, there was greater variability for the distance between a noun and its full-noun repetition across registers. The relevant data are provided in Table 4.

Table 4: A comparison of relevant attributes of discourse context, by Biber (1998).

	Average distance between full-noun repetitions	Average distance between pronouns and their antecedents.
Conversation	9.0	3.0
Public Speeches	10.0	3.5
News Reportage	13.5	3.0
Academic Prose	10.0	2.5

In a separate study that was reported in the same chapter, Biber (1998) examined the variability in the tense and voice of verbs across the various sections of research articles (Introduction, Methods, Results, Discussion). Because these sections vary in their purpose,

Biber (1998) examined these sections in order to determine how the purpose of a text is related to the tense and voice of the verbs within it. One important finding was that the proportion of agentless passives varied significantly across the sections (p <.001). Because verbs in passive sentences generally take one fewer internal argument than their active counterparts, this could suggest variability in verb subcategorization frequency as well.

Roland and Jurafsky (2002) compared the percentages of passive sentences, verbs without subcategorization, and sentences that were direct quotations across two written corpora (the Brown Corpus and the Wall Street Journal Corpus) and one spoken corpus (the Switchboard Corpus), and found substantial differences among the corpora on all three measures. Also, as Table 5 shows, each of the measures varied at least as much between the written and spoken corpora as between the two written corpora.

Table 5: Comparison of percentage of passive sentences, zero subcategorization, and direct quotation across three corpora, from Roland and Jurafsky (2002).

Data Source	% Passive	% 0 Subcategorization	% Direct Quotation
Switchboard	2.2	18	0
Wall Street Journal	6.7	13	4
Brown	7.8	8	6

While there is no known direct evidence that verb subcategorization frequency varies across discourse contexts, the previously described research studies provide evidence that this is likely the case. Roland (2001) found variability in verb transitivity for nine of the sixty-four verbs that he investigated across corpora. In addition, the results of both Biber (1998) and Roland and Jurafsky (2002) indicate that corpora of different modalities vary according to characteristics that are thought to influence verb subcategorization. Most relevant to the issue of variability in verb subcategorization is Roland and Jurafsky's (2002) finding that the zero subcategorization appears to vary in a meaningful way among corpora.

However, any data gathered from spoken language cannot give us direct evidence about how listeners may or may not adjust their expectations based on this information. In addition, while the corpora come from various spoken and written sources, they have been gathered from a relatively narrow range of communicative contexts. The extent to which the language in these corpora is representative of the language that an individual typically encounters on an everyday basis, or in their cumulative experience, is unclear.

1.3 THEORIES AND MODELS OF DISCOURSE CONTEXT

It is widely believed that a coherent discourse context facilitates the comprehension of each successive clause or sentence. A number of theories and models have been developed that explain how discourse-level comprehension influences local sentence processing (Graesser, Mills, & Zwaan, 1997; Graesser, Olde, & Klettke, 2002; Kintsch, 1988, 1998; Trabasso & Bartolone, 2003; Trabasso & Wiley, 2005). While most current models and theories would predict that the presence of a coherent preceding discourse context would lead to faster processing of ensuing sentences than would a preceding series of unrelated sentences, they differ regarding the precise mechanisms by which this facilitation is proposed to occur.

However, before discussing the differences among various theoretical positions, it is important first to discuss the major commonalities among them. First, each of the theories to be described posits that discourse comprehension involves integrating information at various levels of representation. One level of representation that is common among various theoretical positions is the *situation model* which is the "mental microworld of what the story is about,"

(Graesser et al., 2002, p. 231). Second, most of the theories and models to be described hold that local sentence processing is facilitated when it is related in some way to the prior discourse context.

Constructionist Theory (Grasser et al., 2002; Graesser et al., 1997; Graesser, Singer, & Trabasso, 1994) describes the process of building situation models when a narrative text is comprehended. It holds that a wide range of inferences are generated during discourse comprehension, and that these inferences are based on the context of the discourse, general world knowledge, and the pragmatic context (the reason why the narrative is being told). The range of inferences is then narrowed, based on two assumptions: the coherence assumption and the explanation assumption. According to the *coherence assumption*, the comprehender selects certain inferences, in an attempt to build a situation model that both establishes a link between a current statement and a recent proposition (local coherence), and between that statement and information several sentences or pages earlier in the discourse (global coherence).

Evidence for the impact of local coherence on sentence processing comes from Zwaan, Magliano, and Graesser (1995). They found a significant increase in reading times for sentences in a narrative, if the content of the sentence involved a break in coherence on any one of five dimensions: 1) protagonist (the protagonist in the current sentence was not one of the protagonists in working memory); 2) temporality (the event described in the sentence involved a break in the timeline of the narrative); 3) causality (the event described in the sentence was not causally related to the content in working memory); 4) motivation (the action described in the sentence is not part of the character's plan in working memory), and 5) spatiality (the event described in the sentence was in a different location from the content in working memory). In addition, Zwaan et al. (1995) found additional increases in participants' reading times when

continuity breaks occurred on more than one dimension. This would suggest that reading times for sentences following a series of scrambled sentences, across which there would be numerous breaks in continuity, would be greater than for those for the same sentences following coherent discourse contexts.

Additional evidence for the role of local coherence on sentence processing comes from McKoon and Ratcliff (1992; Experiment 2). The investigators varied both local and global consistency across four versions of a story. Participants read two of the four versions of each story. A single target word was then presented, and the participants had to decide whether the word had appeared in the story. Local, but not global, consistency was found to be significantly related to participants' response times in identifying target words as being related to the story.

Graesser and colleagues (2002) state that a second assumption, *the explanation assumption*, also reduces the range of inferences available to a comprehender. According to this assumption, comprehension is also driven by "why" questions. That is, readers and listeners attempt to explain why certain information is presented, or why certain actions, events or states occur. These authors state that inferences are more likely to be made if they are related to more recent events on the causal chain, or if they are related to multiple events within the narrative.

Alternatively, Trabasso and Wiley (2005) and Trabasso and Bartolone (2003) propose that narratives are composed of a series of episodic units, called *goal-attempt-outcome* episodes, which consist of a character's goal, his attempt at fulfilling his goal, and the outcome of this attempt. These episodic units are interconnected in a hierarchal structure, in which a character's failure to fulfill a goal at any level in the hierarchy results in either the generation of subordinate goals, or in goal abandonment. If subordinate goals are generated, then a successful outcome of subordinate goal allows the character to attempt the goal one level above the previous goal in the

hierarchy. On the other hand, if the character is successful in fulfilling his initial goal, this success can result in the formation of superordinate goals.

Trabasso and Wiley (2005) and Trabasso and Bartolone (2003) explain that comprehenders form a coherent representation of the narrative as they track characters' goals and subgoals, their success and failure in fulfilling these goals, and their responses to their success or failure (including the formation of subordinate goals and alternative plans) over the course of the narrative. Evidence for this position comes from Trabasso and Nickels (1992). They found that, when individuals between age five and adulthood retold narratives, goal-attempt-outcome units formed the bases of the retellings.

In addition, Trabasso and Wiley (2005) propose that narratives can also be analyzed in terms of their causal networks, by examining each clause and evaluating its causal relationship to other clauses in the network. Like Graesser and colleagues, Trabasso and Wiley (2005) also hold that comprehenders achieve a coherent representation of the discourse context by tracking causal relationships among events in the narrative.

In a related study of discourse comprehension, Langston and Trabasso (1999) used a connectionist model to mimic this process. The model has a long-term storage component, called a *text representation* that contains nodes and connections among nodes. Each node corresponds to a clause in the discourse, and has an activation level that fluctuates over time, as the text representation is constructed. In addition, each connection between nodes has a connection strength. These connection strengths change as the model incorporates each new node and its connections are integrated into the text representation. At each point in processing, the activation values of the nodes and their connection strengths reflect the reader's history of comprehension at that point.

Langston and Trabasso (1999) introduced text into the model one node at a time to simulate the process of reading comprehension. They measured the activation values of the nodes and the connection strengths among them during and after the incorporation of each new node. They compared the results obtained from this model to empirical data from research participants, and found that the most reliable predictors of the accessibility of a particular node (as determined by participants' response times) was the connection strength of that node to other nodes in the network, and the number of other nodes to which that node was connected. Although the extent to which this model mimics the reading comprehension in humans in unclear, it may suggest that discourse cohesion plays a role in the ease with which incoming information is processed.

Kintsch's Construction-Integration (C-I) model (1988; 1998) is similar to Constructionist theory in that it involves the activation of nodes representing propositions and connection weights among them. However, in his C-I model, Kintsch (1988; 1998) describes a two-stage process that is somewhat different from the one described by Trabasso and Wiley (2005) and Langston and Trabasso (1999). The first stage, Construction, is a strictly bottom-up process in which a very broad range of propositions, both correct and incorrect are generated on the basis of the prior discourse context and general world knowledge. Both general knowledge and the prior discourse context are represented in the same associative knowledge net, and generating propositions involves activating their associated nodes on this knowledge net. One source of evidence for bottom-up contributions to discourse comprehension comes from Just and Carpenter (1980), who found that even proficient readers densely sample words in a text, as indicated by their eye fixations.

According to Kintsch (1988; 1998), Construction is a multi-stage process that first involves forming the concepts and propositions that are directly related to the linguistic input. Then, each of these concepts and propositions are elaborated by selecting a few of its most closely associated neighbors from the knowledge net. Next, additional inferences, needed to establish coherence, are generated. Finally, connection strengths are assigned to all of the pairs of elements that have been created.

In the second stage, Integration, general knowledge and prior discourse context again interact to narrow the range of inferences that are considered by the comprehender. This process thereby eliminates unwanted elements from the discourse representation. Comprehension is assumed to be organized into construction-integration cycles, with each cycle roughly corresponding to a single phrase or short sentence. The end result of each cycle is a new pattern of activation with the highly activated nodes representing the discourse representation that is formed on each processing cycle (Kintsch 1988, 1998).

Kintsch (1988) provides the following example, from Frazier and Rayner (1982), to demonstrate the CI model's explanation of how propositions representing various subcategorizations for a given verb are activated, and how the appropriate one is selected. This example is especially pertinent to the current study, because it involves the type of ambiguity that is the subject of the current study.

Example: The linguists knew the solution to the problem would not be easy (p. 135).

According to Kintsch (1988), both the NP and S interpretation of "to know" are formed during the construction process. However, the NP interpretation receives greater activation, because it involves minimal attachment. When reading further into the sentence reveals that the NP interpretation is implausible, this proposition simply disappears during the Integration phase.

Comprehenders do not need to actively reinterpret the sentence and repair the error. Because both propositions were generated during the Construction phase, the integration of the S (rather than the NP) interpretation would not increase the processing time for that sentence.

According to the C-I model (Kintsch 1988; 1998), regardless of whether the example sentence was preceded by a coherent discourse context or by a series of unrelated sentences, both propositions would have been constructed. However, the activation value of the incorrect interpretation would be reduced to zero, and this would cause it to be removed as a possible interpretation during the integration phase. Therefore, the presence of a coherent representation should have no impact of on the ease of sentence processing.

The three approaches described in this section each offer a unique viewpoint regarding the process of discourse comprehension, and the manner in which a coherent discourse representation influences current sentence processing. While both the Constructionist Theory and the works of Trabasso and Wiley (2005) and Langston and Trabasso (1999) would suggest that a coherent prior discourse context that is related to the current sentence would facilitate current sentence processing, Kintsch's C-I model (1988; 1998) would hold that this type of facilitation would not occur.

1.3.1 The effect of discourse context on sentence parsing

If verb subcategorization frequency varies across discourse contexts, then it would be advantageous for listeners and readers to use discourse context information to influence their predictions for verb subcategorizations. It seems likely that individuals use this type of information in making these predictions, because there is evidence that they use discourse

context information in making other types of parsing decisions. Evidence from a number of research studies suggests that discourse context is an important source of information that is used by comprehenders in resolving various types of temporary ambiguity. In addition, there is evidence that discourse context influences the prediction of upcoming words. Both of these issues are related to the possibility that discourse context also influences the prediction of upcoming verb subcategorizations. Predicting subcategorizations of a verb involves more than predicting upcoming words. Because all of the possibilities for upcoming subcategorizations are made available simultaneously at the point of processing the verb, Trueswell et al. (1993) and other researchers have referred to a verb's potential to take two or more different subcategorizations as a type of local structural ambiguity.

1.3.2. The effect of discourse context on resolving structural ambiguities

Evidence from three studies (Altmann & Steedman, 1988; Britt, Perfetti, Garrod, & Rayner, 1992; Spivey & Tanenhaus, 1998) suggests that prior discourse context can influence the resolution of temporary structural ambiguities. These studies provide evidence that the processing of a particular syntactic structure is facilitated when the prior discourse context supports that particular reading of the sentence. Table 6 (Appendix A) provides a summary of the results of these studies.

Altmann and Steedman (1988) used a self-paced reading procedure to investigate the influence of discourse context on the processing of prepositional phrases that were structurally ambiguous between an NP and a verb phrase (VP) attachment. These prepositional phrases

occurred in sentences that were either strongly NP-biased or VP-biased, based on their lexical content. Examples of these sentence types are shown below.

NP attached: The burglar blew open the safe with the new lock and made off with the loot.

VP attached: The burglar blew open the safe with the dynamite and made off with the loot (p. 224).

Sentences of these two types were preceded by contexts, which supported one of two structurally appropriate readings of the sentence. Reading times for sentences in which the context supported the attachment bias of the sentence were significantly shorter than those for sentences in non-supporting contexts. These results provide evidence for the influence of discourse context, because a non-interactive account would have predicted shorter reading times for the minimally attached VP-attached condition following both supportive and non-supportive contexts.

In a second experiment, Altmann and Steedman (1988) sought to determine whether the influence of discourse context occurred during first pass sentence parsing, or as part of a later reanalysis process. They found that the difference in reading times occurred at the prepositional phrase where the ambiguity was resolved (disambiguating region), thus supporting the interactive viewpoint.

However, the results of Altmann and Steedman (1988) would have been more convincing if a neutral condition has been incorporated into their experimental design. In their study, reading times in supporting contexts were compared to those when the context supported the other interpretation of the sentence. Therefore, the results may overestimate the extent to which the discourse context influenced sentence processing.

In their first two experiments, Britt, Perfetti, Garrod, and Rayner (1992) used a self-paced moving window procedure (Experiment 1) and eye tracking methods (Experiment 2) to determine whether discourse context could override syntactic parsing preferences. They compared the same two types of sentences as Altmann and Steedman (1988). Each of these sentence types was preceded by both neutral and biasing passages. The biasing contexts were designed to bias the participants toward one syntactic attachment over the other. The results of Experiment 1 indicated that, in the neutral context, the participants read the disambiguating region significantly more slowly in the low NP attached condition than in the high VP attached condition. This difference was eliminated in the biasing context condition. In Experiment 2, the pattern of results for the first pass reading times was similar to the reading time results from the first experiment. This was taken as evidence that discourse context has an immediate effect on resolving syntactic ambiguities, rather than affecting only a later reanalysis stage of processing.

The results of the first two experiments indicate that discourse context can influence local attachment decisions. In Experiment 3, Britt et al. (1992) used a self-paced moving window procedure to determine whether context could also influence attachment decisions across a major constituent boundary. Sentence with high VP attachments or with reduced relative clauses were used as stimuli. Examples of these sentence types are shown below.

High VP attachment: The woman rushed to the hospital without taking her laundry.

Reduced relative condition: The woman rushed to the hospital had given birth safely

(p. 306).

In addition, the sentences with high VP and low NP attached prepositional phrases from Experiment 1 were also included as a separate condition in this experiment, in order to provide a direct comparison between the effects of discourse context on local and non-local attachments

decisions. Each of the sentences in the four conditions was presented in isolation and in a biasing context.

The results for the low NP attached/high VP attached comparison replicated those from Experiment 1. However, for the reduced relative clause/high VP comparison, reading times in the disambiguating region were significantly longer for the reduced relative conditions both in isolation and following the biasing context. These results would have been expected based on a non-interactive viewpoint, because the VP-attachment involves a minimal attachment. This suggests that discourse context effects did not override syntactic parsing preferences in the case of reduced relative/high VP attachment ambiguity.

Britt et al. (1992) suggest that local attachment decisions are affected by discourse context, but that ambiguities that involve attachments across a major constituent boundary are not. However, it is also possible that the longer reading times associated with the relative clauses may have been related to sentence processing difficulty, rather than the influence of discourse context. Sentences with relative clauses of the form used in the study (without *who was* or *that was*) are infrequent, and their use may have been unexpected by the participants, thereby increasing reading times.

Spivey and Tanenhaus (1998) examined the role of discourse context on the processing of sentences that were temporarily ambiguous between a reduced relative clause reading and main clause reading, and which resolved to a reduced relative clause reading. These temporarily ambiguous sentences were created by using verbs which were morphologically ambiguous between simple past tense and the passive participle. In Experiment 1, an additional set of unambiguous sentences was created in which the verb form was morphologically unambiguous. All of the verbs used in the sentences had strong transitive preferences, so that the isolated

sentences would have a main clause bias. Examples of these two sentence types are shown below.

Ambiguous reduced relative:

The actress selected by the director believed her performance was perfect.

Unambiguous reduced relative:

The actress chosen by the director believed her performance was perfect (p. 1523).

Each of these sentence types followed one of two types of context. In the two-referent context condition, two possible referents for a given NP (e.g., two actresses) are introduced, thus providing for the presupposition requirements of the reduced relative reading. In the one-referent context, the NP is associated with a unique referent in the discourse, and therefore supports a main clause (nonmodification) reading.

Spivey and Tanenhaus (1998; Experiment 1) recorded participants' eye movements as they read across four regions of the sentences: initial NP, verb region (ambiguous region), *by* phrase, and main verb plus one word. The investigators considered the *by* phrase region to be an intermediate region, in which there was strong probabilistic disambiguation, but which was still syntactically ambiguous. The main verb region was considered to be the disambiguating region.

In the one-referent condition, total reading times at the verb were greater in the ambiguous sentence than in the unambiguous sentences. This difference was reduced in the two-referent condition, although the authors do not comment on whether the difference continued to be statistically significant. For first pass reading times, no significant effect of discourse context was found for the region of the verb. However, for the *by* phrase, in the one-referent condition, there was a significant difference between conditions that was not significant in the two-referent condition.

Spivey and Tanenhaus (1998; Experiment 2) was similar to Experiment 1. However, the eye fixation times for the syntactically ambiguous sentences were compared to those for an unreduced relative clause. The authors found that reading times for the reduced relative sentences were significantly longer when they were preceded by one-referent contexts than in any of the other three conditions. For first-pass reading times, there was a significant interaction of context and relative clause reduction for both the verb region and the by phrase region. The two-referent context facilitated the processing of the reduced, but not the unreduced relative clauses. Taken together, the results of Spivey and Tanenhaus (1998) suggest that discourse context influences syntactic ambiguity resolution during first pass sentence parsing.

Altmann and Steedman (1988), Britt et al. (1992) and Spivey and Tanenhaus (1998) provide consistent evidence that discourse context can influence local attachment. Britt et al. (1992; Experiment 2) and Spivey and Tanenhaus (1998) provide strong evidence that the effect of discourse context on resolving structural ambiguities occurs during first pass parsing, rather than as part of a later reanalysis process. While the results of Britt et al. (1992; Experiment 3) suggest that discourse context does not significantly influence attachments that involve crossing a major constituent boundary, Spivey and Tanenhaus (1998; Experiment 1) found a significant influence of discourse context on participants' reading times for the same type of temporarily ambiguous structure.

Because verb subcategorizations constitute a type of local ambiguity, the results of each of these studies suggest that discourse context might influence the resolution of this type of ambiguity as well. However, these studies investigated only a few types of local ambiguity resolution, and only one type of attachment decision involving a major constituent boundary.

Therefore, the extent to which these results generalize to other types of temporary structural ambiguities is unclear.

1.3.2 The effect of discourse context on the meaning of homonyms

Vu et al. (2000) and Hare, McRae, and Elman (2003) provide evidence that discourse context can also influence the resolution of the temporary ambiguity that can occur when a homonym is encountered in a sentence. The results of these studies, which suggest that the activation of various meanings of a homonym is influenced by prior discourse context, are summarized in Table 7 (Appendix A).

Vu et al. (2000; Experiment 1) investigated the influence of global (discourse) context on the resolution of ambiguous words. They constructed sentences that ended with a homonym. For each sentence, two passages were constructed to bias the meaning of the homonym toward either its subordinate or dominant meaning. After reading each passage, the participants named a target word that was related to one of the homonym's meanings, and their naming times were recorded. These were compared to a baseline condition, in which passages were randomly matched with targets. Participants responded more quickly to targets that were related to the homonym meaning that was supported by the passage (dominant or subordinate) than to unrelated targets. However, response times to the unsupported homonym meaning were not facilitated.

In Experiment 2, Vu et al. (2000) sought to determine whether discourse context influences the initial activation of homonyms or only influences selection during later stages of processing. The protocol for the second experiment was identical to that for Experiment 1, except that the ambiguous word was presented for only 80 ms, to limit the opportunity for later

integrative processes to occur. The results of this experiment were identical to those from Experiment 1, suggesting that discourse context influences the initial processing of homonyms.

Vu et al. (2000; Experiment 3) investigated whether homonym meaning activation was due to an influence of discourse context that was distinct from sentence-level effects. They separated the first sentences of all of the passages and paired them with "null" sentences. Otherwise, the protocol was the same as Experiment 1. Vu et al. (2000) found no significant main effects or interactions, and concluded that the homonym activation effect that was seen in the first two experiments was due to discourse-level rather than sentence-level processes. However, there were several null sentences between the initial sentence and the target word. This separation may have lessened any sentence-level effects, and contributed to the nonsignificant results that were seen in this experiment.

Hare et al. (2003) investigated whether discourse context influences the activation of a particular verb homonym meaning, and whether the activation of this homonym meaning influences which subcategorizations are expected to follow the verb. They studied homonyms that had two meanings: one with an S preference and one with a direct object (DO) preference. These verbs were placed in the final sentences in two different types of contexts: one that was intended to bias the reader toward an S interpretation and one to bias the reader toward a DO continuation. The target sentences were temporarily ambiguous between an S and a DO reading, but all resolved toward the S reading. In addition, they created an unambiguous condition by inserting the complementizer *that* before the S, thereby eliminating the ambiguity.

Hare et al. (2003) measured participants' reading times across the region of the verb, the disambiguating region, and the post-disambiguating region. Across the verb region, reading times were faster for verbs in the DO-biasing context. Across the disambiguating region, there

was a significant interaction between context and ambiguity. For the ambiguous, but not the unambiguous condition, reading times were slower in the disambiguating region following the DO-biasing context than after the S-biasing context. These results suggest that the appropriate homonym of the verb was influenced by context, which in turn adjusted the expectation for its preferred subcategorization frame. Following the DO-biasing context, the S was unexpected, and this adjustment in expectations resulted in an increase in reading time.

However, reading times for the ambiguous sentences following their biasing context were compared to those following a context that biased the reader toward the other verb meaning. Therefore, the difference between the conditions may be the result of the combined influence of facilitating influence of the biasing context and the inhibitory influence of the alternate context. The inclusion of a neutral context would have made the results of this study more convincing.

The combined results of Vu et al. (2000) and Hare et al. (2003) suggest that discourse context can influence the resolution of the type of ambiguity that occurs when a homonym is encountered in a sentence. Vu et al. (2000) also attempted to separate these discourse-level effects from sentence-level effects, but their results were confounded by the fact that their biasing sentences were several sentences away from the target word.

The findings of Hare et al. (2003) suggest that, as discourse context influences the activation of certain verb meanings, readers' expectations for upcoming verb subcategorizations are also adjusted. While this is the only known study that has investigated this particular phenomenon, there is additional evidence that discourse context can influence a listener's predictions about upcoming words in discourse. This evidence is presented in the following section.

1.3.3 The influence of discourse context on the prediction of upcoming words

In addition to evidence that suggests that discourse context can influence the resolution of at least two types of temporary ambiguity, there is also evidence that discourse context can influence the prediction for upcoming words in a sentence. The results of four studies suggest that the processing of words that are related to the prior discourse context is facilitated (Cook & Myers, 2004; Hess, et al., 1995; Schwanenflugel & White, 1991; Sharkey & Sharkey, 1992). These results are summarized in Table 8 (Appendix A).

Schwanenflugel and White (1991) investigated the interaction of local (sentence-level) and paragraph-level (discourse context) information in facilitating the processing of upcoming words in text. In their first two experiments, they examined the effects of discourse context on words that were either highly constrained or not constrained by the local context (i.e., the word was either expected or unexpected). Examples of these sentence types are shown below.

Locally expected: The hikers slowly climbed up the *mountain*.

Locally unexpected: The hikers slowly climbed up the *stairs* (p. 162).

The investigators placed sentences of each of these types in discourse contexts that were either consistent or inconsistent with the local context. They accomplished this by making the first sentence (Experiment 1) or the fourth sentence (Experiment 2) of each paragraph either consistent or inconsistent with the local context. The other sentences in the paragraph did not bias the reader toward any expectation for the final word. After reading each paragraph, participants made lexical decisions regarding the final word in the paragraph.

Schwanenflugel and White (1991) found an identical pattern of results for both experiments. There was no significant main effect for discourse-level consistency. However,

locally expected words yielded significantly faster lexical decision times than locally unexpected words. A significant interaction between consistency and local expectancy indicated that discourse-level information had a significant effect on word processing when the word was locally expected, but not when it was locally unexpected. This suggests that discourse-level information did influence the processing of target words in high-constraint sentences. The similarity of results across the two experiments was taken to indicate that the lack of a significant main effect of discourse context was not related to the distance between the discourse-level information and the final word.

Schwanenflugel and White (1991; Experiment 3) investigated whether the significant effect of discourse-level information that was observed for the locally-expected condition was due to facilitative processes occurring prior to recognition of the word or to later integrative processes. For this purpose, they replaced the lexical decision task with a word naming task. A significant effect of discourse-level consistency was found in both the locally expected and unexpected conditions, suggesting that discourse context plays a role in word processing during the initial stages of sentence processing.

Sharkey and Sharkey (1992; Experiment 2) investigated the influence of discourse context on word processing, and aimed to separate the effects of discourse context from any influence that might be attributable to word priming. Lexical decision targets were placed in discourse contexts that were either related or unrelated to the target word. They attempted to reduce the potential effect of word priming by placing at least two words between any related word and any of the lexical decision targets. Target words in related text contexts yielded significantly faster lexical decision times than those in unrelated text contexts, even when the efforts were made to reduce the effect of word priming. However, as will be explained later in

this paper, the two-word separation between the lexical decision target and a related word may not have eliminated the effect of word priming.

Sharkey and Sharkey (1992) conducted two additional experiments in order to study the perseverance of word priming. In Experiment 3, participants read sentences, during which they responded to a lexical decision target that was placed either immediately after the prime (0-lag) or following two or four intervening words (2-lag and 4-lag). Lexical decision times were compared to a condition in which the target was not related to any word in the sentence. They found no effect of context (related vs. unrelated) or delay, and no interaction. There was no evidence of priming at any of the delay conditions.

In Experiment 4, Sharkey and Sharkey (1992) explored the possibility that the unexpected results in Experiment 3 were due to "sentence processing factors" (p. 562). The protocol was similar to Experiment 3, except that word lists were used that were scrambled versions of the sentences that were used in Experiment 3. Planned comparisons revealed a significant difference in lexical decision times between the 0-lag condition and the other two conditions. The authors interpreted these results as evidence that word priming had no influence on the results of Experiment 2.

However, the results of Experiment 3 appear to be unusual in that word priming, which is a well-studied phenomenon, was not found even when the target was placed immediately after the prime. Because the authors report no effort to ensure that none of the words in the unrelated condition were actually unrelated to the target, it is possible that the lack of a significant difference between the two context conditions was due to word priming occurring in both conditions. The results of Experiment 4, while providing some indication that word priming may decrease across intervening words, are limited in the extent to which they can be used to interpret

the results of Experiment 2, because the stimuli used in Experiment 2 did not included scrambled sentences. Finally, Balogh, Zurif, Prather, Swinney and Finkel (1998) found that, during the syntactic process of gap filling, an antecedent could be reactivated when the gap position followed the antecedent by as many as seven syllables. This suggests that the two-word separation that was used in Experiment 2 was not sufficient to control for semantic priming.

Hess et al. (1995; Experiment 2) examined the effects of global (discourse-level) and local (sentence-level) context on word processing. They created four conditions in which either the global and/or local context was related to the target word. Both global and local related contexts were associated with shorter naming latencies. However, the investigators were concerned that this effect might have been related to an inhibitory influence in the condition in which neither the global nor the local context was related to the target word.

In Experiment 3, Hess et al. (1995) addressed this concern by replacing the local unrelated/global unrelated condition with a neutral context. They found a significant effect of global, but not local context. The remainder of the experiments yielded similar findings, in spite of Hess et al.'s (1995) attempts to control for the following factors that might have reduced the likelihood of finding a significant main effect of local context: low power (Experiment 4), the participants' possible sensitivity to experimental manipulations (Experiment 5), the redundancy of the neutral condition (Experiment 6), the grammatical awkwardness of some of the items (Experiment 7) and differences among the items (in terms of frequency, regularity, length, and semantic content) across quartiles of the experiment (Experiments 8 & 9). They concluded that global (but not local) context has an impact on the ease of processing of upcoming words.

Cook and Myers (2004; Experiment 1) investigated the influence of discourse context on the processing of individual words, and the time course in which incoming discourse is linked to

general world knowledge. Participants' eye fixations were recorded as they read paragraphs that included two encounters with an individual playing the role of "the performer." At each encounter, the performer role was filled by someone who was either appropriate or inappropriate for the scripted role. In addition, during the second encounter, the role was fulfilled either by the same performer as in the previous encounter, or by a different performer. Four conditions, in which each of the first and second encounters were either filled appropriately or inappropriately, were intended to measure two variables: consistency with world knowledge and consistency with discourse context.

For both the first and second encounters, first pass reading times were shorter when the target was appropriate than when it was inappropriate. For the second encounter, first pass reading times were shorter when the performer was the same as in the first encounter. In the post-target region following the second encounter, there was an interaction between consistency with world knowledge and consistency with context. First pass reading was significantly longer when the second encounter was appropriate than when it was inappropriate, but only when the first encounter was appropriate. Second pass reading times in the target region for the second encounter were significantly longer for inappropriate targets regardless of whether they matched the first encounter. Cook and Myers (2004) interpreted these results as indicating that consistency with the discourse context (the prior encounter with the word) influenced the initial processing of the sentence, whereas the effect of world knowledge (appropriateness) was seen slightly later in the process of integrating the word into the sentence.

Cook and Myers (2004; Experiment 2) manipulated the nature of the context, rather than the prior occurrence of the target word. Target words were placed in either a neutral context or in one that justified the character taking on the inappropriate role. For the first pass reading times,

there was an interaction between appropriateness and context. The inappropriate targets yielded longer fixation times, but only in the neutral context. Second pass fixation times were longer for the inappropriate targets for both the neutral and supportive contexts. Cook and Myers (2004) interpreted these results as further support for their conclusion that discourse context influences initial word processing, but that world knowledge continues to influence the ease with which a word is integrated into the discourse representation.

In general, the results of the studies suggest that discourse context has an influence on a reader's prediction for upcoming words, either during first-pass sentence parsing, or during a later reanalysis process. Sharkey and Sharkey (1992), Hess et al. (1995) and Cook and Myers (2004) reported fairly consistent findings of a discourse-level influence, and both Hess et al. (1995) and Cook and Myers (2004) found evidence for discourse-level effects during first-pass sentences parsing. However, the findings of Schwanenflugel and While (1991) suggest that discourse-level context only influences word processing when the discourse-context is consistent with the sentence-level context.

The variability of the results across these studies may be related to the nature of the stimuli that were used. Two of the studies attempted to investigate the effect of discourse context by either manipulating a single sentence in the prior discourse context (Schwanenflugel & White, 1991) or by changing a single word in the context (Cook & Myers, 2004). In the other two studies (Hess, Foss & Carroll, 1995; Sharkey & Sharkey, 1992), the entire contexts were varied. While manipulating entire contexts might seem to be a more valid means of studying the influence of discourse context, it does create the possibility for unwanted sources of variability (e.g., differences in lexical items) that are not typically seen when a single, well-defined variable is manipulated.

There was little agreement among the studies regarding the nature of the interaction between sentence-level and discourse-level context effects on predicting upcoming words. The results of Schwanenflugel and White (1992; Experiment 2) suggest that discourse-level context only influences word processing when the discourse-context is consistent with the sentence-level context. However, in their third experiment, discourse-context was found to influence word processing, regardless of whether the sentence-level context was related to the target word. These results were very different from those obtained by Hess et al. (1995; Experiments 4-9) who found a significant effect of discourse context, but no consistent interaction between discourse-level and sentence-level influences on predicting upcoming words.

In addition to examining the influence of discourse context on word processing, Schwanenflugel and White (1991) also investigated how the varying the distance between discourse-level information and the target word affects the listener's ability to use this information to predict the upcoming word. They demonstrated that discourse-level information influences the prediction for an upcoming word across at least three unrelated sentences.

Cook and Myers (2004) found a significant interaction between discourse-context information and world knowledge in facilitating the prediction of upcoming words. This suggests that discourse-level information can interact with other information sources in predicting upcoming structures. If discourse context information were found to interact with verb (lexical) information in predicting upcoming subcategorizations, it would provide another example of the listener's ability to integrate sources of information.

1.4 SYNTACTIC PRIMING

Another potential source of influence on the ease with which verb subcategorizations might be processed is syntactic priming. *Syntactic priming* occurs when a previously processed syntactic structure facilitates processing of the same syntactic structure. The results of several research studies indicate that speakers tend to repeat syntactic structures that they have previously used (Bock, 1986; Bock & Loebell, 1990; Cleland & Pickering, 2006; Hartsuiker & Kolk, 1998; Pickering & Branigan, 1998; Potter & Lombardi, 1998) or have heard another speaker use (Cleland & Pickering, 2003; Branigan, Pickering, & Cleland, 2000; Potter & Lombardi, 1998). In addition, there is evidence that the comprehension of a syntactic structure is facilitated in cases where the same structure has been previously encountered (Branigan, Pickering, & Stewart, 1995; Frazier, Taft, Roeper, Clifton, & Ehrlich 1984; Noppeney & Price, 2004; Luka & Barsalou, 2005). Finally, the results of several studies suggest that the influence of syntactic priming might extend beyond local priming effects among adjacent sentences, and may persist across a number of unrelated sentences (Bock & Griffin, 2000; Kashak, Loney and Borreggine, 2006).

1.4.1 Local syntactic priming in production

The phenomenon of syntactic priming across adjacent sentences has been demonstrated in a number of studies. There is evidence that speakers are more likely to produce a particular syntactic structure when they have recently produced the same structure (Bock, 1986; Bock & Loebell, 1990; Hartsuiker & Kolk, 1998; Pickering & Branigan, 1998; Potter &

Lombardi, 1998) or have heard or read a sentence with that structure (Branigan, Pickering, & Cleland, 2000; Cleland & Pickering, 2003; Potter & Lombardi, 1998). The results of these studies are summarized in Table 9 (Appendix A).

Bock (1986; Experiment 1) asked participants to repeat transitive and dative sentences (priming sentences). The transitive sentences were either active or passive, and the dative sentences were either prepositional dative (PO) or double object (DO) dative. Examples of each of these sentence types, from Bock (1986), are shown below.

Prepositional Object (PO): The corrupt inspector offered a deal to the bar owner.

Direct Object (DO): The corrupt inspector offered the bar owner a deal (p. 359).

Following each sentence, the participants described a picture that depicted an action that could be described by the type of construction (either transitive or dative) that was used in the preceding sentence. For both transitives and datives, the proportion of picture descriptions that conformed to the target construction was determined. For PO and DO datives, as well as for both types of transitives, the proportion of each construction was greater following a priming sentence with the same syntactic construction.

In Experiments 2 and 3, Bock (1986) explored the possibility that the significant results of Experiment 1 were due to similarities in the humanness characteristics between the prime and target sentences, rather than due to the influence of syntactic priming. In Experiment 2, the materials and procedure were similar to those used for Experiment 1, except that only transitives were studied, and the humanness of the agent and patient were varied. In Experiment 3, an ongoing recognition task was used in order to encourage deeper processing on the part of the participants. For both experiments, passives were significantly more frequent following passive primes than following active primes. However, active sentences were no more likely following

active primes than following passive primes. In addition, there were no significant interactions with the human agency factor, indicating that the degree of priming was not significantly influenced by human agency.

Bock and Loebell (1990; Experiment 1) investigated whether syntactic priming effects such as those seen in Bock (1986) could be explained in terms of the conceptual similarities between the prime and target sentences. They examined structural priming of PO datives, DO datives, and prepositional locatives. Prepositional locatives were included, because they are structurally similar to PO datives, but conceptually distinct. Evidence that prepositional locatives could prime PO datives would indicate that structural priming operates independently from conceptual factors. An example of the prepositional locative construction, from Bock and Loebell (1990) is shown below.

Prepositional Locative: The wealthy widow drove an old Mercedes to the church (p. 7).

Following sentences of each type, participants described pictures of "dative events". Bock and Loebell (1990) found that PO sentences and prepositional locative sentences were equally effective in priming PO picture descriptions, but that significantly fewer PO picture descriptions were produced following DO primes. Conversely, there were significantly more DO picture descriptions following DO sentences than following the other two sentence types. This suggests that syntactic priming is not related to conceptual similarity between the prime and target sentences.

Bock and Loebell (1990; Experiment 2) examined structural versus conceptual influences in the priming of transitive sentences, using three types of primes: passives (with a *by* phrase), locatives (with a *by* phrase), and actives. Passives and locatives shared a syntactic structure, but were conceptually distinct. Example of these sentence types are shown below.

Passive: The construction worker was hit by the bulldozer.

Locative: The construction worker was digging by the bulldozer.

Actives: The construction worker drove the bulldozer (p. 18).

Participants repeated sentences with these three constructions, and described a picture following each sentence. The proportion of passives produced in the passive and locative conditions was similar, and the proportion in each was greater than in the active condition. However, there was no significant difference in the number of active picture descriptions across the three prime types, although it was numerically greater following the active primes.

The purpose of Experiment 3 (Bock & Loebell, 1990) was to explore the possibility that other sentence features, such as the prosody or the phonological form of the closed class words, might be responsible for the priming effect seen in the first two experiments. Participants repeated sentences of three different constructions: PO and DO dative constructions, and an infinitive form. The PO and infinitive forms were identical up to the word *to,* where they diverged. The number of syllables and the stress pattern of the PO sentences and infinitives were equated. Examples of the PO and infinitive constructions, from Bock and Loebell (1990) are shown below.

Prepositional dative: Susan brought a book to Stella.

Infinitive: Susan brought a book to study (p. 23).

Bock and Loebell (1990) found that the proportion of PO constructions was significantly greater following the PO primes than following the other two sentence types. There was no significant difference in the number of PO constructions following the infinitive and DO primes. This suggests that the priming occurred between syntactic representations, and was not related to lexical, metrical, or other types of information.

Hartsuiker and Kolk (1998) examined whether evidence for syntactic priming, such as that found by Bock (1986) and Bock and Loebell (1990), could be extended to two other types of constructions that occur in Dutch, but not in English: the medial dative and a passive construction with a sentence final passive participle. Examples of the English translation of the medial dative and the alternative passive form are shown below.

Medial dative: The sailor writes to his girlfriend a long letter.

Alternative passive form: The walker is by the mud dirtied (p. 151).

Hartsuiker and Kolk (1998; Experiment 1) examined structural priming for two types of transitives that occur in English (active and passives with a clause-final by-phrase (P1)), as well as an alternative passive form with a sentence-final passive participle (P2), which occurs in Dutch. In addition, they examined priming for PO datives, DO datives, and medial datives (MD). Participants read aloud sets of three priming sentences, all having the same target syntactic structures, and then described a picture. For the picture description, there was significant effect of prime type for both types of passive transitives, but not for active transitives. For the datives, a significant effect of prime type was found for the DO and MD constructions, but not for the PO constructions. They contrasted their results with those of Bock (1986; Experiment 1) who did obtain syntactic priming for actives.

The purpose of Hartsuiker and Kolk (1998; Experiment 2) was to explore the possibility that no priming effect for active transitives was found in Experiment 1 because there were twice as many passive as active primes. In addition, because no significant priming effect was found for PO passives in Experiment 1, they replicated this part of the previous experiment. The materials and procedure were similar to those from Experiment 1, except that the P2 structure was not included. They found no significant priming effect for the passive transitive structure,

and a marginally significant (significant by *participants* but not by *items*), but negative priming effect for the active transitive structure. For the datives, there was significant priming effect for PO and MD structures, but not for the DO construction.

In Experiment 3, Hartsuiker and Kolk (1998) explored the possibility that the lack of a significant positive priming effect for the active transitives in the first two experiments might have been due to the sentences being presented in the written rather than auditory modality, as was done by Bock (1986). They hypothesized that because, in their earlier experiments, the stimuli from each trial remained on the screen until the next trial, participants might have used the words that remained on the screen to avoid using the same syntactic structure. In this experiment, the sentences were presented auditorily, and no P2 or MD constructions were presented. They found no significant effect of prime type for either active or passive transitives, but they did find a significant priming effect was found for both types of datives.

In the majority of the previously described experiments, the same verb was used in the prime and target sentences. Pickering and Branigan (1998) sought to determine whether verb subcategorizations could be primed apart from their associated verbs. They stated that this would provide information about whether verb subcategorizations are individually associated with specific verbs in the lexicon, or are shared among verbs with which they are associated.

In Pickering and Branigan (1998; Experiment 1), participants wrote completions to pairs of sentence fragments, in which the verb in each fragment could take either a PO or DO completion. One fragment served as the prime fragment, and the other as the target fragment. For the prime fragment, the postverbal noun phrase was manipulated to make either the PO or DO completion more likely. In addition, the target sentence included either the same verb or a different verb as the priming sentence. A significant interaction between prime completion and

target completion indicated that participants were significantly more likely to produce target completions that were of the same type as the prime completion. In addition, there was a significant three-way interaction of prime completion, target completion, and verb, indicating that the interaction between target completion and prime completion was more pronounced when the prime and target contained the same verb. A separate two-way ANOVA (separating levels of the verb factor) provided additional support for this finding. There was a significant interaction between prime completion and target completion when the prime and target sentences include the same verb. However, this interaction was significant only by items, and not by participants, when the prime and target sentences included different verbs.

Pickering and Branigan (1998; Experiment 2) further explored syntactic priming for verb subcategorizations across sentences in which the verb was not the same in the priming and target sentences. Participants read sets of three sentence fragments: two priming fragments and a target fragment. The postverbal noun phrases of both priming fragments were manipulated, so that both priming sentences favored either DO or PO completions. Target sentence completions were scored in the same manner as in Experiment 1. A significant interaction between prime completion and target completion indicated that syntactic priming occurs, even among sentences that do not share verbs.

In Experiments 3, 4, and 5, Pickering and Branigan (1998) examined whether the magnitude of syntactic priming was greater across sentences in which the verbs shared the same tense (Experiment 3), aspect (Experiment 4) or number (Experiment 5). For all three experiments, the procedure was the same as for Experiment 1, except that particular features of the verb were either the same or different between the prime and target sentences. The investigators found that none of these verb features significantly influenced the strength of

priming. In Experiment 3, there was no significant three-way interaction among prime completion, target completion, and tense, suggesting that syntactic priming is unaffected by differences in tense between the prime and target. In Experiment 4, the three-way interaction among prime completion, target completion, and aspect was significant by items, but not by participants. In Experiment 5, no significant three-way interaction was found among prime completion, target completion, and number.

Potter and Lombardi (1998; Experiment 1) investigated how the occurrence of a particular syntactic structure might interfere with the recall of an alternative structure. They asked participants to read target sentences with PO and DO constructions, and to recall them after a brief interval. Between the presentation and recall of the target sentences, participants read a prime sentence that had a different syntactic structure. This structure was either the alternative dative form, a mismatching structure similar to a dative, or a control sentence. They found that, when the prime sentence included the dative form that was not used in the target sentence, the errors that participants made in recalling the target sentences tended to reflect the syntactic structure of the prime sentences. The prime sentences with surface structures similar to datives had significantly less influence on the recall of the target, but had significantly more influence than the control sentences.

Potter and Lombardi (1998) concluded that the priming of a particular syntactic structure can not only increase the likelihood of occurrence of that structure, but can also serve to decrease the likelihood that an alternative structure will be produced. However, there are a limited number of ways to express the propositional content of the DO and PO target sentences. This raises the question of whether the alternative structure was actually primed, or simply interfered with accurate recall of the target. Retroactive interference has been shown to affect performance on a

variety of recall tasks (Eakin, 2005; Martens & Wolters 2002), and it seems possible that the results of this study are related to a type of retroactive syntactic interference.

Cleland and Pickering (2006) used the same stem completion method as Pickering and Branigan (1998), to determine whether syntactic priming of DO and PO constructions occurs across written and spoken modalities. In Experiments 1 and 2, the primes were either written or spoken and the target was spoken. In Experiment 3, the primes were either written or spoken, and the target was written. In Experiments 1 and 3, the same verb was used in the primes and targets, and in Experiment 2, different verbs were used. All experiments made cross-condition comparisons of the prepositional object (PO) target proportion, a measure of the relative proportions of PO and DO target responses. For all experiments, the PO target proportion was significantly higher following PO primes than following DO primes. In addition, none of the experiments showed a significant *prime structure* by *modality* interaction, indicating that priming was similar when the prime and target were produced in the same or different modality. A further analysis that compared the results of Experiments 1 and 2 found a marginal *prime structure* by *verb* interaction ($p < .06$), which the authors interpreted as suggesting that the magnitude of priming was greater when the verb was repeated.

Overall, the results of the studies described in this section (Bock, 1986; Bock & Loebell, 1990; Cleland & Pickering, 2006; Hartsuiker & Kolk, 1998; Pickering & Branigan, 1998; Potter & Lombardi, 1998) suggest that production of a particular syntactic structure increases the likelihood of that structure being produced in upcoming sentences. However, it is also possible that the nature of some of the tasks used in these studies (e.g. Hartsuiker and Kolk, 1998), in which participants were asked to repeat sentences (or sets of sentences) with a given syntactic

structures, may have drawn the participants' conscious attention to the structure and, therefore increased the chance that they would repeat it.

Syntactic priming was shown to be more influential for some structures than for others. In particular, there is little evidence to support syntactic priming of simple active sentences (Bock, 1986; Bock & Loebell, 1990; Hartsuiker & Kolk, 1998). However, none of the studies in which a lack of priming of active sentence was found addressed the possibility that this might be related to the high frequency of this construction within the language. It seems possible that there was no difference in the proportion of active sentences across conditions because the construction occurred frequently across all conditions, and this limited the extent to which priming could cause an increase in its frequency. Alternatively, it seems possible that the greater number of exposures that participants have had to the active transitive structure (across their lifetime) may have made them less susceptible to the priming of that structure.

The influence of syntactic priming was demonstrated to be independent of animacy, conceptual, phonological, or prosodic similarities between the priming and target sentences (Bock, 1986; Bock & Loebell, 1990). Differences in tense, aspect, or number between the verbs in the target and prime sentences also did not appear to influence the strength of priming (Pickering & Branigan, 1998).

Finally, Pickering and Branigan (1998) found that, while the syntactic priming of verb subcategorizations is more robust when the same verb is used in the priming and target sentences, subcategorizations can be primed apart from their associated verbs. Cleland and Pickering (2006) found similar results, although the effect of verb similarity failed to reach statistical significance. These results are noteworthy, because they suggest that verb

subcategorizations are stored separately in the lexicon from the verbs with which they are associated.

1.4.2 Syntactic priming in comprehension

In addition to evidence for syntactic priming in production, there is also evidence for syntactic priming in comprehension. Specifically, there is evidence that, when a syntactic structure is repeated across sentences or within the same sentence, this leads to faster reading times for the later occurring repetitions of the structure (Frazier, Taft, Roeper, Clinton, & Ehrlich, 1984; Branigan, Pickering, &, Stewart, 1995; Noppeney & Price, 2004). Additional evidence for syntactic priming in comprehension comes from Luka and Barsalou (2005), who examined the relationship between syntactic repetition and grammaticality ratings. The results of these studies are summarized in Table 10 (Appendix A).

Frazier, et al. (1984) examined the extent to which parallelism within conjoined sentences facilitates sentence processing. Participants read a series of sentences, each of which was comprised of two conjoined clauses. The second clause of each of these sentences was either of the same or of a different construction than the first. They examined five sentences with five different types of conjoined clauses. The first three of these types were referred to as syntax sentences, and the last two as animacy sentences. Examples of these are shown below.

1.) Active/passive

The tall gangster hit John	and the short thug hit Sam.
John was hit by the tall gangster	and Sam was hit by the short thug

2.) Minimal/non-minimal attachment

John believed Tom's stories (were literally true)

and Mary believed Jim's stories (were fictitious)

3.) Nonshifted/shifted heavy NP

Mary wrote a long note about her predicament to her mother.

Mary wrote to her mother a long note about her predicament.

4.) Agent/theme

Jack rolled down the hill and George did too

The rock rolled down the hill and the bucket did too.

5.) Animate/inanimate

John phoned the doctor/library and his friend phoned the lawyer/museum (p. 424).

Frazier et al. (1984) asked participants to read sentences with clauses of each of these types using a self-paced procedure, and recorded the reading times for each segment. A significant main effect for parallel form indicated that reading times were faster for the second segment when it was of the same syntactic form as the first segment. There was no significant effect for construction type. In addition, no significant main effects or interactions were reported for the animacy variables.

Frazier et al. (1984) concluded that exposure to a particular syntactic structure increases the availability of that structure, thereby facilitating its processing. The lack of a main effect or interactions involving the animacy variable mirrors the findings of Bock (1986), which also suggest a lack of influence of animacy on syntactic priming. However, Frazier et al. (1984) did not control for lexical, metrical, and other factors that may have been more similar between the parallel sentences than between those of different syntactic forms. Therefore, it seems possible

that these other factors might have contributed to the significant influence of parallel form that was observed in this study. And, while the results of Bock and Loebell (1990) suggest that metrical and conceptual factors do not play a role in structural priming across sentences, it is possible that they do influence intrasentential structural priming.

In addition to the influence of structural repetition on reading times, there is evidence that exposure to a sentence that contains a structural ambiguity can influence the ease with which similar structural ambiguities are processed. Branigan, Pickering, Liversedge, Stewart, and Urbach (1995) describe the results of an unpublished study (Branigan, Pickering & Stewart, 1995) in which participants read sentences that were temporarily ambiguous between an early closure and a late closure. They found that participants' reading times were faster for early closure sentences when they followed another early closure sentence. Likewise, reading times for late closure sentences were faster when they followed another late closure sentence. The authors also found similar priming effects for two other types of locally ambiguous sentences: reduced relative/main clause ambiguities and complement clause/relative clause ambiguities.

However, Branigan et al. (1995a) also report that the results of this study only demonstrated this type of priming effect in comprehension for pairs of sentences in which one of the sentences produced a strong garden-path effect. They were not able to demonstrate a significant influence of prior exposure on two other types of ambiguities: reduced complement "NP/S" ambiguities and PP-attachment ambiguities. They also found no reliable priming effects for active/passive pairs or with subject/object pairs. These results suggest a potentially limited role for syntactic priming in sentence comprehension, and imply a stronger influence of syntactic priming on some structures than others. However, the experiment is not described in sufficient

detail to fully evaluate the findings, and to determine whether alternative explanations could account for the results.

Noppeney and Price (2004) used a self-paced reading paradigm to measure reading time for the two types of temporarily structurally ambiguous sentences: an early closure versus late closure clause boundary ambiguity and a reduced relative/main clause ambiguity. Examples of these types of sentences are shown below.

Clause boundary ambiguity

a.) *Late closure*: Before the director left the stage, the play began.

b.) *Early closure:* After the headmaster left, the school deteriorated rapidly.

Reduced relative/main clause ambiguity

a.) *Simple active*: The artist left his sculptures to the British museum.

b.) *Reduced relative*: The child, left by his parents, played football (p. 703).

Sentences of these four types followed sets of four sentences that were either all structurally similar to the target (primed condition) or were of each of the four different target structures (unprimed condition). For both types of sentences, Noppeney and Price (2004) found that reading times were faster after similar than dissimilar blocks. This provides further evidence for the role of structural priming in sentence comprehension.

Luka and Barsalou (2005) measured the extent to which exposure to a syntactic structure facilitates the comprehension of that structure. They used grammatical acceptability judgments as measures of ease of processing and the influence of repeated exposure. They proposed that this technique would be a different, albeit less direct means of examining the process of structural facilitation. This method was based on earlier studies, including one by

Monahan, Murphy, and Zajonc (2000), which suggest that individuals tend to exhibit a preference for stimuli to which they have been repeatedly exposed.

Prior to their first experiment, Luka and Barsalou (2005) conducted a norming study in which 40 participants rated sentences with various constructions for grammaticality, on a scale that ranged from ungrammatical (1) to perfectly grammatical (7). In Experiment 1, a separate group of participants rated a set of ungrammatical, moderately grammatical, and highly grammatical sentences. The participants had read half of the sentences during an earlier phase of the experiment, and half of the sentences were new. Although sentences that had received ratings along the entire grammaticality continuum were used in this experiment, only the results involving moderately grammatical sentences were analyzed, in order to avoid ceiling effects and the inclusion of sentences that were clearly ungrammatical. Results showed that participants gave significantly higher ratings to the moderately grammatical sentences that they had read earlier than to the new items.

Luka and Barsalou (2005; Experiment 2) examined whether exposure to sentences that were structurally or conceptually similar to a test sentence would increase the acceptability ratings for the test sentences. They assigned participants to one of two conditions. In the first condition, half of the test sentences corresponded to a syntactic structure that had been presented earlier in the experiment. Three structural variants of each syntactic type were read during the first phase of the experiment, in order that "priming effects might accrue over repeated exposures" (p. 441). In the second condition, half of the test sentences corresponded to three positional variants of the test sentence that had been read earlier in the experiment. These positional variants had the same lexical and conceptual content as the test sentence, but were structurally distinct. For both conditions, half of the test sentences were new to the participants.

The investigators found that the grammaticality ratings were significantly higher for sentences where structurally similar sentences had been previously encountered. However, there was no significant difference in ratings between new sentences and those for which positional variants had been previously encountered.

In Experiment 3, Luka and Barsalou (2005) examined the effect of number of exposures for identical items and structurally related items. One group of participants rated new sentences and those that had occurred once, or had been identically repeated three or five times in the earlier reading phase (token condition). The other group rated the same sentences after reading zero, one, three or five sentences that were structurally-similar, but not identical, to the test sentence (type condition). These sentences were previously rated as moderately or highly grammatical by another group of participants. The authors found that the moderately grammatical sentences were more strongly influenced by familiarity than the highly grammatical sentences, although this difference may have been at least partially attributable to ceiling effects. Type repetition was shown to be effective in increasing grammaticality ratings but token repetition was not. In addition, there was a significant linear trend across the levels of exposure of the type condition, indicating that exposure to multiple variants of a grammatical structure boosted grammaticality ratings in a predictable way.

The aim of Luka and Barsalou (2005; Experiment 4) was to re-examine the possible influence of token repetition on sentence processing. They hypothesized that the lack of a significant influence of token repetition in Experiment 3 might have been due to the participants' not attending to the entire stimulus sentence, once it had been recognized as a repetition. In addition, they wanted to verify that the structural facilitation effect could be observed after a single exposure to a syntactically-related sentence. Participants read aloud sentences that were

either identical to or structurally-similar to half of the test sentences. Both identical and structural repetition significantly influenced the participants' grammaticality ratings. Again, the effect of exposure was greater for the moderately grammatical sentences than for the highly grammatical sentences.

In Experiment 5, Luka and Barsalou (2005) re-examined the influence of multiple identical repetitions during reading. In addition, they attempted to remedy a shortcoming of the earlier experiments that compared the influences of token and type repetition. The participants' ratings in the token repetition condition might have been more greatly influenced by their awareness of repetition than the ratings of the participants in the type repetition condition. The reading phase was made more similar across the conditions by requiring the participants to read four repetitions of each sentence. Then, they rated sentences that were either identical or structurally-related to those that had been read earlier, or that were new. The investigators found that both identical and structurally-related repetitions influenced grammaticality ratings, and that moderately and highly grammatical sentences were similarly influenced.

The results of these studies provide some evidence for syntactic priming in language processing, both across sentences and within clauses of the same sentence. In addition, both Frazier et al. (1984) and Luka and Barsalou (2005) found that this effect was independent of nonsyntactic factors, such as animacy and conceptual similarity. While the results of the unpublished study described in Branigan et al. (1995) indicated that this effect was limited to pairs of sentences in which one of the sentences exhibited a strong garden-path effect, the lack of details regarding this experiment make this claim difficult to judge.

The results of Luka and Barsalou (2005) and Noppeney and Price (2004) suggest that the influence of syntactic priming can increase across repeated exposures to the same syntactic

structure, although neither of these studies explicitly examined the magnitude of this influence. In addition, while the results of Luka and Barsalou (2005) indicate that participants generally rated more familiar structures as more grammatical, this evidence of an increase in ease of processing is only indirect.

1.4.3 Shared syntactic priming between comprehension and production

In the previous sections, evidence was presented which supports the existence of syntactic priming in both production and comprehension. In this section, evidence will be presented that syntactic priming can occur from comprehension to production (Branigan, Cleland, & Pickering, 2000; Cleland & Pickering, 2003; Potter & Lombardi, 1998). That is, speakers and writers are more likely to produce a particular syntactic structure after having recently heard or read a sentence with the same structure. Evidence for comprehension to production priming is important, because it suggests that syntactic priming results from the activation of a residual syntactic representation that is shared between comprehension and production, as was proposed by Pickering and Branigan (1999). Investigators in earlier studies (Bock, 1986; Bock & Loebell, 1990) proposed that syntactic priming involves the activation of the procedure involved in producing a sentence of a particular structure, and therefore could not be shared between comprehension and production. The results of the studies that will be described in this section are summarized in Table 11 (Appendix A).

Potter and Lombardi (1998; Experiments 2 and 3) compared the influence of comprehension-to-production syntactic priming to that of production-to-production syntactic priming. Participants silently read and then recalled sentences in which there was a prime clause

and a target clause, each of which included either PO or DO construction. The prime clause either occurred before (Experiment 2) or after (Experiment 3) the target clause, so that the prime was either repeated before the target, or was heard by the participant but not repeated. For both experiments, the dependent variable was the proportion of recalled target clauses in which the structure of the prime clause was substituted for that of the target clause. Regardless of whether the prime clause was repeated before the repetition of the target, the participants' errors in the repetition of the target clause reflected the syntactic structure of the prime clause. This suggests that syntactic priming can influence production of a target, even when it has not already been produced by the participant.

Branigan, Pickering, and Cleland (2000) examined the extent to which participants in dialogue co-ordinate the syntactic structure of their utterances. Pairs of individuals (a participant and a confederate who was posing as a participant) took turns describing pictures and selecting pictures that matched their partner's description. The experimental pictures showed ditransitive actions. Pairs of prime and target cards were created by pairing each of the cards from the confederate's set (the prime card) with a card from the participant's set (the target). The verb to be used in the picture description was printed under each picture. The prime and target cards were either associated with the same or a different verb. In addition, half of the prime cards were assigned PO descriptions and half were assigned DO descriptions.

Branigan et al. (2000) found a significant main effect for prime type. Participants produced a greater proportion of each structure following a description that served to prime that structure. A significant interaction between prime type and verb identity indicated that this effect was stronger when the same verb was used in the prime and target sentences. This second finding lends further support to the results of Pickering and Branigan (1998; Experiment 1) and

Cleland and Pickering (2006) who also found that priming effects were greater when the verb was shared between the priming and target sentences.

The purpose of Cleland and Pickering (2003) was to examine syntactic priming for a relative clause construction and a pre-nominal construction, and to determine whether the greater syntactic priming effect that had been found across sentences sharing verbs could be generalized to other types of phrasal heads (nouns and adjectives). They used the same confederate scripting method that had been used by Branigan et al. (2000). Examples of the two constructions that were examined in this study are shown below.

Relative clause: The square that's red.

Pre-nominal construction: The red square (p. 219).

Cleland and Pickering (2003; Experiments 1-3) found that the proportion of relative clause and pre-nominal constructions produced by the participants were greater following priming constructions of the same type. In Experiment 1, they also varied whether the noun or adjective was the same or different between the priming and target constructions. The proportion of target constructions produced by the participants was significantly greater when nouns were shared between the prime and target. However, the influence of shared adjectives between the prime and target was not significant. In Experiment 2, the noun in the prime was either the same as the one used in the target, was semantically related to it, or was unrelated. No significant difference in the proportion of target constructions produced by the participants was found between the same noun and related noun conditions, but there was a significant difference between the related noun and unrelated noun conditions. In Experiment 3, the nouns in the target and prime sentences were either the same, were phonologically similar (sharing initial and final phonemes) or were phonologically dissimilar. There was a significant difference between the

same noun and phonologically similar conditions, but not between the phonologically similar and phonologically dissimilar conditions. These results suggest generalization to nouns of the findings of previous studies for greater influence of syntactic priming across sentences that share verbs. The lack of significant findings across phonological similarity conditions supports results from other studies (e.g., Bock & Loebell, 1990), that syntactic priming is not generally influenced by this type of information.

The combined results of Potter and Lombardi (1998; Experiments 2 and 3), Branigan et al. (2000) and Cleland and Pickering (2003) strongly suggest that syntactic priming is shared between comprehension and production. This implies that syntactic priming involves the activation of a syntactic representation, and is not simply a procedure for producing a particular syntactic structure.

In addition, there is evidence that the magnitude of syntactic priming is greater for prime-target pairs that share verbs or nouns, or have semantically similar nouns (Cleland & Pickering, 2003). This finding, along with the findings of other studies (Branigan, Pickering, & Cleland, 2000; Pickering & Branigan, 1998) that suggest syntactic priming does occur among sentences with different verbs, might imply an interaction between syntactic processing and activation of the combinatorial properties that are held to be associated with lexical items in a sentence.

1.4.4 Accounts of syntactic priming

As mentioned earlier, various theorists account for syntactic priming in different ways. Bock (1986) and Bock and Loebell (1990) proposed a *framing assumption*, in which syntactic structure serves as the frame of the sentence, and words are anchored onto this frame. According

to this account, syntax serves as a skeleton for the production of the sentence, and shapes the process involved in producing the sentence. Therefore, syntactic priming involves the activation of a particular frame, or procedure for building the sentence.

Alternatively, Pickering and Branigan (1999) argue that syntactic priming involves the activation of a residual syntactic representation that is common to both comprehension and production. Supporting evidence for this view comes from several sources. First, evidence from the studies that were reviewed in the previous section suggests that syntactic priming can occur from comprehension to production. Second, the results of Cleland and Pickering (2006) suggest that syntactic representations are shared between spoken and written language. Both of these findings would be difficult to account for under the framing assumption. Finally, the findings of Luka and Barsalou (2005) for increased grammaticality ratings for familiar structures also suggest the existence of an underlying structural representation.

The conceptualization of syntactic priming as involving the underlying representation of a target syntactic structure is an important one for the argument for the existence of a cumulative syntactic priming effect. As will be discussed later, researchers such as Kashak, Loney, and Borrenggine (2006) propose that listeners and readers are sensitive to the relative frequency of these same types of syntactic structures.

1.4.5 Cumulative syntactic priming

While the studies described in the previous sections provide evidence that syntactic priming among adjacent sentences (local syntactic priming) influences both sentence comprehension and production, a few recent studies (Bock & Griffin, 2000; Kashak & Glenberg,

2004; Kashak et al., 2006) have also found evidence for a *cumulative syntactic priming* effect, in which the influence of syntactic priming accumulates across trials or sentences. There are two types of relevant evidence. First, syntactic priming is long-lasting, and can persist across as many as ten unrelated sentences (Bock & Griffin, 2000). Second, the frequency with which a particular structure occurs across recently presented sentences is significantly associated with the likelihood that participants will use to structure to complete upcoming sentence fragments (Kashak et al., 2006).

Both the persistence of syntactic priming and the increased magnitude of its effect with greater frequencies of presentation suggest that syntactic priming involves more than the transient activation of a residual syntactic representation (Pickering & Branigan, 1998, 1999). Bock and Griffin (2000) propose that syntactic priming reflects a long-term implicit learning process, in which the process of comprehension is gradually adjusted in response to changes in the input. Each exposure to various types of language input (e.g., a particular syntactic structure) increases the ease with which that particular type of input is processed. A similar proposal was made by Mitchell et al. (1995). Kashak and Glenberg (2004) also suggest that syntactic priming plays a role in children's acquisition of particular syntactic constructions. The results of the studies that will be discussed in this section are summarized in Table 12 (Appendix A).

The proposal that exposure to a particular linguistic structure results in long-term changes to the processing system is not unique to studies of syntactic priming. There is evidence from studies of *repetition priming* that items to which participants have been previously exposed tend to be processed more rapidly and are associated with higher response accuracy than new items. This effect has been shown to occur in the absence of participants' ability to recognize or recall the target words (Cave, 1997).

Repetition priming has been viewed by some theorists as involving long-term changes in processing because it has been shown to persist across unrelated trials and over long time periods. For example, Humphreys, Besner, and Quinlan (1988; Experiment 1) found evidence for repetition priming in a word identification task across six intervening trials. Cave (1997) found no significant decline in the degree of repetition priming on participants' performance on a picture identification task across intervals as long as 48 weeks. Across each interval, the pictures that participants had seen previously were associated with faster naming times than new pictures. This priming effect at 48 weeks was even found in a subgroup of 49 participants who recognized less than 10% of the items.

The long-term changes of repetition priming for lexical items might suggest that exposure to various syntactic structures (including verb subcategorizations) would have a long-term influence on the processing system as well. As stated at the beginning of the paper, it is widely held that verb subcategorization preferences result from a listener's or reader's history of experience with a particular verb (Boland, 1993; Jennings et al., 1997; Mitchell et al., 1995; Trueswell, et al., 1993).

If a verb's association with a particular subcategorization (subcategorization preference) results from an individual's long-term experience with that verb, this suggests that this cumulative experience was acquired through a large number of individual experiences with the verb co-occurring with various subcategorizations. The preferred subcategorization would be the one with which the verb was most frequently associated during these individual trials. In order for these individual experiences to result in changes in the long-term association between the verb and a particular subcategorization, it seems likely that individuals incrementally adjust their

expectations, based on the short-term frequency information that is available in everyday language input.

Jurafsky (1996) proposed a probabilistic model for syntactic disambiguation that is relevant in this regard. According to his model, sentence processing involves maintaining multiple interpretations of an ambiguous sentence, and each interpretation is ranked according to its relative probability. For verb subcategorizations, this ranking involves sensitivity to the long-term frequency of its association with the verb. However, Jurafsky (1996) also proposed that other, more local factors such as the prior sentence context and the nature of the immediately preceding discourse context can influence the relative ranking of these structures, as well.

Bock and Griffin (2000) examined the persistence of syntactic priming across varying numbers of intervening sentences. In each of two experiments, participants repeated sets of auditorily presented sentences. One sentence in each set was intended to prime a target structure (either simple passive or PO dative). Following each set, participants described pictures that could be described easily using the intended structure. The number of filler sentences (lag conditions) between the priming sentence and the picture description was varied. In Experiment 1, the priming sentences were either immediately before the target, or there were either one or two filler sentences between the prime and the target. In Experiment 2, the prime was either immediately before the target, or there were either four or ten filler sentences between the prime and the target.

Bock and Griffin (2000; Experiment 1) found that the proportion of picture descriptions that conformed to the target structure was significantly greater for the primed (versus unprimed) condition, across all three lag conditions. In addition, the difference between the primed and unprimed conditions did not differ significantly across the lag conditions. For Experiment 2, a

significant difference was also found between the proportion of conforming structures that the participants generated in the primed and unprimed conditions. However, this difference was only found when the priming sentence was either immediately before the target, or when there were ten intervening sentences between the prime and the target. There was no significant difference between the primed and unprimed conditions when there were four intervening sentences between the prime and the target. When the results for the PO datives and passives were examined separately, the investigators found that the lack of a significant priming effect in the condition with four intervening sentences was due to an unexplained negative priming effect for the passive sentences in this condition.

Kashak and Glenberg (2004; Experiment 1) examined the role of structural priming in learning by exposing adult participants to a novel construction. They examined the extent to which incremental increases in the number of exposures to the construction resulted in changes in reading times across that construction. Participants were trained on the *needs* construction, which is "a feature of American English spoken and in the northern midlands dialect region" (p. 450). Participants read passages, sentence by sentence, from a computer screen. Each passage began with a series of introductory sentences. Following these, participants read one of two types of critical sentences: a *needs* sentence or a standard sentence. Examples of these sentence types, from Kashak and Glenberg (2004), are shown below.

Needs construction: The meal needs cooked, given that dinner is in an hour.

Standard construction: The meal needs to be cooked, given that dinner is in an hour

(p. 451).

To monitor comprehension, the sentence following the critical sentence was manipulated, so it was either consistent or inconsistent with the meaning of the passage. The authors measured

reading times for this sentence, because earlier research studies have indicated that sentences related to the meaning of a text are easier to comprehend than those that are unrelated.

Participants read a total of 24 critical sentences, 12 of which included the *needs* construction. An ANOVA conducted on the reading times for the critical sentences across sentence types (*needs* versus standard construction) and time periods (trials 1-4, 5-8, 9-12) yielded a significant interaction between sentence type and time period. The participants read the *needs* construction significantly more slowly during trials 1-4, but this difference was significantly reduced for trials 5-8 and trials 9-12. For the follow-up sentences, they read inconsistent sentences more slowly than consistent ones for both sentence types and across all time periods. The faster reading times on the *needs* construction (relative to the standard construction) suggests that the participants quickly learned the new syntactic construction. The finding that they read the follow-up sentences more quickly on consistent than inconsistent trials suggests that the participants comprehended the *needs* construction across all of the sets of trials.

The purpose of Kashak and Glenberg (2004; Experiment 2) was to determine whether readers can extend their knowledge of a novel construction to a new verb, *wants*. In addition, they replaced the sentence-by-sentence reading procedure with a word-by- word procedure in order to obtain a more "fine grained" view of the learning of a novel construction. In the first phase of the experiment, half of the participants read sentences with the *needs* construction, and the other half read standard versions of the same sentences. Following this, the participants read ten *wants* sentences.

The results of this experiment supported the major finding of Experiment 1. Once again, there was a significant *time* by *training* interaction for the disambiguating region. Participants in the *needs* training condition read sentences more slowly than those in the control condition in

trials 1-3, but this difference was not significant for trials 4-7 or trials 8-10. In addition, participants in the *needs* training condition read *wants* sentences significantly faster than those in the control condition, suggesting that their acquisition of a new construction was generalized to a different verb.

In Experiment 3, Kashak and Glenberg (2004) sought to determine how exposure to the *needs* construction would influence the processing of a construction with which the participant was already familiar. First, participants were trained either on the *needs* construction or on the standard construction. Then, the participants continued to read either the *needs* construction or the standard construction, but also began to read a modifier construction. The investigators chose this construction because they predicted that readers who are unfamiliar with the *needs* construction would initially misinterpret it as this modifier construction. An example of the modifier construction is provided below.

Modifier construction: The ceramic tile wall needs washed stickers to be put on it (p. 467).

Kashak and Glenberg (2004; Experiment 3) found that participants in the *needs* condition read the modifier sentences more quickly than those in the standard condition. However, the reading times for the *needs* and standard sentences during this phase of the experiment did not significantly differ. This suggests that the initial misinterpretation of the *needs* construction during training helps to facilitate later processing when the modifier interpretation turns out to be the correct reading.

In Experiment 4, Kashak and Glenberg (2004) tested their explanation for the results of Experiment 3. As in Experiment 3, the participants received prior training on the *needs* construction. However, one group of participants received additional instructions prior to the *needs* training that would reduce the likelihood that they would consider the modifier

interpretation while reading the *needs* sentences. Participants who received the additional instructions slowed down when moving from the *needs* construction to the modifier construction, relative to those who did not receive the additional instructions. The authors interpreted the results of these last two experiments as evidence for an *episodic processing account* of learning. According to this account, a construction becomes easier to process as episodes of processing that construction accumulate in memory.

While these results were taken as evidence for the influence of syntactic priming on sentence processing, they may be confounded by other aspects of learning. For example, although exposure to the syntactic structure likely played a role in the participants' increased reading times, the participants were also becoming more familiar with the meaning of this construction. Increased knowledge of the meaning of the construction could have led to faster processing times, and could have been at least partially responsible for the generalization to the *wants* construction.

Kashak et al. (2006) examined the extent to which frequency and recency of experience with a particular syntactic construction influence the likelihood that participants would produce that construction. Experiment 1 consisted of two phases: a Recent Experience phase and a Priming phase. In the Recent Experience phase, participants completed sentence fragments which were intended to induce either a prepositional object (PO) or double object (DO) completion. There were three conditions. In the equal exposure (EE) condition, participants produced an equal number of each construction, with tokens of each construction placed randomly within each set. In the equal exposure-block (EE-B) condition, participants produced an equal number of each construction, but tokens of each construction occurred either entirely in the first or second half of the trials. In the unequal exposure (UE) condition, participants were

exposed to tokens of only one construction. Following this, half of the participants in each condition were assigned to either PO or DO priming trials. During these trials, they completed a total of six priming stems, and following each of these, completed one target sentence stem.

Kashak et al. (2006) classified each of the participants' target stem completions as DO, PO, or other. They found a significant interaction between Recency of Experience, Prime Type and Target Response. For both PO and DO target responses, participants in the EE and EE-B conditions produced more target responses following primes that matched the target than those that did not. However, this difference was not found for the UE condition. In addition, the results of an interaction contrast indicated that the influence of structural priming was equally strong in the EE and EE-B conditions, and stronger in these conditions than in the UE condition. These results suggest that the frequency of exposure to a certain construction, but not the temporal distribution of exposure to that construction, influenced the likelihood of that structure being produced.

In Kashak et al. (2006; Experiment 2), the procedure was similar to the one used in Experiment 1. This experiment addressed the possibility that the run of one construction at the end of the EE-B condition did not reduce the influence of priming for the other construction because the run was not long enough to influence priming. For this reason, the participants in this experiment completed twenty (rather than ten) prime stems in the recent experience phase. The authors hypothesized that, if the relative frequency of a particular structure is related to the influence of structural priming on production, then the influence of priming in this experiment should be as strong as when there were only ten prime stems. An additional level of frequency was also added, by substituting a UE-75 condition for the EE condition. In the UE-75 condition, 75% (15) of the twenty trials included the target construction. Results indicated that, for both DO

and PO primes, participants produced more target completions in the UE-100 condition (exposure to only one construction) than in the UE-75 condition, and more in the UE-75 condition than in the UE-50 condition.

The results of these studies suggest that the influence of syntactic priming on the likelihood that a structure will be produced accumulates across trials (Kashak & Glenberg, 2004; Kashak et al., 2006). Investigators from studies that were described in earlier sections of this paper (Pickering & Branigan, 1998; Luka & Barsalou, 2005; Noppeney & Price, 2004) also manipulated the number of constructions of a particular type that was used to prime a target construction. However, neither Pickering and Branigan (1998) nor Noppeney and Price (2004) reported on the effect of this manipulation. Luka and Barsalou (2005) reported a linear increase in the influence of priming with increased numbers of exposures to a particular structure. Kashak et al. (2006) found that this frequency-based effect even occurred in cases in which sentences of other structures were presented between the prime and target sentences.

Bock and Griffin (2000) provide evidence that the influence of syntactic priming is long-lasting, and not simply an influence that occurs among adjacent sentences. In addition, Kashak et al. (2006) found that the frequency, but not the recency of a particular construction is related to the likelihood that the construction will be used to participants in a sentence completion task. This suggests that the influence of syntactic priming is more long-lasting than previously assumed (e.g., Bock & Loebell, 1990).

Syntactic priming influences the processing of structurally ambiguous sentences, as well (Kashak & Glenberg, 2004). This was illustrated for the *needs* sentences that were temporarily ambiguous between the correct reading and a modifier construction reading. Kashak and Glenberg's (2004) finding that exposure to the *needs* construction can influence reading times

for the *wants* construction coheres with the findings of other researchers, such as Pickering and Branigan (1998), which suggest that the subcategorizations are shared among the verbs with which they are associated.

One limitation of the literature on syntactic priming is that it has been studied on a relatively narrow range of constructions. Therefore, the extent to which the results of these studies can be generalized to other constructions is unknown. In addition, the studies that provide some evidence for cumulative syntactic priming are limited by the fact that they examined this phenomenon across sets of unrelated sentences. It is unclear whether similar results would occur for natural language production, which almost always involves some meaning-related coherence among sentences.

1.5 CONCLUSION

As stated earlier, Gibson's Discourse Locality Theory (DLT, 1998; 2000) describes the process of sentence comprehension as a continual updating of syntactic structure as new lexical items are encountered. Predictions for upcoming syntactic candidates are made by comparing the activation levels of the candidate structures. The activation level of each structure is determined by lexical, plausibility, and discourse context constraints. However, the theory does not specify the precise nature of these constraints, and how they might interact in influencing a comprehender's predictions for upcoming structures, including verb subcategorizations.

This paper has discussed two potential sources of influence (constraint) on predictions for upcoming verb subcategorization: discourse context effects and cumulative syntactic priming. Evidence suggests that both of these factors can influence the ease with which syntactic

structures are processed. However, no known study of either discourse context or syntactic priming has considered the influence of the other. None of the reviewed studies on discourse context effects controlled for the proportion of the various syntactic structures within the sentences that composed the discourse context. Likewise, none of the studies of syntactic priming has examined its influence within a natural context.

If syntactic priming does influence the prediction for upcoming verb subcategorizations, it is unknown whether the magnitude of its influence would be the same within a discourse context as has been demonstrated for a series of unrelated sentences. Discourse context involves the formation of a global representation, which reflects the accumulated meaning of the discourse. Participants who read, listen to, or complete a series of unrelated sentences are not required to attend to this type of global representation. This might allow them to attend more closely to the syntactic structure of the sentences that are presented. Similarly, syntactic priming experiments, in which a small number of target structures are repeated frequently in the context of a limited number of filler sentences, might allow participants to attend more closely to target structures than would be the case within a more natural context.

Because verb subcategorization information is argued by some theorists (Pickering & Branigan, 1998) to be stored as part of the verb's lexical entry, the prediction for alternative subcategorizations based on this type of verb-specific information represents an important interface between syntactic and lexical levels of processing. Therefore, it would be particularly interesting, as well as informative to models of sentence processing, to better understand how syntactic information as well as discourse context information influence readers and listeners in making this prediction.

In addition, this study also has implications for the development of augmentative and alternative communication (AAC) systems. In order to enhance communication rate and to promote keystroke savings, some designers of AAC systems (e.g. Mayer-Johnson) have included algorithms that allow systems to make predictions about which words a user is likely to use next, based on the portion of the message that has been constructed up to that point. This feature, called *linguistic prediction*, is based on various types of syntactic information about the language. To date, these predictions are based on sentence-level information, and the use of this information has resulted in keystroke savings as great as 46% in some programs (Hunnicutt & Carlberger, 2001).

However, much of the evidence presented in the Review of the Literature strongly suggests that discourse context also influences the prediction for upcoming structures, including verb subcategorizations. If discourse level information could be used to increase the accuracy of predictions made by AAC technology, this could lead to further keystroke savings for individuals with severe communication disorders.

Finally, the outcome of this investigation also might have implications for further research in the area of language disability. According to Nippold (1998), older children and adolescents with language-learning disability often have associated reading difficulties. Because of their lack of exposure to grade-level reading materials, they are often restricted in their experience with certain vocabulary that is more prevalent is written than in spoken language.

One class of words that appears more often in written than in spoken language is metacognitive verbs (e.g., believe, doubt, predict; Nippold, 1998). Because many of these verbs subcategorize for S and NP, it seems possible that individuals with language disabilities, due to their lack of experience with these verbs, are at less accurate than their peers in using the

subcategorization preferences of these verbs to predict upcoming structures. In addition, because they have less experience with the texts in which these verbs are embedded, they may also experience difficulty in using inter-sentential sources of information in influencing their predictions.

If the results of this study indicate that the syntactic predictions of individuals without a history of language impairment are influenced by either discourse context information or syntactic priming, then it would be important to determine whether these factors influence the predictions made by individuals with language impairments. If individuals with language impairments do not make use of the relevant types of information, this might suggest that exposing these individuals to these more literate language forms, and teaching them make more efficient use of inter-sentential information might help to improve the efficiency of their language processing.

2.0 PURPOSE

The purpose of this investigation was to examine whether syntactic priming and the influence of discourse context independently or interactively influence readers' predictions about upcoming verb subcategorizations. Specifically, the study addressed the following questions.

1.) *Does the proportion of verb subcategorizations in the input influence reading times for verb subcategorization?* It was hypothesized that reading times for subcategorizations following verbs would be influenced by the proportion of verb subcategorizations in the prior discourse context.

2.) *Is there a significant difference in reading times for verb subcategorizations that follow coherent discourse contexts (designed to increase expectation for those subcategorizations) and those that follow a series of scrambled sentences?* It was hypothesized that reading times for verb subcategorizations would be faster following the discourse contexts.

3.) *Does the influence of the proportion of verb subcategorizations in the input differ significantly between coherent discourse contexts and scrambled sentences (across which a global representation is less likely to be generated)?* It was hypothesized that the influence of the proportion of verb subcategorizations in the input would be greater for scrambled sentences than for coherent discourse contexts.

3.0 METHODS AND PROCEDURES

3.1 OVERVIEW OF THE EXPERIMENT

The first purpose of this experiment was to determine the extent to which syntactic priming influences readers' predictions for verb subcategorizations. This research question was addressed by comparing participants' reading times across the disambiguating and post-disambiguating regions of temporarily ambiguous sentences under conditions in which the number of exposures to a particular verb subcategorization (S) in the preceding sentences was varied. The second purpose was to determine whether the influence of a discourse context that was designed to increase the expectation for a particular verb subcategorization would influence readers' predictions for that subcategorization. This question was addressed by comparing participants' reading times across the disambiguating and post-disambiguating regions of temporarily ambiguous sentences when the preceding sentences either constituted a coherent discourse context (designed to increase readers' expectation for a sentential complement), or comprised a series of scrambled sentences (in which all but the first and last sentences were selected quasi-randomly from a different coherent text). A final purpose was to examine whether syntactic priming exerts the same degree of influence for predicting verb subcategorizations within a discourse context, as has been demonstrated across sets of unrelated

sentences. This question was addressed by examining the interaction of the previously-described syntactic priming and discourse context factors.

These questions were addressed using a within-subjects design. Each participant was randomly assigned to one of four stimulus sets. Each stimulus set comprised one-quarter of the items in each of four conditions, which were defined by their proportion of target subcategorizations and the presence or absence of a coherent discourse context. Before beginning the experimental task, participants completed six practice trials, which served to familiarize them with the task. The experimental task comprised a total of 40 trials: 16 experimental trials and 24 filler trials. For the experimental trials, participants read a series of eight sentences, in which the matrix verb in each sentence was followed by an S or NP. Examples of these sentence types are provided below.

Verb + S: Diana maintained her opinion was well-informed.

Verb + NP: Diana maintained her opinion about the political candidate.

For the filler trials, participants read a series of sentences, ranging in number between 5 and 12. The majority of these sentences differed in syntactic construction (including the types of subcategorizations) from those used in the experimental conditions. Half of the filler sets comprised a coherent discourse context and half did not, and the proportion of sets which either did or did not comprise a coherent discourse context was equated across filler sets of varying lengths.

For both the experimental and filler trials, sentences were presented one at a time, with the participant controlling the speed of presentation by pressing a button to display each successive sentence. Following each set of sentences, a target sentence was presented one word at a time, with the participant pressing a button to display each successive word. For the

experimental trials, reading time for each word in the disambiguating and post-disambiguating regions of the target sentence was measured as the amount of time between button presses. Comprehension questions appeared following one-third of all trials, to encourage to participants to read the sentences for comprehension.

3.2 PARTICIPANTS

3.2.1 Sample size

Forty participants between the ages of 20 and 35 were recruited from the Indiana, PA community. An *a priori* power analysis for a two-way within-subjects ANOVA indicated that 36 participants would be needed to obtain an effect size (Cohen's *d*) of 0.5 with a power of 0.8 and alpha set at .05 (Cohen, 1988) to detect both main effects and interactions. This sample size was determined through the use of a formula provided by Cohen (1988; pg. 396) for calculating the sample sizes needed for a two-way ANOVA.

The effect size for main effects was estimated from results reported for the influence of syntactic priming or discourse-context influences on sentence parsing (Hess, et al., 1995; Kashak & Glenberg, 2004; Kashak et al, 2006; Luka & Barsalou, 2005; Noppeney & Price, 2004). The only study of discourse context effects for which effect size values were available (Hess et al.,1995) consistently reported large effect sizes (Δs ranging from 1.876 – 8.051; M = 4.401). There was more variability in the effect sizes reported for the studies of syntactic priming (Δs ranging from 0.222- 8.462; M = 2.406). The generally large effect sizes reported in these studies suggested that large effect sizes could be predicted for the main effects in the current experiment.

However, because no data were reported regarding the effect sizes for the interaction of these two factors, there was no strong evidence for estimating the effect size for the interaction. Therefore, a more conservative (but relatively large) estimated effect size of 0.5 was chosen. An additional four participants (10%) were recruited in an attempt to ensure adequate power, even if data were lost due to outliers, equipment failure, etc.

3.2.2 Recruitment

Participants were recruited through the Department of Special Education and Clinical Services at Indiana University of Pennsylvania. In addition, flyers were distributed at Indiana University of Pennsylvania. The flyers that were distributed are shown in Appendix B. In addition, the investigator requested permission from Indiana University of Pennsylvania faculty members to recruit participants from their classes, by reading a description of the study. The scripts that were used for recruitment of participants are provided in Appendix C.

3.2.3 Screening Measures

3.2.3.1 Exclusionary Criteria Once potential participants were recruited, a number of exclusionary criteria applied. First, based on a questionnaire, they were excluded if they reported a native language other than American English, or reported themselves to be fluent in a language other than English. This was because the language experience of these individuals, and therefore the predictions that they would make regarding upcoming syntactic structures, might have varied meaningfully from those with a monolingual American English background. In addition,

individuals were excluded if they indicated a history of language and/or other learning disability or traumatic brain injury (TBI). The reason to exclude such individuals was to control for the possibility that they may respond differently to the experimental items than those with a normal language and learning history. The final item on the questionnaire addressed whether the participant had any physical limitations (e.g. weakness) that might have affected their ability to press a button with their dominant hand. The questionnaire addressing each of these issues is provided in Appendix D.

Potential participants were screened for visual acuity using a modified version of the Snellen Chart (Snellen, 1862). From a computer screen, they were asked to read lines of letters taken from the Snellen Chart at an equivalent font size and distance, and under the same lighting conditions, as was used during the experiment. To be included for participation in the study, they were required to read accurately all of the letters from each line of the Snellen Chart. The vision criterion excluded participants whose vision may have interfered with their ability to read the sentences used in the experimental task.

The Slosson Oral Reading Test -3rd Edition (Slosson-R3; Slosson & Nicholson, 2002) was used to screen participants for reading ability. This assessment required participants to read sets of 20 isolated words that had been graded for reading level. While basal and ceiling rules dictated the number of words that were read by each participant, no participant was required to read more than 100 words. Participants scoring below the 10th grade level on this measure were excluded from participation in this study.

The Slosson-R3 (Slosson & Nicholson, 2002) was selected because it has been nationally standardized on over 1,000 participants across 30 states. In addition, it offers concurrent validity with the two more comprehensive measures of reading ability, the Reading Comprehension

subtest of the Peabody Individual Achievement Test –Revised/Normative Update (PIAT-R/NU; Markwardt, 1997) and the Passage Comprehension subtest of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001) of .90 or higher (Slosson et al., 2002).

The tenth grade reading level was selected as the criterion because the text used in the experimental conditions did not exceed the eighth grade reading level. This two grade-level difference between the participants' minimum reading level and the maximum reading level required for the text ensured that participants did not experience difficulties with the experimental materials due to working at the upper limits of their reading ability.

Although the Slosson-R3 (Slosson & Nicholson, 2002) correlates well with more comprehensive measures of reading achievement, it is a direct measure of only oral decoding skill. It provides no direct evidence of the individual's comprehension of the target words. For this reason, and because there is strong evidence for a relationship between vocabulary and reading ability (Nation & Snowling, 1999; Perfetti, Landi, & Oakhill, 2005), participants were also administered the Peabody Picture Vocabulary Test – 4th Edition (PPVT-4; Dunn & Dunn, 2007). This measure was chosen because it samples words representing various parts of speech (including verbs) across various difficulty levels. In addition, it was standardized on over 4,000 participants, and the norm sample reflects the 2004 United States Census data in terms of gender, race/ethnicity, socioeconomic status, and geographic region. Finally, it demonstrates internal consistency (split-half) reliability of .94, and test-retest reliability of .93 (PPVT-4 technical manual; Dunn & Dunn, 2007). Participants scoring more than one standard deviation below the mean on this measure were excluded from participation in the study.

3.2.3.2 Handedness Inventory All qualifying participants were tested for handedness using a modified version of the Edinburgh Handedness Inventory (Oldfield, 1971). This inventory requires participants to provide responses to a variety of items regarding their preferred hand for use in 10 different activities, and whether they ever use the other hand for these activities. The modified version included all of the items from the original inventory, but presented them on a computer screen, rather than using a pencil-and-paper format. The inventory was presented under the same font size and lighting conditions as was used for the experimental task. Participants used a computer mouse to mark their responses to each of the items.

3.3 STIMULI AND PROCEDURES

3.3.1 Stimuli

3.3.1.1 Verb Selection Criteria For this experiment, verbs were chosen that are associated with no more than one predominant meaning. The New Oxford American Dictionary (2nd edition; 2005) was consulted and two criteria applied for determining whether the verb is associated with a single meaning. First, as suggested by Rodd, Gaskell, and Marslen-Wilson (2002), a word with a single meaning should have no more than one distinct entry, but may have multiple alternatives (senses) within a single entry. Second, the verbs chosen for study have no more than one etymological origin, as indicated within its entry in The New Oxford American Dictionary (2nd edition; 2005).

In addition, the HAL (Hyperspace Analogue to Language) frequency counts (Balota, Cortese, Hutchinson, Neely, Nelson, Simpson, & Treimann, 2002; Lund & Burgess, 1996) were consulted in order to determine the frequencies of all candidate verbs (across senses). The HAL corpus consists of approximately 131 million words obtained from 3,000 Usenet newsgroups. Verbs that occurred extremely infrequently within this corpus (those with a frequency of less than 200 occurrences per 1,000,000 words) were not used in the experiment. The reason for this criterion was to reduce processing difficulty associated with reading words that are extremely infrequent, and perhaps unknown to participants. In addition, all of the verbs in the study ranged from one to three syllables in length.

Finally, verbs that rarely or never occur without a sentential complementizer, according to the data from three research studies (Garnsey et. al., 1997; Jennings et al., 1997; Trueswell et al., 1993), were also eliminated from consideration. Verbs were eliminated if, according to any one study, more than 80% of their associated sentential complements are preceded by a complementizer.

However, for some verbs, data were not available regarding the proportion of their associated sentential complements that were preceded by complementizers. For these verbs, the Internet Frequency Estimate Procedure (Nixon, 2006), which is described in more detail in the next section, was used to determine this proportion. This procedure involved locating instances of the verb using Internet searches, and tallying each instance in which the verb's associated sentential complement did or did not occur with a complementizer. Verbs were eliminated from consideration when more than 80% of their associated sentential complements were preceded by complementizers.

These verbs were eliminated because the target sentence in which they were to be used did not include a sentential complementizer. Therefore, eliminating these verbs reduced the possibility that the expectation for a complementizer would influence participants' reading times.

3.3.1.2 Determining subcategorization bias for verb senses All of the verbs selected subcategorized for both a noun phrase (NP) and a sentential complement (S). Because there is evidence that the various senses associated with individual verbs vary meaningfully in terms of their bias (Hare et al., 2004), a preliminary study was conducted in which each verb sense was classified as equi-biased, S-biased, NP-biased, or other.

A modified version of the Internet Frequency Estimate Procedure (Nixon, 2006) was used to determine the subcategorization bias of each verb sense. This procedure was conducted as follows. For each verb that met the aforementioned criteria for inclusion in the study, each verb sense was determined by consulting The New American Oxford Dictionary (2nd edition; 2005). Each numbered alternative within each dictionary entry was taken as a separate verb sense.

An independent rater who frequently conducts Internet searches was provided, in writing, with each verb and definitions of the particular senses of the verbs. The rater was be asked to provide up to 10 words that could be used as search terms for a web search for each verb sense. The definitions were presented to the rater quasi-randomly, with the restriction that no two senses of the same verb occurred consecutively.

The web searches were conducted using Google. For each web search, the verb and the search terms selected by the rater were entered. Boolean searches were conducted, using conjunction "or" between the search terms, so that the "hits" included at least the verb and one of the search terms.

Because websites located by Google are listed in order of relevance, the first 50 websites on the list of those retrieved were selected. For each of the selected websites, instances of the particular sense of the verb were located. Other senses of the same verb (or other target verbs) found on the website were ignored.

Websites for which there was reason to believe that the proportion of subcategorizations following verbs was not representative of the proportion of their use within the language were excluded from consideration. To replace these websites, additional websites were selected in the order in which they were retrieved. Similarly, certain instances of verbs were also not included in the analysis. The types of websites and verb tokens that were excluded, and the rationale for not including them are shown in Tables 13 and 14 (Appendix F).

If less than 100 instances of a particular verb sense occurred across the 50 websites, additional websites were selected in the order in which they appeared. Additional websites were selected, 50 at a time, until a minimum of 100 instances of a particular verb were found. This process was continued until a minimum of 100 instances was located across no more than 250 websites. Verb senses for which 100 instances could not be found across 250 websites were deemed "very infrequent" and were eliminated from consideration. This was done to eliminate verb senses with which a participant may have had limited prior experience, as well as to eliminate error (in determining verb sense bias) due to a limited sample.

The minimum of 100 instances was based on methods used by Hare et al. (2004) and Roland and Jurafsky (2002). In both studies, verb bias was determined, based on generally no more than 100 instances of any particular verb (across senses). Based on this, it was concluded that examining 100 instances of each sense of each verb would be sufficient to determine their bias.

For each located instance of a particular verb sense, the subcategorization following each verb was classified as: NP, S, prepositional phrase (PP), infinitive "to", zero subcategorization, or other. The total number of each verb subcategorization type following each verb sense was tallied. The data for the verbs included in the study are shown in Appendix E.

Using the criteria described in the Table 15, each of the remaining verb senses was classified as: NP-biased, S-biased, Equi-biased, or Other, based on the percentage of verb subcategorizations of a particular type that follow the verb sense. All verb senses that are classified as "Other" were not considered for inclusion.

Table 15: Criteria for classifying verb senses according to their subcategorization bias (based on Garnsey et al. (1997) and Trueswell et al. (1993)).

Classification	**Criteria**
NP-biased	NP is most prevalent completion associated with verb sense % NP completions - % S completions > 25%
S-biased	S is most prevalent completion associated with verb sense. % S completions - % NP completions > 25%
Equi-biased	NP and S are two most prevalent completion associated with verb sense |% NP completions - % S completions | ≤ 15 %
Other	Most prevalent completion type associated with verb sense is not NP or S. OR 15% < |% NP completions - % S completions | < 25%

Five NP-biased, five S-biased, and six equi-biased verb senses were selected for use in the target sentences of the experimental items. An equal number of verb senses of each of the included bias categories (equi-biased, S-biased, and NP-biased) appeared each condition, and within each stimulus set.

There are several distinct advantages of using this procedure over traditional procedures for determining subcategorization bias. As with other corpus-based approaches, it eliminates the concern about the relatively poor correlations between the subcategorization frequencies

obtained from isolated norming experiments and those found in spontaneous corpora (Lapata et al., 2001; Merlo, 1994).

In addition, the Internet websites include discourse contexts of various purposes and levels of formality, factors that may meaningfully affect subcategorization frequency (Biber, 1998; Roland, 2001; Roland & Jurafsky, 2002). In Biber (1998), the frequently-used corpora that were surveyed (e.g., the London-Lund corpus) reflected only a narrow range in terms of formality and purpose.

3.3.1.3 Target sentences for experimental items Each selected sense of each verb served as the matrix verb in one target sentence. The *target sentence* was the sentence for which self-paced reading measures were taken during the experimental trials. Each target sentence consisted of an NP, the matrix verb, and a sentential complement. The sentential complement occurred without a complementizer, such as *that* or *who*, in order to keep the sentence temporarily ambiguous between NP and S readings. In addition, the experimental target sentences ranged from 9 to 14 words in length.

Reading times across a disambiguating and post-disambiguating region of the target sentence were the dependent measures, because reading times (first pass or total reading times) across these regions have been the most affected by verb characteristics in other self-paced reading studies (e.g., Ferriera & Henderson, 1990; Trueswell et al, 1993). The *disambiguating region* is the region during which the temporarily ambiguous sentence resolves to one of its possible interpretations. In the current experiment, the disambiguating region was represented by the embedded verb. The *post-disambiguating region* is the region following the disambiguating region, during which the resolution of the ambiguity may continue to influence reading times.

For this study, the post-disambiguating region contained the following word, or two words, if the word immediately after the embedded verb is a function word (such as an article, preposition, or pronoun). This division of the sentence into regions is similar to the one used by Ferreira and Henderson (1990). These regions are illustrated in the example below, from these authors.

The judge believed the attorney | *made* | *a mistake.*

Disambiguating region | Post-disambiguating region

3.3.1.4 Narrative texts

Number of Experimental items A total of 64 experimental items (16 in each condition) were used in the current experiment. Sixteen of these items were quasi-randomly assigned to each of four stimulus sets. An *a priori* power analysis for a two-way within-subjects ANOVA indicated that 9 items in each condition (cell) were needed to obtain an effect size (Cohen's *d*) of 0.5 with a power of 0.8 and alpha set at .05 (Cohen, 1988) to detect both main effects and interactions. As with the determination of participant sample size, the number of items needed was determined through the use of a formula provided by Cohen (1988; p. 396). An additional seven items were added to ensure adequate power, even if data were lost due to outliers, equipment failure, etc.

Versions of narrative texts: Experimental items For the experimental items, each target sentence followed each of four versions of one narrative text. Each narrative text (N = 16) was composed of eight sentences. For each text, four different versions were created. These were defined by the number of matrix verbs within the text that were followed by sentential complements (S), and by whether the text was a coherent discourse context or a series of scrambled sentences. These versions were: 1) a discourse context, six S version, 2) a discourse

context, two S version, 3) a scrambled sentence, six S version, and 4) a scrambled sentence, two S version.

Text versions 1 and 2 (coherent discourse) introduced a character who was presented with a statement or situation and expected to react with an opinion or with a statement of belief or knowledge. This was intended to increase readers' expectations for an S to follow the matrix verb of the target sentence, in order to describe the character's reaction to the statement or situation. The experimental items are shown in Appendix H.

Three independent raters read each coherent discourse version of each text, and indicated whether they felt that the character's reaction to the situation reflected a "typical" or "logical" reaction. Any texts that were indicated as being illogical by any of the raters were rewritten. This was done to reduce the possibility that participants' reading times would be influenced by potentially illogical responses of the main character.

Text versions 3 and 4 (scrambled sentences) included the same first and last sentences as versions 1 and 2. However, the intervening six sentences were quasi-randomly selected from different coherent texts, with no more than one sentence from the same text. This was done in order to reduce the likelihood that a global representation will be generated by the participant.

In addition, efforts were also made to reduce other (confounding) factors that might have caused participants to respond differently in the discourse context and scrambled sentence conditions. First, the last sentence to resolve to an S was the same distance (in number of intervening sentences) from the target sentence for all four versions of the text. This was done to control for any potential influence of recency on the processing of the sentential complement in the target sentence. In addition, the first and last sentences of all four versions of each text were the same. This was done to reduce the possibility that readers might have responded differently

to passages that did not begin in a manner that is typical for a narrative text, and to reduce the possibility that local (rather than discourse-level) factors might have influenced participants' reading times for the target sentences.

The four versions of each text also included two different syntactic priming conditions. For two versions of the text, six of the eight main verbs (75%) resolved to an S, and two resolved to an NP. For the other two versions, two of the eight main verbs (25%) resolved to an S, and six resolved to an NP.

Four stimulus sets were created, with one version of each text included in each stimulus set. Therefore, there was a total of 16 experimental items in each stimulus set. Within each stimulus set, the trials for the 75% and 25% conditions were presented alternately, to reduce the potential for cumulative syntactic priming across trials. An example of an experimental items (text and target sentences) that was used in this experiment is provided below.

Example: Discourse context/ 75% S, 25% NP condition.

> As he returned from skiing with friends, Don discovered that one of his skis was missing. Although it was getting dark and snowing heavily, he decided he should go back and find the ski before it was covered with snow. His friends warned him that it would be dangerous to leave the cabin so late in the day. They promised they would help him look for his ski in the morning. However, Don demanded they let him go. Walking through the snow, Don noticed a shape on the hillside that he thought was his ski. However, as he approached the item, he noticed that it was just a large stick. Turning around, he doubted he could find his way back.
>
> Target: Don realized his mistake had been to ignore his friends' advice.

Verbs used in experimental narrative texts All of the verbs used in the main/matrix clauses in all four versions of each experimental narrative text subcategorized for both NP and S. Using such verbs was intended to reduce variability in experience across the stimulus item, such as that observed by Kashak and Glenberg (2004; Experiment 3). As described previously, these authors required participants to read sentences that were temporarily ambiguous between a *needs* construction and a *modifier* construction, and which always disambiguated to the *needs* construction. Despite their lack of direct exposure to the *modifier* construction, participants' subsequent processing of this construction was facilitated. This suggests that the reading time differences that might be attributed to the frequency of a particular subcategorization in the preceding sentences might also be influenced by the other subcategorizations associated with the verbs that are encountered. This is because the participant might have predicted the alternative subcategorization at some point during sentence processing.

In constructing the texts, the frequency with which particular verbs appear with a complementizer was considered. For verbs that typically occur with a sentential complementizer more than 80% of the time (Garnsey et. al., 1997; Jennings et al., 1997; Trueswell et al., 1993), complementizers preceded any sentential complement that followed them, in order to reduce this potential source of processing difficulty. However, in order to keep the process of disambiguation similar for sentences with and without complementizers, the number of post-verbal NPs preceded by the demonstrative adjective *that* were equal to the number of post-verbal sentential complements preceded by the complementizer *that* within each stimulus set.

Number of filler items In addition to the experimental items, 24 filler items were created, in which the narrative text ranged in length from five to twelve sentences. The decision to include 50% more filler items than experimental items was based on the methods used in similar

experimental studies (studies of the influence syntactic priming or discourse context) in which a significant influence of the examined variable was found. Studies in which the experimental items from one experiment served as filler items in another experiment were not included in this survey of methods, because the number of filler items may have been chosen for practical, rather than theoretical reasons.

In the examined studies of syntactic priming (Hartsuiker & Kolk, 1998; Kashak & Glenberg, 2004; Cleland & Pickering, 2006; Potter & Lombardi, 1998), it was determined that the ratio of the number of filler items to the number experimental items ranged from 1.5:1 to 4:1. For the studies of the influence of discourse context (Hess et al., 1995; Schwanenflugel & White, 1991; Spivey & Tanenhaus, 1998), this ratio ranged from 0.6:1 to 3.125:1. One major reason why these investigators stated that they included a large number of filler items was so that the participants would remain unaware of the purpose of the experiment and/or the type of construction under investigation.

Because half of the experimental stimuli used in the current study were coherent paragraphs, this suggested that the participants would be less likely to attend to a particular syntactic construction than may have been the case in previous studies of syntactic priming in which unrelated sentences served as stimuli. For this reason, a ratio in the low- to mid- range of the values found in prior studies (rather than a value on the higher end of this range) was judged to be appropriate. There were a total of 24 filler items, 50% more than the number of experimental items.

Development of filler items Like the experimental items, half of the filler items constituted a meaningful discourse context, and half were comprised of a series of scrambled sentences. As for the experimental items, the scrambled sentence versions were created by replacing all but the

first and last sentence of coherent discourse contexts with sentences that were quasi-randomly selected from other coherent discourse contexts. Each sentence (except for the first and last) was selected from a different discourse context.

For these trials, the syntactic constructions of the majority of sentences in main text (at least 80%) differed from those used for the experimental items. A small proportion (20%) of sentences with the target constructions (verb + NP or verb + sentential complement) was used in the text of each of the filler trials. This was done to reduce obvious differences between the experimental and filler trials. A variety of other subcategorizations and combinations thereof (PP, Ø, PP NP) were used in the filler trials.

As in the experimental items, some of the filler texts (20%) introduced a statement or situation to which a main character was expected to react. However, unlike the experimental items, the reaction of the character was stated within the text, and therefore was not expected to occur in the target sentence. This allowed for greater variability in the syntactic construction of the target sentences of these trials. The content of the remaining 80% of the filler texts varied from this format, to prevent the participants from anticipating the content of the stimuli before they were presented.

The target sentences of the filler trials used a different syntactic construction from those the used in the experimental trials (NP + verb + sentential complement). These target sentences were compound, complex, or compound-complex, in terms of their syntactic construction and ranged from 9 to 14 words in length.

The filler items were placed quasi-randomly throughout each stimulus set, with the constraint that no stimulus set would have more than two consecutive experimental items. An example of a filler item that was used in this experiment is provided below.

One day, a woman was shoveling snow in her driveway, when her favorite pin fell into the snow. Before she realized what had happened, it became buried under the snow. Despite an hour of searching, she was unable to find it. That night, she explained to her husband that she had lost the pin. A few days later, he was able to find the pin in the snow, but it was badly damaged. He took it to a jeweler to have it repaired. Meanwhile, the woman, unaware of her husband's discovery, decided to buy a new pin, just like the one she had lost. Wearing her new pin home, she soon found, beside a note from her husband, a badly repaired version of the same pin. The old pin was in poor condition, and she hoped that her husband would not be disappointed if she wore the new pin instead. However, after reading the note, she decided to return the new pin to the store.

Target sentence: She didn't want to make him feel that his effort was unappreciated.

Practice trials Prior to beginning the experimental task, participants completed a series of six practice trials. These practice items were similar to the items used for the filler trials. They consisted of a narrative text and a target sentence. The narrative texts ranged from 5 to 12 sentences in length. Three of these trials comprised a coherent narrative text, and the other three consisted of a series of scrambled sentences. Like the filler trials, the target sentences following each of these trials differed in their syntactic construction from those used in the experimental trials.

3.3.1.5 Comprehension questions Comprehension questions were constructed for one-third of all experimental and filler trials. These trials were selected quasi-randomly, with the constraint

that no more than two consecutive trials were followed by a comprehension question. In addition, comprehension questions followed four (two-thirds) of the six practice trials. The larger proportion of comprehension questions following practice trials was intended to allow participants to have additional practice, as well as to receive feedback, for this portion of the task.

All of the questions that followed the scrambled sentence text versions probed details that were explicitly stated in individual sentences of the text, rather than requiring integration of information across sentences in the texts, because these trials did not comprise coherent discourse contexts. In contrast, one-third of the questions following the discourse context versions probed details from individual sentences; the remaining two-thirds required integration of information across sentences in the text. This proportion allowed the experimenter to determine whether a participant comprehended the discourse-level information on these trials. Using both questions that probed detail-level information and those requiring integration of information across sentences for the discourse context versions reduced any obvious distinction in question type used following the discourse context and scrambled sentence conditions.

The questions that were intended to probe details addressed information presented in both early and late portions of the text. Twenty-five percent of these probed information presented in the first half (i.e., first five sentences) of the text; seventy-five percent probed information in the second half of the text. The reason for including more questions about the second half of the text was to encourage participants to focus on the content of the text until the end.

Each question required a yes/no response, registered by a button-press. An example comprehension question (for the example experimental item) is provided below.

Comprehension Question: Did Don's friends agree with his decision?

The accuracy of responses to these items was measured only to determine whether the participants were processing the texts throughout the experiment. This measure was not used in any statistical analyses. However, if a participant answered the questions following the coherent discourse contexts with less than 80% accuracy, their data were excluded from the analyses. The criterion of 80% accuracy was selected based on a the results of a study (Nicholas & Brookshire, 1995), which examined the response accuracy of adults with and without brain damage to questions regarding texts of similar length and reading level to the ones used in the current study. The performance of non-brain damaged adults ranged from 95-100% accuracy for questions probing implied main ideas, and from 80-100% for questions probing stated details.

The current study included questions probing both main ideas and details. However, participants in the current study were required to respond to the comprehension questions while completing a novel task that required them to read both discourse contexts and scrambled sentences. The unusual nature of this task might have been detrimental to their performance in responding to the comprehension questions. Thus, the 80% criterion was selected, as it represents the lower end of the performance range of non-brain damaged individuals on detail-level questions.

3.3.2 Procedures

All procedures occurred at Indiana University of Pennsylvania. The screening and experimental procedures were described to the participants, and an informed consent document was discussed and signed.

3.3.2.1 Screening procedures First, each participant filled out a brief background history questionnaire (see Appendix D). Second, each participant was screened for vision, reading ability, receptive vocabulary, and handedness, using the procedures described in the Participants section (section 3.2) of this paper. After the screening procedures had been completed, a random numbers table was used to quasi-randomly assign qualifying participants to one of the four stimulus sets with the condition that an equal number of participants (10) were assigned to each stimulus set. If a participant was randomly assigned to a stimulus set that had been filled (that had already been completed by 10 participants), a random numbers table was used to reassign that person to another set.

3.3.2.2 Experimental apparatus One version of each experimental text was assigned to each stimulus set in a quasi-random manner, with the condition that an equal number of discourse context and scrambled sentences text versions occurred within each stimulus set. This resulted in 16 experimental items in each stimulus set. In addition, the same 24 filler items were used for all four stimulus sets. The experimental task was presented on a Toshiba Satellite laptop computer using the E-prime software program (Schneider, Eschman, & Zuccolotto, 2002). All of the stimuli were presented in black, 24-point Times New Roman font on a white background. All sentences from the narrative texts and all target sentences were presented half-way between the top and bottom of the computer screen.

Each trial (including the practice trials) began with a black screen with white text, which said: “Ready? Press the space bar to begin.” The participants initiated the first sentence of the text, and each successive sentence, by pressing the space bar. Each stimulus sentence was presented in black font on a white background. As each sentence appeared, the previous sentence

disappeared. One reason for preventing participants from viewing more than one sentence at a time was that this would have made it less likely that they would consciously take note of the proportion of the types of subcategorizations used in the experimental items. In addition, it ensured that the sentences were read in the intended order, and that participants did not reread any previous sentences, because rereading earlier sentences may have caused unintended recency effects.

When the participant pressed the space bar after the final sentence of a stimulus text, a series of dashes appeared on the screen, one representing each word in the target sentence. With an additional press of the space bar, the first word of the target sentence appeared. At this point, each successive key press initiated the presentation of each word. As each word was presented, the previous word was replaced by dashes. Presenting the target sentence in this manner allowed for the use of self-paced reading as a measure of reading time, while retaining the left-to-right reading pattern that is used for English text.

Following one-third of all trials, participants were presented with the comprehension questions. The comprehension questions were presented on a white background, half-way between the top and bottom of the computer screen in black, 24-point Times New Roman font. Each question remained on the screen until the participant responded.

Participants answered by pressing one of two buttons on the computer keyboard that were labeled “yes” and “no” (A= yes; L = no). These two keys were chosen because they are placed relatively far apart on the keyboard. Choosing widely spaced keys was intended to reduce the likelihood that participants would press the wrong key in error. In addition, these keys are in the center row of the keyboard, and near the end of the row. It was believed that this would help the participants to locate these keys quickly and easily.

3.3.2.3 Experimental procedures Before beginning the experimental task, participants completed six practice trials in order to familiarize them with the task, and with the sentence-by-sentence and word-by-word modes of presentation. Comprehension questions followed four of the six practice trials. Before beginning the practice trials, the participants were given the following instructions.

> You are going to read several sets of sentences, which will range in length from 5-12 sentences. Some of the sets of sentences will tell a coherent story, but others will not. Also, after some of the sets of sentences, there will be questions for you to answer, but you will not know ahead of time which sets of sentences will be followed by a question.
>
> As you read the sets of sentences, you will see only one sentence at a time. It is important that you read each sentence as quickly as possible, because I will be measuring your reading speed, but I would also like to you read carefully enough that you can answer any questions that appear. I would like you to begin with the index finger of your (right or left) hand on the space bar. After you read each sentence, press the space bar and the next sentence will appear. When you come to the final sentence, you will see only one word each time you press the space bar, rather than an entire sentence. Please move through the words as quickly as you can. If a question follows the item, you will respond by pressing either the "yes" or "no" key to answer the question. I have been telling you that I want you to read fast, but you also need to read carefully enough that you can answer the questions when they appear. Do you have any questions before we practice?
>
> First, let's start by doing some practice sets, so that you can get used to the way the sentences will be presented.
>
> [Completed first practice sentence set]
> *[A small number of participants paused for a second or more after reading the last word in at least one of the sentences within the text, or after reading a word in the target sentences. These participants were given the following instructions, "It looks that you have the basic idea, but I want you press the button even more quickly after you read each sentence or word."*
>
> *[If the participant inaccurately answered the comprehension questions, he or she was given the following instructions, "You're doing a good job of reading through the sentences, but I need you focus more closely on the content of the sentences, so that you can answer the questions."]*
>
> Now that you have the idea, I would like you to complete a few more practice sets on your own before you do the real thing.
>
> [Completed the remaining five practice sets].

> Okay, I think you're ready to start the actual experiment. There will be 40 sets of sentences. As you did with the practice sets, press the space bar as soon as you finish reading each sentence or word, but read carefully, so that you can answer the questions when they appear.

3.3.2.4 Session organization and testing environment All of the screening and experimental procedures were conducted in one, 90-minute session in a quiet, well-lit room. For the screening and experimental tasks, participants were seated at a desk or table. For the experimental task, they were seated approximately 18 inches from the computer screen. The distance of the laptop from the floor was individually adjusted for each participant, so that the middle of the screen was at eye-level for each participant. All participants responded to the experimental task with their dominant hand (as determined by the Edinburgh Handedness Inventory; Oldfield, 1971). They were asked to begin the experimental task with the index finger of their dominant hand resting against the space bar, and to press the space bar with this finger to progress through the sentences and words during the experimental task.

The session began with reviewing and signing consent forms and completing the brief history questionnaire. This was followed by the vision, reading, and vocabulary screenings. If the participant passed all of the screenings, and did not indicate any exclusionary criteria on the history questionnaire, they completed a handedness questionnaire. Following this, experimental testing began.

4.0 STATISTICAL ANALYSES

The dependent measures in this study were reading times across disambiguating and post-disambiguating regions of the experimental sentences. Two, 2-way analyses of variance (ANOVAs) were used to compare these reading times across two context conditions (discourse context and scrambled sentence) and two frequency conditions (2S and 6S). Separate analyses were conducted for data from the disambiguating and post-disambiguating regions. Each analysis was conducted with an α level of .05. A one-tailed test was conducted, because the predictions for both main effects and for the interaction were directional as well.

The results of two self-paced reading studies (Ferreira & Henderson, 1990; Trueswell et al., 1993) suggest that verb subcategorization preference may influence reading times in either the disambiguating or post-disambiguating region. Because the present study addressed how these expectations may be influenced by discourse context or syntactic priming, reading times across these regions were targeted for statistical analysis. Findings for an influence of discourse context in either region would have implications for supporting or refuting the previously described viewpoints regarding the time course over which its influence is believed to occur.

According to the interactive viewpoint on language processing, discourse context would be predicted to have an initial influence on sentence parsing. This suggests that it would likely influence reading times across the disambiguating region (Altmann & Steedman,1988; Britt et al, 1992; Grodner, Gibson, & Watson, 2005; Hess et al., 1995; Schwanenflugel & White, 1991)

According to the autonomous viewpoint, however, non-syntactic information, such as discourse context information would not influence the initial processing of sentences, but would be influential primarily at a later stage of processing. Although the autonomous viewpoint could account for the influence of discourse context in either region (i.e., it could explain earlier occurring non-syntactic influences as being related to a reanalysis stage of processing), the fact that the disambiguating region is short (only one or two words in length) in the current experiment suggests that any later-occurring non-syntactic influences would be more likely to appear in the post-disambiguating region. As mentioned in the review of the literature, the possibility for a later-occurring influence of discourse context on sentence parsing is supported by the results of both Boland and Blogett (2001) and Cook and Myers (2004; Experiment 1). In addition, Kintsch's CI model (1988; 1998) also predicts a later-occurring influence of discourse during the Integration stage of processing.

The predictions for the influence of syntactic priming are less controversial. The two accounts of syntactic priming described earlier, the framing assumption (Bock, 1986; Bock and Loebell, 1990) and the syntactic representation account (Pickering and Branigan, 1999), both suggest that syntactic priming has an initial influence on sentence processing. This is supported by the previously described findings of Kashak and Glenberg (2004; Experiments 1-4).

In addition, both autonomous and interactive viewpoints on sentence processing suggest that syntax plays an early role in the process of sentence parsing. This strongly suggests that the influence of syntactic priming would likely occur in the disambiguating region, rather than in the post-disambiguating region.

Because it was predicted that the influence of syntactic priming would be seen in the disambiguating region only, and that the influence of discourse context might occur in either the

disambiguating or post-disambiguating region, this suggested that any interaction between these influences would be restricted to the disambiguating region. For either region, a significant main effect of frequency in the absence of other significant effects would suggest that cumulative syntactic priming exerted an independent influence on predicted verb subcategorizations. Similarly, a significant main effect of discourse context in the absence of other significant effects would suggest that discourse context exerted an independent influence on predicted verb subcategorizations. Finally, a significant interaction between frequency and discourse context would suggest that the cumulative syntactic priming exerts a significantly different degree of influence after a coherent discourse context than after a series of unrelated sentences.

5.0 RESULTS

5.1 EXCLUDED AND QUALIFYING PARTICIPANTS

A total of 44 individuals were recruited for participation in this study. Individuals were excluded from participation if they reported a native language other than English, were fluent in more than one language, or reported a history of language and/or learning difficulties. Three individuals were excluded because they reported a native language other than English; one was excluded due to a history of learning disability.

No participants were excluded for failure to answer the comprehension questions with at least 80% accuracy. Similarly, no individuals were excluded because they did not pass the reading, vocabulary, and vision screenings.

The forty individuals (20 males and 20 females) who qualified for participation ranged in age from 20 to 34 years (mean = 22;5, SD = 3.0 years). Their educational levels (e.g., 12 = high school graduate; 16 = bachelor's degree) ranged from 13 to 16 (mean = 15.07, SD = .85). Participants were not compensated for participation in this study. The majority of participants were university undergraduate and graduate students who were recruited toward the end of finals week of fall semester, and during the next week. It is likely that they were willing to participate without compensation, because they had free time between their final exams and returning home for the holidays.

5.2 PRIMARY ANALYSES

5.2.1 Analyses based on raw data

Tables 16 and 17 show the descriptive statistics for the participants' reading times by coherence condition (discourse context and scrambled sentences) and frequency condition (6S and 2S) for both the disambiguating and postdisambiguating regions of the target sentences.

Table 16: Reading times (ms) for the disambiguating region.

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	508.23	272.37	152-1757
6S/discourse context	553.52	342.38	150-1904
2S/scrambled sentences	559.31	298.42	146-1797
6S/scrambled sentences	572.56	301.41	148-1942

Table 17: Reading times (ms) for the postdisambiguating region.

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	675.28	572.94	138-6880
6S/discourse context	735.34	507.83	149-5265
2S/scrambled sentences	741.46	469.32	174-4286
6S/scrambled sentences	690.57	297.09	169-1728

The main effects of frequency (6S vs 2S) and coherence (coherent discourse context vs scrambled sentences) and the interaction between these factors across the disambiguating and postdisambiguating regions were examined using two, 2 (coherent discourse context vs. scrambled sentence) x 2 (6S vs. 2S) two-way analyses of variance (ANOVAs). For the disambiguating region, this did not yield a significant main effect for coherence either by participants, $F(1,39) = 1.558$ or by items, $F(1, 15) = 3.219$, $p > .05$. In addition, there was no

main effect of frequency either by participants, $F(1,39) < 1$, or by items, $F(1,15) = 1.788$, $p > .05$, suggesting that neither factor significantly influenced participants' reading times. Finally, there was no significant interaction between these factors, either by participants, $F(1,39) < 1$ or by items, $F(1,15) < 1$.

Similarly, for the postdisambiguating region, there was no significant main effect of coherence either by participants, $F(1,39) < 1$ or by items, $F(1,15) < 1$. There was also no significant main effect for frequency either by participants, $F(1,39) < 1$ or by items, $F(1,15) < 1$. There was no significant interaction between coherence and frequency by participants $F(1,39) = 1.328$, $p > .05$, although the interaction was significant by items, $F(1, 15) = 6.122$, $p < .05$. For the *by items* analysis, reading times following the scrambled sentence conditions were faster in the 6S condition than in the 2S condition. Conversely, reading times following the coherent discourse context conditions were faster in the 2S condition than following the 6S condition (Tables 18 and 19).

As shown in Tables 18 and 19, the power of the experiment to detect significant main effects and interactions was low for all comparisons for the disambiguating regions (ranging from .110 to .390), and low to medium for the comparison for the postdisambiguating regions (ranging from .052 to .638).

Table 18: Analysis of variance results by items and participants for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	22449.829	22449.829	.677	.416	.017	.126
Coherence	1	52789.307	52789.307	1.558	.219	.038	.230
Freq * Coherence	1	10316.141	10316.141	.646	.427	.016	.123
Error (freq)	39	1293857.687	33175.838				
Error (coherence)	39	1321348.459	33880.730				
Error (freq*coherence)	39	623064.375	15976.010				
By items (N = 16)							
Frequency	1	13569.338	13569.338	1.788	.201	.106	.240
Coherence	1	19498.631	19498.631	3.219	.093	.177	.390
Freq * Coherence	1	3929.723	3929.723	.578	.459	.037	.110
Error (freq)	15	113861.655	7590.777				
Error (coherence)	15	90867.151	6057.810				
Error (freq*coherence)	15	102005.080	6800.339				

Table 19: Analysis of variance results by items and participants for reading times across the postdisambiguating region.

Variable	*Df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	1255.800	1255.800	.024	.877	.001	.053
Coherence	1	3703.219	3703.219	.065	.800	.002	.057
Frequency* Coherence	1	127675.875	127675.875	1.328	.256	.033	.203
Error (frequency)	39	2022494.559	51858.835				
Error (coherence)	39	2225121.078	57054.387				
Error (frequency*coherence)	39	3750685.609	96171.426				
By items (N = 16)							
Frequency	1	350.626	350.626	.021	.888	.001	.052
Coherence	1	1799.881	1799.881	.101	.755	.007	.060
Frequency* Coherence	1	49417.290	49417.290	6.122*	.026	.290	.638
Error (frequency)	15	256296.409	17086.427				
Error (coherence)	15	266657.164	17777.144				
Error (frequency*coherence)	15	121075.055	8071.670				

Note. *p < .05

5.2.2 Analyses with outliers eliminated

No significant main effects were found based on the initial analyses. One potential factor that might have contributed to the lack of significant findings is the inclusion of outliers. According to Howell (1992), analyses of variance are particularly sensitive to the inclusion of outliers (p. 308). For this reason, the initial analyses were repeated with the outliers eliminated. A participant's reading time was considered to be an outlier if it was more than three standard deviations above or below the mean reading time for that condition.

The number and percentage of reading times that were eliminated in each condition and each target region, as well as the number of participants contributing outliers to each condition are shown in Table 20. Descriptive statistics for these data trimmed of outliers are shown in Tables 21 and 22.

Table 20: Number and percentage of outliers eliminated in each condition and for each target region, and the number of participants contributing outliers.

Condition	*Number of Outliers Eliminated*	*Percentage of Total Data Points*	*Number of Participants*
Disambiguating Region			
2S/DC	1	.625%	1
6S/DC	3	1.875%	2
2S/SS	4	2.5%	3
6S/SS	1	.625%	1
Postdisambiguating Region			
2S/DC	1	.625%	1
6S/DC	3	1.875%	3
2S/SS	3	1.875%	2
6S/SS	1	.625%	1

Table 21: Reading times (ms) for the disambiguating region (with outliers eliminated)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	505.85	251.52	152-1222
6S/discourse context	529.92	299.05	150-1439
2S/scrambled sentences	539.72	246.97	146-1211
6S/scrambled sentences	565.08	281.96	148-1517

Table 22: Reading times (ms) for the postdisambiguating region (with outliers eliminated)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	636.32	291.75	138-2202
6S/discourse context	688.32	309.97	149-2209
2S/scrambled sentences	694.20	301.29	174-1663
6S/scrambled sentences	684.04	286.30	169-1453

As in the initial analyses, the main effects for frequency and coherence and the interaction between these factors across the disambiguating and postdisambiguating regions were examined using two, 2 (coherent discourse context vs. scrambled sentence) x 2 (6S vs. 2S) two-way ANOVAs. For both the disambiguating and postdisambiguating regions, no significant results were found. However, the power of these analyses to detect a significant main effects and interactions was again low for comparisons involving both the disambiguating (ranging from .050 to .496) and postdisambiguating regions (ranging from .165 to .449. The results and power of these analyses are shown in Tables 23 and 24.

Table 23: Analysis of variance results (with outliers eliminated) by items and participants for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	29498.477	29498.477	1.006	.341	.025	.165
Coherence	1	46899.669	46899.669	1.836	.211	.045	.262
Frequency* Coherence	1	20.435	20.435	.001	.974	.000	.050
Error (frequency)	39	1143069.204	29309.467				
Error (coherence)	39	996107.011	25541.205				
Error (frequency*coherence)	39	657229.151	16852.030				
By items (N = 16)							
Frequency	1	20765.191	20756.191	3.837	.069	.204	.450
Coherence	1	15764.036	15764.036	4.347	.055	.225	.496
Frequency * Coherence	1	309.025	309.025	.066	.801	.004	.057
Error (frequency)	15	81146.955	5409.797				
Error (coherence)	15	54398.170	3626.545				
Error (frequency*coherence)	15	70123.800	4674.920				

Table 24: Analysis of variance results (with outliers eliminated) by items and participants for reading times across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	14428.919	14428.919	1.090	.303	.027	.175
Coherence	1	32075.704	32075.704	2.484	.123	.060	.337
Frequency * Coherence	1	52053.219	52053.219	1.316	.258	.033	.201
Error (frequency)	39	516304.253	13238.571				
Error (coherence)	39	503607.884	12913.023				
Error (frequency*coherence)	39	1542153.266	39542.391				
By items (N = 16)							
Frequency	1	7943.823	7943.823	1.091	.313	.068	.165
Coherence	1	9864.835	9864.835	1.106	.310	.069	.167
Frequency * Coherence	1	15711.317	15711.317	3.827	.069	.203	.449
Error (frequency)	15	109199.136	7279.942				
Error (coherence)	15	133791.688	8919.446				
Error (frequency*coherence)	15	61574.573	4104.972				

5.2.3 Analyses with proportion data

The estimates for required sample size that were used in the current experiment were based upon effect sizes from similar studies. It was possible that the lack of significant findings in the previously described analyses might have been, at least in part, due to greater variability in the participants' reading times in the current study than for those in the studies on which the sample size estimates were based. To reduce some of this variability, the analyses were conducted again with proportion scores as the dependent measure. To derive each proportion score, a participant's reading time across a target (disambiguating or postdisambiguating) region was divided by his or her reading time for the first sentence of the preceding sentence set. Reading times for the first sentence in each set were chosen as the baseline measures because reading times for these sentences would not be influenced by the coherence or frequency factors of the texts in which they occur.

As in the previous analyses, the first step was to eliminate any outliers. Any proportion score was considered to be an outlier, if it was more than three standard deviations from the mean proportion score for that condition. The number and percentage of proportion scores that were eliminated in each condition and each target region, as well as the number of participants who contributed outliers to each condition are shown in Table 25. The descriptive statistics for the outlier-trimmed data are shown in Tables 26 and 27.

Table 25: Number and percentage of outlying proportion scores eliminated in each condition and for each target region, and the number of participants contributing outliers.

Condition	*Number of Outliers Eliminated*	*Percentage of Total Data Points*	*Number of Participants*
Disambiguating Region			
2S/DC	2	1.25%	2
6S/DC	2	1.25%	2
2S/SS	2	1.25%	2
6S/SS	3	1.875%	3
Postdisambiguating Region			
2S/DC	3	1.875%	3
6S/DC	1	.625%	1
2S/SS	4	2.5%	4
6S/SS	3	1.875%	3

Table 26: Reading times (ms) for the disambiguating region (for proportion scores)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	.117	.053	.059-.247
6S/discourse context	.133	.072	.045-.276
2S/scrambled sentences	.133	.072	.045-.276
6S/scrambled sentences	.131	.061	.052-.270

Table 27: Reading times (ms) for the postdisambiguating region (for proportion scores)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	.151	.055	.068-.238
6S/discourse context	.177	.073	.067-.333
2S/scrambled sentences	.167	.076	.041-.353
6S/scrambled sentences	.158	.060	.073-.284

Using these proportion scores, the main effects of frequency and coherence and the interaction between these factors across the disambiguating and postdisambiguating regions were examined using two, 2 (coherent discourse context vs. scrambled sentence) x 2 (6S vs. 2S) two-way analyses of variance (ANOVA). For both the disambiguating and postdisambiguating regions, these analyses yielded no significant main effects or interactions. The results of these analyses, including power calculations, are shown in Tables 28 and 29.

Table 28: Analysis of variance results (for proportion scores) by items and participants for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	.001	.001	.368	.548	.009	.091
Coherence	1	.002	.002	.577	.452	.015	.115
Frequency * Coherence	1	.002	.002	2.648	.112	.064	.355
Error (frequency)	39	.121	.003				
Error (coherence)	39	.125	.003				
Error (frequency*coherence)	39	.034	.001				
By items (N = 16)							
Frequency	1	.001	.001	3.593	.077	.193	.426
Coherence	1	.001	.001	3.593	.077	.193	.426
Frequency * Coherence	1	.001	.001	2.064	.171	.121	.270
Error (frequency)	15	.003	.000				
Error (coherence)	15	.003	.000				
Error (frequency*coherence)	15	.010	.001				

Table 29: Analysis of variance results (for proportion scores) by items and participants for reading times across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	.003	.003	.822	.370	.021	.143
Coherence	1	.000	.000	.016	.901	.000	.052
Frequency* Coherence	1	.010	.010	3.607	.065	.085	.457
Error (frequency)	39	.148	.004				
Error (coherence)	39	.098	.003				
Error (frequency*coherence)	39	.109	.003				
By items (N = 16)							
Frequency	1	.001	.001	1.094	.312	.068	.165
Coherence	1	.000	.000	.044	.837	.003	.054
Frequency * Coherence	1	.004	.004	4.248	.057	.221	.488
Error (frequency)	15	.018	.001				
Error (coherence)	15	.016	.001				
Error (frequency*coherence)	15	.015	.001				

5.2.4 Transformation of Proportion Scores

One factor that might contribute to the lack of significant findings in the analyses with proportion scores would be the failure of the distribution of these scores to meet the normality and homogeneity of variance assumptions for analysis of variance. For this reason, these factors were examined. The results of Levene's Test for homogeneity of variance indicated that this assumption was met for the disambiguating region (p = .444), but not for the postdisambiguating region (p = .018). To determine whether the assumption of normality was met, the values for skewness and kurtosis were examined. These are shown in Table 30.

Table 30: Skewness and kurtosis values for distribution of proportion scores before transformation.

	Conditions	*Skewness*	*Std. Error*	*Kurtosis*	*Std. Error*
Disambiguating Region	2S/DC	.923	.564	.642	1.091
	6S/DC	.799	.564	-.217	1.091
	2S/SS	.867	.564	-.173	1.091
	6S/SS	.965	.564	.219	1.091
Postdisambiguating Region	2S/DC	.294	.564	-1.306	1.091
	6S/DC	.472	.564	-.197	1.091
	2S/SS	.791	.564	1.229	1.091
	6S/SS	.504	.564	-.489	1.091

For a perfectly normal distribution, both skewness and kurtosis values are zero. However, a distribution can be considered to be normal, if the skewness and kurtosis values are within the standard error. As shown in Table 30, the skewness values indicate a positively skewed distribution in five of the eight cases. However, the kurtosis values extend beyond the range of the standard error in only two cases. Because the most apparent threat to the normality assumption was the positive skewing, it was determined that a logarithmic (base 10) transformation would be the most useful in achieving a normal distribution (Howell, 1992). The results of this transformation on the normality of the proportion data distributions are shown in Table 31.

Table 31: Skewness and kurtosis values for distribution of proportion scores after transformation.

	Conditions	*Skewness*	*Std. Error*	*Kurtosis*	*Std. Error*
Disambiguating Region	2S/DC	.171	.564	-1.043	1.091
	6S/DC	-.026	.564	-.970	1.091
	2S/SS	.111	.564	-1.065	1.091
	6S/SS	.127	.564	.400	1.091
Postdisambiguating Region	2S/DC	-.180	.564	-.934	1.091
	6S/DC	-.362	.564	-.507	1.091
	2S/SS	-.859	.564	1.692	1.091
	6S/SS	-.119	.564	-.888	1.091

As indicated by the values in Table 31, this transformation resulted in all but one of the skewness values being within the standard error. In addition, the values for kurtosis indicated that only one of these values was now outside of the standard error. Finally, Levene's Test was conducted on the transformed data, and indicated that the assumption of homogeneity of variance was met for both the disambiguating region (p = .559) and postdisambiguating regions (p =.491). Because the homogeneity of variance assumption was met, and ANOVA is robust to the assumption of normality in this case, it was deemed to be appropriate to reanalyze these data using ANOVA. Tables 32 and 33 show the results of ANOVAs conducted on these transformed data for the disambiguating and postdisambiguating regions.

Table 32: Analysis of variance results (for transformed proportion scores) for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	.015	.015	.388	.537	.010	.093
Coherence	1	.013	.013	.382	.540	.010	.093
Frequency * Coherence	1	.017	.017	1.673	.203	.141	.243
Error (frequency)	39	1.484	.038				
Error (coherence)	39	1.336	.034				
Error (frequency*coherence)	39	.396	.010				
By items (N = 16)							
Frequency	1	.009	.009	1.346	.264	.082	.192
Coherence	1	.008	.008	2.467	.137	.141	.313
Frequency * Coherence	1	.002	.002	.277	.606	.018	.078
Error (frequency)	15	.099	.007				
Error (coherence)	15	.049	.003				
Error (frequency*coherence)	15	.127	.008				

Table 33: Analysis of variance results (for transformed proportion scores) for reading times across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	power
By participants (N = 40)							
Frequency	1	.041	.041	1.506	.227	.037	.224
Coherence	1	.013	.013	.762	.388	.019	.136
Frequency * Coherence	1	.035	.035	1.890	.177	.046	.268
Error (frequency)	39	1.055	.027				
Error (coherence)	39	.662	.017				
Error (frequency*coherence)	39	.729	.019				
By items (N = 16)							
Frequency	1	.013	.013	1.073	.317	.067	.163
Coherence	1	.001	.001	.194	.666	.013	.070
Frequency * Coherence	1	.015	.015	1.558	.231	.094	.215
Error (frequency)	15	.179	.012				
Error (coherence)	15	.111	.007				
Error (frequency*coherence)	15	.146	.010				

These analyses of variance again yielded no significant main effects or interactions for either the disambiguating or postdisambiguating region.

5.3 ANALYSES WITH PER PHONEME READING TIMES

For both the planned and post hoc analyses, there were few significant findings. One possible explanation for this lack of significant findings is related to the variability in the lengths (in graphemes and phonemes) of the target regions of the experimental sentences. To explore this possibility, each of the previously described analyses was repeated with per phoneme reading time as a control variable.

According to both dual-route (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), and connectionist (Seidenberg, 2005; Seidenberg & McClelland, 1989) models of reading, the process of reading involves, at least in some instances, translating orthographic symbols (or combinations of orthographic symbols) into phonemes. The dual-route model states that orthographic symbols are converted into speech in two different ways. In the lexical route, which is used for reading familiar words, a set of orthographic symbols composing a written word activates the entry for the corresponding spoken word in the individual's lexicon. For these familiar words, phonological mediation is not necessary. For sets of orthographic symbols for which no such lexical route is available (unfamiliar words and nonwords), the nonlexical route is used to indicate relationships between orthographic symbols and phonemes. If the resulting spoken word is in the individual's lexicon, its entry will be activated.

On the other hand, Seidenberg's connectionist model consists of three layers: an orthographic layer, a phonological layer, and a semantic layer. Converting orthographic symbols (or combinations of symbols) to phonemes involves activating the orthographic layer and allowing activation to pass between this layer and the phonological layer through the connections that have been established between them through prior experience. In a subsequent stage, activation is passed from the phonological layer to the semantic layer.

This connectionist model, therefore, holds that the series of orthographic symbols that represent written words must be converted to sets of phonemes before the semantic layer is activated to represent meaning. Similarly, the dual-route model proposes that the relationship between orthographic symbols and phonemes is used to decode unfamiliar words (and nonwords) before any meaning-related information is retrieved. This suggests that an individual's reading speed is related, at least for some words, to the rate at which he or she converts orthographic symbols to phonemes.

In addition to these models, a number of hypotheses have been proposed to explain the role of phonology in the reading process (Frost, 1994; Perfetti, Zhang, & Berent, 1992; Tabossi & Laghi, 1992). One of these, the phonology-plus-meaning hypothesis (Perfetti & Tan, 1998; Perfetti, Zhang, & Berent, 1992), asserts that phonological word forms are integral to the word identification process. That is, word identification involves retrieval of both the phonological form of the word and the meaning, rather than the phonological form simply mediating the retrieval of meaning from the orthographic form.

Evidence for this viewpoint comes from Perfetti and Zhang (1995), in which Chinese speakers completed two types of judgment tasks. In one task, they were asked to determine whether two characters had the same meaning. The second task required them to determine

whether two characters had the same pronunciation. For each task, some of the foils (critical foils) were target items from the alternate task. For example, in the meaning task, two foils might have the same pronunciation. Interference was assessed by comparing performance on the critical foils to those on control foils in which the two characters were unrelated. The results suggested that phonological interference occurred earlier than semantic interference. This suggests that this information is activated very rapidly, even when this information is not needed at the grapheme-phoneme level to access meaning.

If the activation of the phonological word form is an integral part of the reading processing, this suggests that the rate at which the words are accessed is related to the length of the spoken word (in phonemes), and that this directly impacts reading speed. For this reason, per phoneme reading rates were used to control of variability in reading speed among participants. To obtain the per phoneme reading times, the reading time for each participant across each target region was divided by the number of phonemes in that region.

Two types of analyses were conducted with phoneme counts controlled. For the analyses with outliers eliminated, there was a significant main effect of coherence by participants, $F(1,39) = 9.667$, $p < .05$, and by items, $F(1,15) = 9.451$, $p < .05$, across the disambiguating region. There were no other significant main effects or interactions for these analyses. For the analyses with proportion data, there were no significant main effects or interactions. These data are shown in Tables 34-37.

Table 34: Analysis of variance results (with per phoneme reading times and with outliers eliminated) for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	623.253	623.253	1.185	.283	.029	.186
Coherence	1	6037.796	6037.796	9.451*	.004	.195	.850
Frequency * Coherence	1	90.556	90.556	.204	.654	.005	.072
Error (frequency)	39	20520.546	526.168				
Error (coherence)	39	24915.170	638.851				
Error (frequency*coherence)	39	17353.969	444.974				
By items (N = 16)							
Frequency	1	430.195	430.195	4.259	.057	.221	.489
Coherence	1	1425.846	1425.846	9.667*	.007	.392	.828
Frequency * Coherence	1	78.242	78.242	.482	.298	.031	.100
Error (frequency)	15	1515.045	101.003				
Error (coherence)	15	2212.349	147.490				
Error (frequency*coherence)	15	2437.317	162.488				

Note. *p < .05.

Table 35: Analysis of variance results (with per phoneme reading times and with outliers eliminated) for reading times across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 40)							
Frequency	1	823.466	823.466	1.672	.204	.041	.243
Coherence	1	620.098	620.098	1.333	.255	.033	.203
Frequency * Coherence	1	1674.053	1674.053	1.626	.210	.040	.238
Error (frequency)	39	19206.781	492.482				
Error (coherence)	39	18147.809	465.328				
Error (frequency*coherence)	39	40157.432	1029.678				
By items (N = 16)							
Frequency	1	249.708	249.708	1.008	.331	.063	.156
Coherence	1	268.546	268.546	1.372	.260	.084	.195
Frequency * Coherence	1	230.713	230.713	1.943	.184	.115	.257
Error (frequency)	15	3714.630	247.642				
Error (coherence)	15	2935.475	195.698				
Error (frequency*coherence)	15	1781.022	118.735				

Table 36: Analysis of variance results (for proportion scores with per phoneme reading times) for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*Power*
By participants (N = 40)							
Frequency	1	5.60E-005	5.60E-005	2.743	.106	.066	.365
Coherence	1	6.70E-005	6.70E-005	.760	.389	.019	.136
Frequency * Coherence	1	2.32E-005	2.32E-005	.748	.392	.019	.135
Error (frequency)	39	.001	2.04E-005				
Error (coherence)	39	.003	8.82E-005				
Error (frequency*coherence)	39	.001	3.11E-005				
By items (N = 16)							
Frequency	1	1.79E-005	1.79E-005	.904	.357	.057	.145
Coherence	1	2.25E-005	2.25E-005	1.214	.288	.075	.178
Frequency * Coherence	1	1.22E-005	1.22E-005	.346	.565	.023	.086
Error (frequency)	15	.000	1.98E-005				
Error (coherence)	15	.000	1.85E-005				
Error (frequency*coherence)	15	.001	3.52E-005				

Table 37: Analysis of variance results (for proportion scores with per phoneme reading times) for reading times across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*Power*
By participants (N = 40)							
Frequency	1	3.90E-005	3.90E-005	1.336	.255	.033	.204
Coherence	1	3.38E-005	3.38E-005	1.853	.181	.045	.264
Frequency * Coherence	1	3.18E-005	3.18E-005	2.204	.146	.053	.305
Error (frequency)	39	.001	2.92E-005				
Error (coherence)	39	.001	1.82E-005				
Error (frequency*coherence)	39	.001	1.44E-005				
By items (N = 16)							
Frequency	1	3.66E-005	3.66E-005	1.403	.255	.086	.199
Coherence	1	1.63E-008	1.63E-008	.001	.980	.000	.050
Frequency * Coherence	1	.000	.000	3.884	.067	.206	.454
Error (frequency)	15	.000	2.61E-005				
Error (coherence)	15	.000	2.47E-005				
Error (frequency*coherence)	15	.001	3.47E-005				

5.4 ANALYSES WITH NONPARAMETRIC STATISTICS

Aside from the variability in the lengths of the target regions, another potential explanation for the paucity of significant findings in the previously described planned and post hoc analyses is that using ANOVA on these data sets might not have been appropriate. That is, the normality and homogeneity of variance assumptions might not have been met. To explore this possibility, the skewness, kurtosis, and homogeneity of variance of each data set were examined.

As stated earlier, data can be considered to be normally distributed, if the skewness and kurtosis values do not exceed the standard error. From the data from each analysis, there were several values that exceeded the standard error (Table 38).

Levene's Test was used to calculate homogeneity of variance. This test provides *p*-values that indicate the level of discrepancy in the variances for the various populations under investigation. In Table 38, there is only one case in which the *p*-value fell below .05, indicating that for most of the comparisons, the assumption of homogeneity of variance was met.

Analysis of variance is robust to the assumption of normality when the assumption of homogeneity of variance is met (Howell, 1992). However, the large number of skewness and kurtosis values that exceeded the standard error for all of the data sets make the appropriateness of analysis of variance questionable. For this reason, both the planned and post hoc analyses were repeated with the use of Friedman's two-way analysis of variance, which is a distribution free measure.

Table 38: Normality and homogeneity of variance characteristics for data from planned and post hoc comparisons.

Normality						*Homogeneity of Variance (Levene's Test)*
Initial Analysis						
Disambiguating Region	Comparison	Skewness	Std. Error	Kurtosis	Std. Error	
	2S/DC	.728	.564	-.902	1.091	$p = .158$
	6S/DC	1.321	.564	1.674	1.091	
	2S/SS	.667	.564	-.819	1.091	
	6S/SS	1.034	.564	-.139	1.091	
Postdisambiguating Region	2S/DC	2.866	.564	9.895	1.091	$p = .589$
	6S/DC	1.036	.564	1.390	1.091	
	2S/SS	1.283	.564	4.722	1.091	
	6S/SS	.919	.564	1.672	1.091	
Analyses with Outliers Eliminated						
Disambiguating Region	2S/DC	.800	.564	-.692	1.091	$p = .310$
	6S/DC	1.381	.564	1.893	1.091	
	2S/SS	.438	.564	-1.434	1.091	
	6S/SS	.978	.564	-.368	1.091	
Postdisambiguating Region	2S/DC	1.333	.564	3.903	1.091	$p = .723$
	6S/DC	1.174	.564	3.106	1.091	
	2S/SS	.152	.564	1.749	1.091	
	6S/SS	.495	.564	.412	1.091	
Analyses with Proportion Scores						
Disambiguating Region	2S/DC	.923	.564	.642	1.091	$p = .444$
	6S/DC	.799	.564	-.217	1.091	
	2S/SS	.867	.564	-.173	1.091	
	6S/SS	.965	.564	.219	1.091	
Postdisambiguating Region	2S/DC	.294	.564	-1.306	1.091	$p = .018$
	6S/DC	.472	.564	-.197	1.091	
	2S/SS	.791	.564	1.229	1.091	
	6S/SS	.504	.564	-.489	1.091	

While Friedman's two-way analysis of variance is a commonly used measure, it tests only for main effects and not for interactions (Howell, 1992). However, these analyses yielded a pattern of results that was similar to those obtained from the previous analyses. When the initial analyses (including outliers) were repeated using this nonparametric test, there was a significant main effect of coherence by items, $X^2_F = 10.125$, $p < .05$, but not by participants, $X^2_F = .502$, $p > .05$, across the disambiguating region. No other main effects were significant (Table 39).

Likewise, for the analyses with outliers eliminated, there was a significant main effect for coherence by items, $X^2_F = 4.500$, $p < .05$, but not by participants, $X^2_F = 2.450$, $p > .05$, across

the disambiguating region, with no other main effects being significant (Table 40). For the analyses with proportion scores, there were no significant main effects (Table 41).

For all of these analyses, the original reading time data, rather than data adjusted for per phoneme reading times, were used. If per phoneme reading time data had been used for these analyses, it is possible that the results would have been different.

Table 39: Results of Friedman's two-way analyses of variance conducted on original data set (with outliers included).

Disambiguating Region			
Variable	*df*	X^2_F	*p-value*
By Participants			
Frequency	1	.450	.502
Coherence	1	.450	.502
By Items			
Frequency	1	2.000	.157
Coherence	1	10.125*	.001
Postdisambiguating Region			
By Participants			
Frequency	1	.200	.655
Coherence	1	.000	1.000
By Items			
Frequency	1	1.125	.289
Coherence	1	.125	.724

Note. *p < .05.

Table 40: Results of Friedman's two-way analyses of variance conducted on original data set (with outliers eliminated).

Disambiguating Region			
Variable	*df*	X^2_F	*p-value*
By Participants			
Frequency	1	1.800	.180
Coherence	1	2.450	.118
By Items			
Frequency	1	.500	.480
Coherence	1	4.500*	.034
Postdisambiguating Region			
By Participants			
Frequency	1	2.579	.108
Coherence	1	.051	.821
By Items			
Frequency	1	2.793	.095
Coherence	1	.310	.577

Note. *p < .05.

Table 41: Results of Friedman's two-way analyses of variance conducted on proportion data.

Disambiguating Region			
Variable	*df*	X^2_F	*p-value*
By Participants			
Frequency	1	.000	1.000
Coherence	1	.050	.823
By Items			
Frequency	1	.000	1.000
Coherence	1	1.125	.289
Postdisambiguating Region			
By Participants			
Frequency	1	2.450	.118
Coherence	1	.450	.502
By Items			
Frequency	1	.500	.480
Coherence	1	.125	.724

5.5 POTENTIAL SOURCES OF VARIABILITY

As stated earlier, one potential explanation for the lack of significant findings in the planned analyses is that there was greater variability among participants and among stimuli for the current study than for the studies on which the effect size estimates were based. To further explore this possibility, several additional analyses were conducted.

5.5.1. Examination of Variability among Participants

Two possible ways in which the group of participants in the current study may have varied more from those in previous studies is in their reading ability and vocabulary skill. To determine whether the variability in participants' reading ability might have been related to greater variability in reading times, Pearson product-moment correlations were conducted to examine the relationship between participant's raw scores on the Slosson-R3 (the number of stimulus words read accurately) and their reading times (in ms) across the disambiguating and postdisambiguating regions.

Also, as stated earlier, there is evidence for a strong relationship between reading ability and vocabulary skill (Perfetti, Landi, & Oakhill, 2005). For this reason, the relationship between participants' raw scores on the PPVT-4 and their reading times following the coherence conditions was also investigated using Pearson product-moment correlations. In addition, these correlations were calculated separately for reading times following the discourse context and scrambled sentences conditions to infer whether participants were forming a coherent global representation when they read the coherent discourse contexts. Because of the relationship

between reading skill and coherence building, a significant correlation was expected for the coherent discourse context condition, if a coherent global representation was formed.

These analyses yielded a significant correlation between PPVT-4 raw scores and reading times across the disambiguating region in the scrambled sentence condition. In addition, there was a significant relationship between Slosson-R3 raw scores and (a) reading times across the disambiguating region following the scrambled sentence condition and (b) reading times across the disambiguating region combined over both coherence conditions. For the coherent discourse contexts, this relationship was not significant. All correlations are shown in Table 42. For all significant negative correlations, these suggest that participants with greater vocabulary skill or reading ability achieved faster reading times across the target regions.

Table 42: Correlations between PPVT-4 and Slosson-R3 raw scores and reading times across the target regions.

	Slosson –R3			*PPVT-4*		
	df	*Correlation*	*p-value*	*Df*	*Correlation*	*p-value*
Disambiguating Region/Scrambled Sentences	38	-.441	.017	38	-.396	.033
Disambiguating Region/Discourse Contexts	38	-.173	.369	38	-.146	.450
Disambiguating Region/Combined	38	-.401	.031	38	-.355	.059
Postdisambiguating Region/Scrambled Sentences	38	.110	.571	38	-.117	.545
Postdisambiguating Region/Discourse Contexts	38	-.048	.804	38	-.093	.631
Postdisambiguating Region/Combined	38	.027	.890	38	-.116	.550

5.5.2 Examination of variability among stimuli

Another factor that might have contributed to the lack of significant findings in the current study is that the experimental stimuli might have varied more than the stimuli used in the experiments on which the effect size estimates were based. One attribute of the experimental stimuli that was not controlled in the current experiment was the number of occurrences of the

target verbs across all components of the experimental stimuli, specifically in the experimental texts, filler items, comprehension questions, or as an embedded verb in the target sentences. As stated in the Review of the Literature, there is evidence from studies of *repetition priming* that items to which participants have been previously exposed tend to be processed more rapidly than new items (Cave, 1997). This may suggest that verbs that were repeated more times across the experimental stimuli might have been associated with faster reading times than those that occurred less often. Instances of each verb within the experimental texts were tallied, and this revealed that the verbs varied greatly in terms of number of prior occurrences (see Table 43).

Table 43: Number of occurrences (outside of the target sentences) of target verbs within the experimental stimuli.

Matrix Verb	# of Occurrences as Embedded Verb	# of Occurrences in Comprehension Questions	# of Occurrences in Experimental Stimuli	# of Occurrences in Filler Items	Total # of Occurrences
Admit	0	0	3	2	5
Announce	0	0	1	0	1
Decide	0	0	10	18	28
Expect	0	0	0	1	1
Explain	0	0	16	15	31
Guarantee	0	0	0	0	0
Indicate	0	0	7	0	7
Know	0	0	0	0	0
Maintain	0	0	9	1	10
Predict	0	1	3	0	4
Prove	0	0	2	1	3
Realize	0	0	4	4	8
Say	0	0	11	8	19
Sense	0	0	2	0	2
Show	0	1	5	1	7
Suggest	0	0	9	5	14

These data were then analyzed to determine whether the number of occurrences of target verbs outside of the target sentences affected participants' reading times across the target regions. To accomplish this, *t*-tests were used to compare the average reading times (across participants and conditions with outliers eliminated) for the disambiguating and

postdisambiguating regions following the five verbs with the highest number of occurrences to those following the five verbs with the lowest number of occurrences. Significant differences in reading times following these two sets of verbs were found for neither the disambiguating (t = .329, p = .532) nor the postdisambiguating region (t = .452, p = .664), suggesting that the number of occurrences of the target verb outside of the target sentences did not have a significant effect on participants' reading times.

5.6 GENDER-BASED ANALYSES

5.6.1 Original Analyses Repeated for Each Gender

Post-hoc analyses were conducted to determine whether female and male participants were differentially influenced by the frequency and coherence conditions. Several research studies report that male and female participants more accurately responded to comprehension questions and recalled more information from passages in which the topic was stereotypically biased toward their gender (Brantmeier, 2003; Bügel & Buunk, 1996; Doolittle & Welch, 1989; Hyde & Linn, 1988). Therefore, it is possible that the male and female participants performed differently in the current study. To examine this possibility, the previously described analyses were conducted separately for each gender to investigate potential gender effects. The descriptive statistics for these analyses are shown in Tables 44-47.

Table 44: Reading times (ms) for the disambiguating region (for data from female participants)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	466.06	253.37	152-1222
6S/discourse context	505.35	282.21	150-1401
2S/scrambled sentences	529.48	248.39	146-1153
6S/scrambled sentences	563.88	329.34	148-1942

Table 45: Reading times (ms) for the postdisambiguating region (for data from female participants)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	594.42	273.86	138-1626
6S/discourse context	689.03	282.60	149-1722
2S/scrambled sentences	714.96	294.89	174-1452
6S/scrambled sentences	684.98	262.86	222-1405

Table 46: Reading times (ms) for the disambiguating region (for data from male participants)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	527.57	253.39	177-1276
6S/discourse context	549.37	311.92	164-1439
2S/scrambled sentences	537.46	253.03	183-1211
6S/scrambled sentences	566.56	266.01	183-1307

Table 47: Reading times (ms) for the postdisambiguating region (for data from male participants)

Condition	*Mean (ms)*	*Standard Deviation*	*Range*
2S/discourse context	636.26	291.79	138-2202
6S/discourse context	686.38	507.83	149-2209
2S/scrambled sentences	694.20	301.29	174 -1663
6S/scrambled sentences	684.04	286.30	169-1453

For female participants, no significant effects were found for their reading times across the disambiguating region (Table 48). For their reading times across the postdisambiguating region, there was no significant main effect for frequency, either by participants, $F(1,19) = 1.311$, $p > .05$ or by items, $F(1,15) = 1.895$, $p > .05$. However, the main effect of coherence was significant by participants, $F(1,19) = 6.347$, $p < .05$, but not by items, $F(1,15) = 1.295$, $p > .05$,

with the reading times following discourse contexts being numerically faster than those following scrambled sentences. Finally, the interaction between frequency and coherence was marginally significant by participants, F (1,19) = 4.378, p = .05, but was not significant by items, F (1,15) = 3.795, p > .05, with the effect of frequency being greater in the scrambled sentence condition than in the coherent discourse context condition (Table 49). Again, the power was generally low, the highest (.67) being for the interaction of frequency and coherence across the postdisambiguating region in the *by participants* analysis.

Table 48: Analysis of variance results (for data contributed by female participants) across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 20)							
Frequency	1	58531.504	58531.504	1.747	.202	.084	.241
Coherence	1	83673.282	83673.282	3.513	.076	.156	.429
Frequency * Coherence	1	23521.226	23521.226	1.302	.268	.064	.192
Error (frequency)	19	634535.442	33396.602				
Error (coherence)	19	452514.977	23816.578				
Error (frequency*coherence)	19	343335.450	18070.287				
By items (N = 16)							
Frequency	1	18196.324	18196.324	1.149	.305	.071	.171
Coherence	1	32980.830	32980.830	4.245	.058	.221	.487
Frequency * Coherence	1	13859.666	13859.666	.981	.471	.061	.153
Error (frequency)	15	116542.511	15837.016				
Error (coherence)	15	237555.243	7769.503				
Error (frequency*coherence)	15	211823.278	14121.552				

Table 49: Analysis of variance results (for data contributed by female participants) across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 20)							
Frequency	1	15853.142	15853.142	1.311	.266	.065	.193
Coherence	1	76333.309	76333.309	6.347*	.021	.250	.667
Frequency * Coherence	1	99346.259	99346.259	4.378†	.050	.187	.511
Error (frequency)	19	229768.042	12093.055				
Error (coherence)	19	228502.959	12026.472				
Error (frequency*coherence)	19	431109.884	22689.994				
By items (N = 16)							
Frequency	1	18973.505	18973.505	1.895	.155	.112	.252
Coherence	1	32120.427	32120.427	1.295	.282	.079	.187
Frequency * Coherence	1	47326.319	47326.319	3.795	.062	.202	.446
Error (frequency)	15	150159.765	10010.651				
Error (coherence)	15	371915.269	24794.351				
Error (frequency*coherence)	15	187061.740	12470.783				

Note. *$p < .05$, $^{\dagger}p = .05$

For the male participants' reading times across both the disambiguating and postdisambiguating regions, the analyses yielded no significant effects. However, as in previously described analyses, the power of these analyses to detect a significant effect was low to medium for the comparisons involving both the disambiguating (ranging from .050 to .457) and postdisambiguating regions (ranging from .050 to .098; Tables 50 and 51).

Table 50: Analysis of variance results (for data contributed by male participants) across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	H^2	*power*
By participants (N = 20)							
Frequency	1	1.775	1.775	.000	.999	.000	.050
Coherence	1	289.117	289.117	.011	.956	.001	.051
Frequency * Coherence	1	25522.536	25522.536	1.831	.171	.088	.251
Error (frequency)	19	479678.959	25246.261				
Error (coherence)	19	506529.305	26659.437				
Error (frequency*coherence)	19	264870.366	13490.546				
By items (N = 16)							
Frequency	1	23002.597	23002.597	3.918	.059	.207	.457
Coherence	1	7582.678	7582.678	.957	.497	.060	.150
Frequency * Coherence	1	3835.959	3835.959	.812	.613	.051	.135
Error (frequency)	15	88065.619	5871.041				
Error (coherence)	15	118818.149	7921.210				
Error (frequency*coherence)	15	70894.009	4726.267				

Table 51: Analysis of variance results (for data contributed by male participants) across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	η^2	*power*
By participants (N = 20)							
Frequency	1	1933.070	1933.070	.130	.723	.007	.064
Coherence	1	529.163	529.163	.044	.837	.002	.055
Frequency * Coherence	1	55.695	55.695	.001	.975	.000	.050
Error (frequency)	19	283178.918	14904.154				
Error (coherence)	19	230318.158	12122.008				
Error (frequency*coherence)	19	1063694.647	55983.929				
By items (N = 16)							
Frequency	1	23002.597	23002.597	3.918	.059	.207	.457
Coherence	1	7582.678	7582.678	.957	.497	.060	.150
Frequency * Coherence	1	3835.959	3835.959	.812	.613	.051	.135
Error (frequency)	15	88065.619	5871.041				
Error (coherence)	15	118818.149	7921.210				
Error (frequency*coherence)	15	70894.009	4726.267				

5.6.2 Analysis of Interest Ratings by Gender

Several research studies suggest that participants' reading comprehension can be influenced by the gender bias of the material (Brantmeier, 2003; Bügel & Buunk, 1996; Doolittle & Welch, 1989; Hyde & Linn, 1988). To explore the possibility that this phenomenon may account for the different patterns of results for the male and female participants, an additional *post hoc* analysis was conducted.

Ten females who ranged in age from 19 to 58 years (mean = 24.6 years; SD =11.83 years) and 10 males who ranged in age from 19 to 61 years (mean = 27.8 years; SD = 14.03) completed a questionnaire rating their level of interest in the experimental texts using a five-point Likert-type scale (1 = not at all interesting; 5 = very interesting). There were two versions of the questionnaire, each including either the 2S or 6S coherent discourse context versions of the texts, and an equal number of each version was distributed to participants of each gender.

The levels of interest ratings of female and male participants for each text were compared using a series of Mann-Whitney U tests. This analysis revealed that, for two of the texts (Items 11 and 12), the female participants found the stimulus items significantly more interesting than did the male participants. There were no significant differences between genders for their levels of interest for any of the other items. For these analyses, power varied greatly, ranging from .047 to .884. However, most values were less than .40.

To explore whether there were gender differences across texts, the interest ratings of each participant were averaged. Using these values, an additional Mann-Whitney U test was conducted. Again, there was no significant difference between genders. As in the previous

analysis, the power for this test (.495) may not have been sufficient to detect a significant difference between groups.

The results of the Mann-Whitney U tests for both the individual texts as well as the averaged scores are shown in Table 52.

Table 52: Results of Mann-Whitney U tests to compare level of interest ratings across genders.

Stimulus Item	*U*	*p-value*	*Power*
1	45.000	.691	.156
2	47.500	.842	.091
3	42.000	.529	.079
4	35.500	.255	.445
5	42.000	.532	.169
6	37.000	.310	.400
7	31.000	.126	.385
8	37.000	.279	.290
9	38.000	.393	.204
10	40.000	.423	.243
11	24.500	.046	.801
12	24.500	.040	.884
13	46.500	.786	.070
14	49.000	.937	.047
15	41.000	.481	.118
16	32.000	.163	.532
Averaged Values	33.500	.212	.495

5.6.3. By Gender Analyses using of Per Phoneme Reading Times

In the previous analyses that were conducted with data from both genders, reading time data was converted to per phoneme reading times to control for variability in the length of the target regions of the experimental target sentences. Because these analyses yielded different results than those conducted with reading time data, it seemed likely that gender-based analyses using per phoneme reading times might yield different results than the previously described analyses, in which overall reading times were the dependent measure.

For the analyses with data from the female participants (with outlying data points eliminated), there was a significant main effect of coherence by participants, $F(1,39) = 8.966$, $p < .05$, and by items, $F(1,15) = 6.384$, $p < .05$, across the disambiguating region. These analyses yielded no other significant main effects or interactions. For the analyses with data from the male participants (with outlying data points eliminated), there were no significant main effects or interactions. The results of these analyses are shown in Tables 53-56.

Table 53: Analysis of variance results for data from female participants (with per phoneme reading times) for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	H^2	*Power*
By participants (N = 20)							
Frequency	1	368.397	368.397	1.213	.277	.030	.189
Coherence	1	1622.354	1622.354	6.384*	.016	.141	.693
Frequency * Coherence	1	289.020	289.020	1.938	.172	.047	.274
Error (frequency)	19	11844.198	303.697				
Error (coherence)	19	9911.768	254.148				
Error (frequency*coherence)	19	5815.770	149.122				
By items (N = 16)							
Frequency	1	341.035	341.035	.565	.464	.036	.109
Coherence	1	4643.556	4643.556	8.966*	.009	.374	.799
Frequency * Coherence	1	73.890	73.890	.163	.692	.011	.067
Error (frequency)	15	9047.209	603.147				
Error (coherence)	15	7768.632	517.909				
Error (frequency*coherence)	15	6808.073	453.872				

Note. *p < .05

Table 54: Analysis of variance results for data from female participants (with per phoneme reading times) for reading times across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	H^2	*power*
By participants (N = 20)							
Frequency	1	175.213	175.213	1.691	.201	.041	.245
Coherence	1	266.821	266.821	2.803	.102	.065	.372
Frequency * Coherence	1	212.288	212.288	3.844	.057	.088	.482
Error (frequency)	19	4144.926	103.623				
Error (coherence)	19	3807.260	95.181				
Error (frequency*coherence)	19	2208.845	55.221				
By items (N = 16)							
Frequency	1	268.252	268.252	1.167	.297	.072	.173
Coherence	1	245.271	245.271	1.259	.279	.077	.183
Frequency * Coherence	1	417.349	417.349	3.932	.066	.208	.459
Error (frequency)	15	3448.635	229.909				
Error (coherence)	15	2921.405	194.760				
Error (frequency*coherence)	15	1591.994	106.133				

Table 55: Analysis of variance results for data from male participants (with per phoneme reading times) for reading times across the disambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	H^2	*power*
By participants (N = 20)							
Frequency	1	174.336	174.336	.832	.367	.021	.144
Coherence	1	597.886	597.886	2.859	.099	.068	.378
Frequency * Coherence	1	717.447	717.447	3.624	.064	.085	.459
Error (frequency)	19	8176.366	209.650				
Error (coherence)	19	8154.802	209.097				
Error (frequency*coherence)	19	7720.730	197.967				
By items (N = 16)							
Frequency	1	531.161	531.161	2.196	.159	.128	.284
Coherence	1	845.193	845.193	3.044	.102	.169	.372
Frequency * Coherence	1	132.702	132.702	.480	.499	.031	.100
Error (frequency)	15	3627.331	241.822				
Error (coherence)	15	4165.379	277.692				
Error (frequency*coherence)	15	4143.319	276.221				

Table 56: Analysis of variance results for data from male participants (with per phoneme reading times) for reading times across the postdisambiguating region.

Variable	*df*	*SS*	*MS*	*F*	*p-value*	H^2	*power*
By participants (N = 20)							
Frequency	1	35.572	35.572	.432	.515	.010	.098
Coherence	1	4.573	4.573	.174	.679	.004	.069
Frequency * Coherence	1	153.430	153.430	3.756	.059	.082	.474
Error (frequency)	19	3461.091	82.407				
Error (coherence)	19	1103.193	26.267				
Error (frequency*coherence)	19	1715.473	40.845				
By items (N = 16)							
Frequency	1	55.175	55.175	.098	.758	.007	.060
Coherence	1	532.166	532.166	.783	.390	.050	.132
Frequency * Coherence	1	1354.462	1354.462	1.142	.302	.071	.170
Error (frequency)	15	8492.807	560.187				
Error (coherence)	15	10197.706	679.847				
Error (frequency*coherence)	15	17795.720	1186.381				

5.6.4 By Gender Analyses using Nonparametric Statistics

Like the per phoneme analyses, the nonparametric analyses that were conducted on data from both genders yielded different results than the original analyses. For this reason, nonparametric statistics were also used to analyze data contributed by each gender.

For the analyses with data from female participants (with outlying data points eliminated), there was a significant main effect of coherence by items, $X^2_F = 4.500$, $p < .05$, but not by participants, $X^2_F = 3.600$, $p > .05$, across the disambiguating region. No other main effects were significant (Table 57). Similarly, for the analyses with data from male participants (with outlying data points eliminated), there was a significant main effect of coherence by items,

$X^2_F = 6.125$, $p < .05$, but not by participants, $X^2_F = .200$, $p > .05$, across the disambiguating with no other significant main effects (Table 58).

Table 57: Results of Friedman's two-way analyses of variance conducted on data from female participants.

Variable	*df*	X^2_F	*p-value*
Disambiguating Region			
By Participants			
Frequency	1	2.500	.114
Coherence	1	3.600	.058
By Items			
Frequency	1	.500	.480
Coherence	1	4.500*	.034
Postdisambiguating Region			
By Participants			
Frequency	1	.000	1.000
Coherence	1	2.500	.114
By Items			
Frequency	1	3.125	.077
Coherence	1	2.000	.157

Note. *p < .05.

Table 58: Results of Friedman's two-way analyses of variance conducted on data from male participants.

Variable	*df*	X^2_F	*p-value*
Disambiguating Region			
By Participants			
Frequency	1	.000	1.000
Coherence	1	.200	.655
By Items			
Frequency	1	1.720	.085
Coherence	1	6.125*	.013
Postdisambiguating Region			
By Participants			
Frequency	1	.100	.752
Coherence	1	.900	.343
By Items			
Frequency	1	3.125	.077
Coherence	1	2.000	.157

Note. *p < .05.

5.6.5 Possible Explanations for Gender-Based Differences

Across almost all of the post hoc analyses, different patterns of performance emerged for the male and female participants. To explore whether alternative explanations can be offered for these differences, a series of analyses were conducted.

5.6.5.1 Possible Differences in Reading and/or Vocabulary Skill First, because previous analyses indicated a relationship between reading times and both reading skill and vocabulary in at least one condition (section 5.7.1), the possibility was explored that differences in performance might be due to differences in the reading or vocabulary skills of the male and female participants. To examine this possibility, *t*-tests were conducted to compare the participants of each gender on their raw scores on the PPVT-4 (Dunn & Dunn, 2007) and on the Slosson-R3 (Slosson & Nicholson, 2002). There was no significant difference between groups for their scores on the PPVT-4, ($t = 1.88$, $p = .068$), although the mean score of the male participants was numerically higher. In addition, there was no significant difference between groups for their scores on the Slosson-R3 ($t = .473$, $p = .639$), although the mean score of the female participants was numerically higher.

5.6.5.2 Possible Differences in Demand Characteristics A second possible explanation for the apparent gender-based differences in performance is related to differences in demand characteristics between the two groups. Of the twenty female participants, eight were currently enrolled in at least one course taught by the investigator. By comparison, only three of the twenty male participants were enrolled in courses taught by the investigator. It is plausible that

individuals who were currently taking a course from the investigator might have been more motivated to perform well on the experimental task. To investigate this possibility, Mann-Whitney U tests were conducted to compare reading times across both target regions for participants who were and who were not taking a course from the investigator. For both target regions, there was no significant difference between the two groups. (Table 59).

Table 59: Results of Mann-Whitney U tests to compare reading times for individuals who were and were not enrolled in a course taught by the investigator.

	Disambiguating Region	Postdisambiguating Region
Mann-Whitney U	95.000	82.000
Significance	.693	.335

5.6.5.3 Possible Differences in Variability within Groups A third possibility was that the lack of significant findings for the analyses of data from male participants might be due to the higher variability of their data, and consequent lower power, than that for the female participants. Separate power analyses were conducted for the data from the male and female participants. For reading times across both the disambiguating and postdisambiguating regions, power values for the analyses of data from male participants ranged from .050 to .457 (mean = .128). For the analyses of data from female participants, these values ranged from .171 to .667 (mean = .327). While the ranges of power values obtained for the two gender groups did overlap, the large difference in the mean power value might suggest that differences in power contributed to the different pattern of results obtained for the two groups.

One factor that might have contributed to the different power values that were obtained for the two groups is the greater variability among male participants' data. To investigate this

possibility, *t*-tests were conducted to compare the variances in the distribution of reading times across the disambiguating and postdisambiguating regions. These analyses indicated that there was significantly greater variability in the reading times across the disambiguating region for the male participants than for the female participants ($t = 3.235$, $p = .002$). However, there was no significant difference in variability for the two groups across the postdisambiguating region ($t = .396$, $p = .692$).

The greater variability in reading times for the male participants across this region might have been related to a greater variability within this group in terms of vocabulary or reading skill. To explore this possibility, the data for individual participants were inspected for the presence of outlying reading and vocabulary test scores that might be related to gender-based differences on the experimental task. Inspection of these data revealed that, for both groups, no participants' mean reading time for either target region was more than two standard deviations from the mean reading time for that region.

Similarly, to investigate whether male and female participants' reading times differed in terms of within-subject variability, each participant's reading times across each trial were examined, and variances were calculated. Four of the five participants who exhibited the greatest variability in performance were male. One male participant's variance in reading times was more than two standard deviations above the mean variance for all male participants. No female participants had a variance in reading time more than two standard deviations from the mean variance for their gender.

6.0 DISCUSSION

6.1 SUMMARY

The purpose of this study was to investigate the potential influences of discourse context and syntactic priming on predicted verb subcategorization. It was hypothesized that both the presence of a coherent discourse context as well as a greater number of similar structures in the previously occurring sentences (syntactic priming) would lead to faster reading times across the disambiguating region of target sentences. In addition, it was predicted that there would be a significant interaction of these factors, such that the influence of syntactic priming would be greater for the scrambled sentence conditions than for the coherent discourse context conditions. Finally, it was hypothesized that discourse context would significantly influence participants' reading times across the postdisambiguating region of the target sentences, but that there would be no significant effect of syntactic priming across this region, nor a significant interaction between these factors.

6.2 LACK OF SIGNIFICANT FINDINGS IN PLANNED ANALYSES

These predictions were not upheld in any of the planned analyses. However, the post hoc analyses in which per phoneme reading times were the dependent measure, and the analyses with nonparametric statistics provide some preliminary evidence for an influence of coherence on sentence processing. While the post hoc nature of these analyses limits their interpretability, they do suggest that the lack of significant findings in the planned analyses might have been due, at least in part, to the choice of statistical procedure that was used to analyze these data.

6.2.1 Participant variability

A second possible explanation for the lack of significant findings is participant variability. The follow-up analyses suggest that variability among participants may have resulted in effect sizes that were smaller than those obtained in similar studies on which the sample size estimates were based. These analyses revealed a significant relationship between Slosson-R3 raw scores and reading times across the disambiguating region (for the initial analyses with outliers eliminated), and specifically, reading times across this region following the scrambled sentences condition. Participants with higher raw scores on the Slosson-R3 had faster reading times across the disambiguating region in the scrambled sentence condition. In addition, there was a significant correlation between PPVT-4 raw scores and reading times across the disambiguating region (for the initial analyses with outliers eliminated) following the scrambled sentences condition. Participants who obtained higher raw scores on either the PPVT-4 or on the Slosson-

R3 had faster reading times across the disambiguating region following the scrambled sentence condition than participants who obtained lower scores on these measures.

Further analyses were conducted to determine whether the participants who obtained lower scores on either the Slosson-R3 or the PPVT-4 had disproportionately longer reading times in the scrambled sentence condition. A finding of this nature would suggest that the reading times of participants who were poorer readers (or who might have greater difficulty with the task due to a weaker receptive vocabulary) were particularly long, because these readers were attempting but failing to create a discourse representation in the scrambled sentence condition. These same readers might have benefitted from the cues to a coherent discourse representation that were present in the coherent condition. To examine this possibility, the mean reading times across the disambiguating region were calculated separately for the participants whose Slosson-R3 and PPVT-4 scores were in higher and lower half of all participants. The difference between the mean reading times in the coherent discourse context and scrambled sentence conditions were calculated for both groups of participants. The results are shown in Table 60.

Table 60: Mean reading times across the disambiguating region for participants scoring in the higher half and lower half of all participants on the Slosson-R3 and PPVT-4.

	Higher half	*Lower half*
Slosson- R3		
Discourse Context	496.47	535.61
Scrambled Sentence	498.75	593.92
Difference	2.28	58.31
PPVT-4		
Discourse Context	485.27	546.06
Scrambled Sentence	519.48	574.43
Difference	34.21	28.37

As shown in Table 60, stronger readers (those scoring in the higher half of all participants on the Slosson-R3) obtained relatively similar reading times across the disambiguating region in the coherent discourse context and scrambled sentence conditions. The poorer readers, conversely, differed more in their reading times in the two conditions, with a 58.31 ms difference in mean reading times between the two conditions. This raises the question of whether the poorer readers were more affected by the lack of a coherent discourse context than were the stronger readers. On the other hand, those scoring in the higher and lower half of participants on the PPVT-4 experienced a similar increase in reading time between the coherent discourse context and scrambled sentence conditions, suggesting that vocabulary level did not play a significant role in participants' responses to the coherence conditions. It is possible that this lack of a significant influence of vocabulary level might be due to the limited task demands

in the area of vocabulary. Because the experimental items were short narratives, it is likely that reading them did not require knowledge of any specialized vocabulary that might have been difficult for the participants who scored more poorly on the PPVT-4. Therefore, all participants might have been familiar with the vocabulary words needed to complete the experimental task.

To determine whether the differences in reading times between the two coherence conditions were significant for each of the reading ability groups, two paired *t*-tests were conducted. These tests compared the mean reading times across the disambiguating region following the scrambled sentence and coherent discourse context conditions, one for the participants whose raw scores on the Slosson-R3 were in the higher half (higher reading ability) and one for those in the lower half (lower reading ability) of all participants. For both the comparison involving the higher reading ability group ($t = .182$; $p = .858$) and the one involving the lower reading ability group ($t = 1.89$; $p = .069$), no significant differences were found in reading times in the coherent discourse context versus scrambled sentence conditions. This suggests that neither group was affected by the lack of a coherent discourse context. However, the small *p*-value for the lower reading ability group may suggest that a larger sample size might yield a significant coherence effect.

In a similar analysis, two paired *t*-tests were used to compare mean reading times across the disambiguating regions in the two coherence conditions, one for those participants whose raw scores on the PPVT-4 were in the higher half (higher vocabulary skill) and one for those in the lower half (lower vocabulary skill) of all participants. Again, neither the higher vocabulary skill group ($t = .021$; $p = .983$) nor the lower vocabulary skill group ($t = 1.336$, $p = .203$) demonstrated a significant difference in reading times between the two coherence conditions.

The findings of a significant relationship between reading times across the disambiguating region and both reading and vocabulary skill can be accounted for by both interactive and autonomous models of sentence processing. The results are consistent with an interactive model of sentence processing, because the significant effect of vocabulary on reading times across the disambiguating region (rather than across the postdisambiguating region) suggests that semantic information plays an early role in sentence parsing.

However, the methods used in this study were not sufficiently sensitive to determine precisely how early semantic information influenced sentence processing, The self-paced reading paradigm used in the current investigation was intended to measure participants' reading times across each of the words in the disambiguating and postdisambiguating regions. However, these methods cannot separate first-pass reading times (that are generally attributed to the initial stage of sentence processing) from reading times during second-pass sentence parsing (that are believed to reflect a reanalysis stage of processing). Thus, these results cannot rule out the autonomous view, because it is possible that semantic information influenced reading times during a reanalysis stage, rather than during the initial stage of processing. In future investigations, eye-tracking methodology would be useful for making this determination because it would enable investigators to separate first-pass and second-pass reading times (Britt et al., 1992; Trueswell et al, 1993).

In addition to their reading and vocabulary skills, the variability in other participant characteristics may have been greater for the current study than for the studies from which effect sizes for this study were estimated. First, for three of the five studies from which effect sizes were estimated (Hess et al., 1995; Kashak et al., 2006; Kashak & Glenberg, 2004), all of the participants were introductory psychology students. Therefore, these participants might have

been more similar to one another in terms of their age, motivation, background knowledge, and other factors that might have influenced performance on the experimental task. Participants in the current study, for example, had a relatively broad range of performance on both the PPVT (standard scores ranged from 87 to 138) and the Slosson-R3 (grade-equivalent scores ranged from ~ 10 to beyond the twelfth grade level).

6.2.2 Stimulus characteristics

6.2.2.1 Stimulus variability In addition to variability across participants, the stimulus items varied both in terms of the length of the target regions, and the number of times the matrix verbs of the target sentences appeared in positions outside of the target sentences (within the experimental or filler texts, filler target sentences, or in the comprehension questions). In one set of post hoc analyses, the issue in variability in length of the target region was controlled by calculating per phoneme reading times. Because decoding involves associating letters and letter combinations with corresponding phonemes, per phoneme reading times were deemed to be an appropriate means of equating reading time across target regions of varying lengths. The use of this metric yielded a significant main effect of coherence across the disambiguating region (for the initial analyses with outliers eliminated).

In addition, the frequency of matrix verbs within the experimental stimuli (outside of the target sentences) was determined by tally, and this revealed that some matrix verbs occurred much more frequently outside of the target sentences than did others. *T*-tests were used to compare the reading times for the target regions following the five matrix verbs that occurred most frequently in the stimuli to those following the five matrix verbs that occurred least

frequently. These did not yield significant findings, suggesting that variability in the frequency of the matrix verbs within the experimental stimuli was not a significant source of stimulus variability, at least when phoneme counts were not controlled.

6.2.2.2 The Nature of the Stimulus Items Aside from issues related to variability, the lack of significant findings in the current study may have been at least partially due to various characteristics of the stimulus items. First, for thirteen of the sixteen experimental target sentences, the disambiguating region consisted of a form of the verb "to be". Because this verb is extremely frequent in English, it is possible that the participants were able to read and process this verb more quickly and automatically than would be the case for other verbs. In addition, this greater familiarity might have caused reading times for this verb to be less greatly influenced by the nature of the prior context than might be the case for other verbs. This possibility could be examined in future investigations by comparing the influence of prior context on reading times for forms of "to be" to its influence on reading times for other verbs.

Second, the reading times of participants were recorded as they pressed the space bar to reveal words that replaced dashes on the computer screen. It is possible that once participants became familiar with the repetitive motor pattern associated with this task, that the duration for each word in the target regions reflected this practiced pattern rather than the amount of time required by participants to read each word. This possibility could be explored in future investigations by comparing the influence of prior discourse context on reading times at different points during the experiment. If the nature of the prior discourse context is related to reading times early in the experiment, but not later, this would support the possibility that participants'

increasing familiarity with the motor pattern reduced the influence of prior discourse context on their response times.

6.3 JUSTIFICATION FOR RESEARCH HYPOTHESES

Finally, several weaknesses in the available literature suggest that the research hypotheses may not have been justified by this literature. First, as mentioned in the Review of the Literature, there is only indirect evidence that verb subcategorization varies across corpora of different purposes, modalities, and registers. Until more direct evidence is available, it remains possible that there is no reason for comprehenders to use frequency-based information to vary their expectation for upcoming structures, because the frequency of these structures does not vary across contexts. The lack of significant main effect of frequency in the current study suggests that this is a possibility.

Second, the studies of cumulative syntactic priming that were described in the literature each required participants to respond to large numbers of unrelated sentences that were presented within a short time (Bock & Griffin, 2000; Kashak et al, 2006; Kashak & Glenberg, 2004; Luka & Barsalou, 2005; Noppeney & Price, 2004). These types of experimental procedures might have alerted at least some participants to the fact that the experiments addressed issues related to syntax. This awareness might have caused participants to attend more closely to the syntactic structures of the sentences that were presented. Because the current experiment did not use this format, the participants might have been less likely to notice and attend to the syntactic structures of the experimental sentences.

Third, for several of the studies that reported an influence of discourse context on word or sentence processing (Altmann & Steedman, 1988; Hare et al, 2003; Spivey & Tanenhaus, 1998; Vu et al, 2000), participants' response times following a condition that was intended to support a particular interpretation were compared to those following a condition that was intended to support an alternative interpretation. There was no neutral condition. Therefore, it is possible that the significant results obtained in these studies might have resulted from a combination of the facilitating influence of the biasing condition and the inhibitory influence of the alternative condition. Because the current study compared reading times following coherent discourse contexts intended to support an S interpretation to those following scrambled sentences (a neutral condition), rather than to reading times following a NP-supporting context, this could at least in part explain the lack of significant findings in the planned analyses.

6.4 IMPLICATIONS OF FINDINGS REGARDING THE POSSIBLE MAIN EFFECT OF COHERENCE

In the planned analyses, the findings of no significant influence of coherence on reading times across the target regions do not comport with the predictions of the majority of the comprehension models (Elman, 1991, 1992; Gibson, 1998, 2000; Graesser, Olde, & Klettke, 2002; MacDonald & Christiansen, 2002) that were described in the Review of the Literature. Two of these models (Graesser, Olde, & Klettke, 2002; Trabasso & Wiley, 2005) would predict a significant influence of coherence on reading times. The other three models (Elman, 1991, 1992; Gibson, 1998, 2000; MacDonald & Christiansen, 2002) would offer less specific

predictions for the outcome of the current experiment, but each proposes that the nature of previously occurring sentences influences current sentence processing.

A finding of no significant effect of coherence on participants' reading times is in line with Kintsch's Construction-Integration model (1988; 1998). As mentioned in the Review of the Literature, this model would predict that there would be no significant difference in reading times for target sentences following the scrambled sentence and coherent discourse context conditions. The model offers no prediction regarding the potential influence of frequency.

If the finding for no significant influence of coherence or syntactic priming is real, this would have implications for at least three models of language processing. As stated in the Review of the Literature, Gibson's Discourse Locality Theory (DLT; 1998, 2000) posits that the selection of a particular syntactic structure is based upon the relative activation levels of the various alternative structures that are available at that point in sentence processing. In addition, the findings of Chen, Gibson, and Wolf (2005) and Warren and McConnell (2006) suggest that parsing might also involve selecting among alternative verb subcategorizations. If a future study replicates the lack of influence of either coherence or syntactic priming on predicting verb subcategorizations, this would suggest that discourse-level information has little or no influence on this selection process, and that the prediction for upcoming structures is guided primarily by sentence-level information. In order to take into account these findings, the DLT (1998; 2000) would need to be revised to state that only sentence-level information contributes to the activation level of the various alternative syntactic structures.

The constructionist models of both Elman (1991; 1992) and MacDonald and Christiansen (2002) state that subcategorizations associated with verbs are learned by the co-occurrence between the verb and certain types of subcategorizations. Connections are established between

the verb and the subcategorizations with which it co-occurs, with greater connections weights associated with the subcategorization that occurs most frequently with the verb. A lack of effect of either coherence or syntactic priming would suggest that neither of these factors in the previously occurring sentences influences the connection weights between the verb and its subcategorizations at the point at which the verb is being processed. To take into consideration such findings, connectionist models would need to be amended, so that the connection weights between each word and the other words with which it is associated in the network would not be influenced by discourse-level information.

The models of Graesser, Olde and Klettke (2002) and Trabasso and Wiley (2005) both describe the process by which discourse-level coherence is established, and predict faster processing of sentences that are related to the prior context. Therefore, if the finding for no significant main effect of coherence can be replicated, this finding would be incompatible with these models. While the explanation offered by these models for the establishment of coherence would not conflict with such findings, the models would need to be revised such that coherence would not facilitate current sentence processing.

It is also possible that the significant main effect of coherence that was observed across the disambiguating region in several of the follow-up analyses in this study is a more accurate reflection of the nature of sentence processing. This finding would be in accordance with the majority of the models of sentence processing and discourse comprehension described earlier, (Elman, 1991, 1992; Gibson, 1998, 2000; Graesser, Olde, & Klettke, 2002; MacDonald & Christiansen, 2002; Trabasso & Wiley, 2005). In addition, a significant main effect of coherence across the disambiguating region would be consistent with interactive models of sentence processing, because it would suggest that nonsyntactic information, such as discourse context

information, plays an early (and possibly an initial) role in sentence parsing. However, while the results suggest that nonsyntactic information plays an early role in sentence parsing, the methods used in the current study were not sensitive enough to ascertain whether this information plays an initial role in sentence parsing. Therefore, the results can also be accounted for by the autonomous viewpoint. Finally, an effect of coherence would be consistent with top-down theories of sentence processing, because it would suggest that readers use contextual information to predict upcoming syntactic structures.

6.5 IMPLICATIONS FOR FINDINGS REGARDING PREDICTED MAIN EFFECT OF FREQUENCY

For both the planned and post hoc analyses, there was no significant main effect of frequency. These results are in contrast to the results of Bock and Griffin (2000), Kashak et al (2006), and Luka and Barsalou (2005). These studies provide evidence for a cumulative syntactic priming effect in which the production or comprehension of a syntactic structure was facilitated in proportion to the number of times the structure had occurred across sets of previously occurring unrelated sentences.

One possible explanation for the lack of a significant main effect for frequency might be related to the inclusion of both scrambled sentence and discourse context conditions in the current experiment. Previous studies of syntactic priming studied its influence only across sets of scrambled sentences. As stated earlier, the influence of syntactic priming was predicted to be greater across scrambled sentences than across coherent discourse context condition. This is

because scrambled sentences do not require the reader to attend to a global representation, thus perhaps allowing them to attend more closely to the syntactic structures of the sentences. If the influence of syntactic priming is greater across scrambled sentences than across coherent discourse context (but not to the extent necessary to yield a significant interaction between coherence and frequency) this could account for the results of the current study.

6.6 LIMITATIONS OF THE STUDY

Aside from the previously described variability among participants and stimuli, as well as stimulus characteristics that might have resulted in the inaccurate effect size estimates described earlier, there are at least two other limitations that should be considered when interpreting the results of this study. First, each of the experimental texts was eight sentences in length. It is unclear whether different results would have been obtained if longer texts had been used. Using varying lengths of text in the initial design of the study would have allowed the investigator to determine whether the inclusion of a greater number of target syntactic structures would have increased the degree to which the frequency of these structures facilitates sentences processing. As stated in the Review of the Literature, the results of Kashak et al. (2006; Experiments 1 and 2) suggest that the influence of syntactic priming accumulates across trials. This implies that a greater number of target syntactic structures (in addition to a greater proportion of structures) in the prior discourse context could have increased the degree to which the relative frequency of these structures influences sentence processing.

Second, potential gender effects were not considered *a priori* in the initial design of this study. Having done so would have allowed the experimenter to develop stimuli that were intended to be more or less interesting to male and female participants. If other relevant variables were equated between males and females, this would have made the gender-based findings of this study more interpretable than was possible through the use of *post hoc* analyses.

6.7 GENDER DIFFERENCES

6.7.1 Results for Male vs. Female Participants

The effects of coherence and frequency on the prediction for upcoming verb subcategorizations were analyzed separately for each gender using a variety of methods. In the first set of analyses, participants' unadjusted reading times were the dependent measure. For these analyses, the patterns of results were different for the male and female participants. For males, there were no significant main effects or interactions for either the disambiguating or postdisambiguating region of the target sentences. For the female participants, reading times across the postdisambiguating region were significantly faster following coherent discourse contexts than following scrambled sentences. This outcome was predicted for all participants, and suggests that coherence does influence a female reader's predictions for upcoming verb subcategorizations.

In addition, there was a marginally-significant (p = .05) interaction between coherence and frequency for reading times across the postdisambiguating region. This interaction is

interesting because its direction suggests that the influence of the frequency of previously occurring syntactic structures was greater in the scrambled sentence condition than in the coherent discourse context condition. However, this interaction was predicted for the disambiguating, not the postdisambiguating region.

The results of prior studies of the cumulative effects of syntactic priming (Bock & Griffin, 2000; Kashak & Glenberg, 2004; Kashak et al., 2006) suggest that the influence of syntactic priming accumulates across trials or sentences. However, these studies examined this effect only across sets of unrelated sentences. The numerically greater influence of syntactic priming in the coherent discourse context condition than in the scrambled sentence condition in this study suggests that further research is needed to determine the extent to which the findings for cumulative syntactic priming across unrelated sentences can be generalized to everyday discourse comprehension.

Additional analyses involved the use of per phoneme reading times rather than total reading times across the target regions as the dependent measure. From these analyses, a similar gender-based pattern emerged. For female participants, there was a significant main effect of coherence. However, in these analyses, this effect was found across the disambiguating region rather than across the postdisambiguating region. There was no significant main effect of frequency nor a significant interaction between frequency and coherence for either region. For the male participants, there were, once again, no significant main effects or interactions found for either region.

These data were also analyzed using nonparametric methods. These analyses yielded a significant main effect of coherence (by items but not by participants) for both female and male

participants. The different pattern of results obtained in these analyses may be related to the rank order data that was used for the nonparametric methods.

A potential reason for the apparent gender-based differences was that greater variability in the reading times for the male participant group might have diminished the effect sizes obtained in this group. A post-hoc *t*-test comparing the variances of the male and female groups revealed that the group of male participants varied significantly more in terms of reading times across the disambiguating region (but not the postdisambiguating region) than did the group of female participants.

Conversely, the results of additional analyses suggest that factors such as receptive vocabulary did not contribute to gender-based differences in performance. There was no significant difference between male and female participants on the PPVT-4, and the males obtained a numerically higher mean score on this measure than did the female participants.

6.7.2 Potential Reasons for Gender Differences

Most of the follow-up analyses yielded gender-based differences in the pattern of results. Although the *post hoc* nature of these comparisons limit their interpretability, at least five possible reasons can be offered for these gender-based differences.

First, while most participants were university students, six participants (five males and one female) were recruited from the Indiana, PA community. Because most of these participants were male, several differences between these participants and the university students might have contributed to the gender-based differences in performance. For example, the non-university students had, on average, an educational level that was more than a year lower than that of the

university students (non-university students: mean= 13.67; university students: mean = 14.84), and none were currently enrolled in an educational program. Thus, it is possible that the non-university students were not doing as much reading on an everyday basis as the university students. In addition, their somewhat lower average educational level might suggest that they had less reading experience, in general, than the university students. Both of these factors might have made the experimental task somewhat more challenging for these participants than for the university students. If these participants found the sentence-by-sentence reading requirements to be more difficult, they might have been less able to attend to the overall meaning that was presented in the coherent discourse contexts.

Second, the gender-based differences that were found in this study could possibly be related to evidence for greater depth of semantic elaboration in females. Meyers-Levy and Maheswaran (1991) suggest that female participants' more accurate performance on an information recognition task might have been attributable to their more detailed elaboration of the content of a message. Similarly, Wirth, Horn, Koenig, Stein, Federspiel, Meier, Michel, and Strik (2007) found that female participants' integration times for words were more greatly influenced by the relatedness of the previously presented word. This might suggest that, in the current investigation, females' processing of the verb subcategorizations in the target sentences in the current experiment might have been more greatly influenced by the coherence of the preceding discourse context.

A third possible explanation for the different pattern of results for male and female participants is that female participants might be more inclined to pattern their responses after others. That is, they might be more likely to replicate the gestural, postural, and perhaps linguistic patterns that are present in the communicative environment. There is evidence that

female participants are more likely than male participants to engage in *postural mirroring*, or sharing the physical posture of those with whom they are interacting (Bernieri, Davis, Rosenthal, & Knee, 1994; Grammar, Kruck, & Magnusson, 1998; La France & Ickes, 1981).

These studies suggest that females, at least when they are engaged in an interaction, are more receptive to nonverbal models. The greater receptivity of females to nonverbal models may suggest that they might be more greatly influenced by linguistic models as well. In the current investigation, participants were provided with either two or six models of a particular syntactic structure. If the female participants were more receptive to these syntactic models, this could, at least in part, explain why there was a significant interaction between coherence and frequency for reading times across the postdisambiguating region for the female, but not for the male participants. And, while this interaction was expected to occur in the disambiguating region, it is possible that the short length of the disambiguating region caused a "spill-over" of this effect into the postdisambiguating region.

Fourth, the gender-based differences might be at least partially due to the female participants' being more compliant with the experimenter's instructions to read the experimental texts for comprehension. There is evidence that female children are more compliant with maternal requests than are male children (Abe & Izard, 1999; Kochanska, Tjebkes, & Forman, 1998). Similarly, the results of a meta-analysis of the relationship between gender and helping behavior (Eagly & Crowley, 1986) suggest that females are more willing to participate in routine acts of personal service and volunteerism. Because serving as a participant in a research study was a volunteer activity, it seems likely that the female participants were more motivated to participate in the experiment. This may have led them to pay more attention to the experimental

stimuli, and therefore to benefit more (in terms of reading times across the target sentences) from the global representation of the coherent texts.

Finally, the possibility was explored that these results might be due to differences in the levels of interest for the topics of the experimental texts. There is evidence that male and female individuals perform differently on measures of comprehension (e.g., multiple choice tasks, number of content units recalled) for materials on various topics (Brantmeier, 2003; Bügel & Buunk, 1996; Doolittle & Welch, 1989; Hyde & Linn, 1988). However, this finding is not entirely consistent. Young and Oxford (1997) reported no significant differences in participants' responses to comprehension questions for passages of various topics for which familiarity had been controlled.

The possibility that the texts used in the current study were biased toward the female participants was explored by having an additional group of male and female participants rate the texts in terms of their level of interest. Only two of the sixteen texts were found to be significantly more interesting to the female participants, and the two texts that were significantly more interesting to the female participants did not differ in any obvious way from the other items. That is, they were not overtly female-biased in their content. In fact, for both of these texts, most or all of the characters were male. Thus, interest level does not appear to influence the gender differences observed in this study.

6.8 IMPLICATIONS FOR FUTURE RESEARCH

The results of this investigation suggest several avenues for future research. First, because the sample size for the current study was based upon inaccurate effect size estimates, an aim of a future study could be to replicate the current study with estimates of the number of participants and items based on a more accurate prediction of the effect sizes that are likely to be obtained. This would involve taking into account factors that would reduce performance variability (e.g., homogeneity in vocabulary and reading level of participants). The effect sizes found in the current study could also serve as a basis for that estimate.

The replication study would differ from the current study in two other ways. The possibility would be explored that a greater number of target syntactic structures would increase the degree to which the frequency of these structures facilitates sentence processing. To accomplish this, texts of various lengths would be included in the experimental stimuli, and the extent to which the frequency of a syntactic structure facilitates the processing of that structure for various text lengths would be examined.

Another way in which the replication study would differ from the current study would be that various potential sources of inter-stimulus and inter-participant variability would be considered *a priori*, and would be incorporated into the initial design of the study. For example, to control for variability across participants' times, per phoneme reading times, rather than overall reading times, would be used as the dependent measure. In addition, efforts would be made to control for potential sources of variability across stimulus items, such as the number of occurrences of the matrix verbs across the experimental stimuli.

Second, once a replication study has been conducted, an additional direction for further investigation would involve determining whether any main effect or interaction involving syntactic priming found in the replication study can be generalized to sentences with a variety of syntactic structures. In addition, it should be determined whether the recent occurrence of sentences with similar (but not identical) syntactic structures can facilitate sentence processing. For example, it is unclear whether complex sentences with *that* complementizers, such as those used in this study, can facilitate processing of complex sentences with *wh-* subordinating conjunctions (e.g., *I don't know who they are*.).

Third, there is evidence that lexical and syntactic information presented within the sentence exerts a greater influence on the processing of verb subcategorizations following equi-biased verbs than on those following strongly S-biased or NP-biased verbs (Garnsey et al., 1997). Thus, it is possible that reading times for subcategorizations of equi-biased verbs are more susceptible to inter-sentential influences as well. Similarly, it is possible that the predictions for subcategorizations following strongly-biased verbs are influenced by verb bias to such an extent that the nature of the preceding discourse context adds little or nothing to such predictions. Therefore, it is suggested that future investigations examine the influences of coherence and frequency separately for verbs with different degrees of bias.

Fourth, efforts should be made to determine whether any results obtained from a replication study can be generalized to more naturalistic contexts. As stated earlier, the experimental texts included sentences with only two different verb subcategorizations. In this way, the experimental texts are much different than reading material that individuals typically encounter on a daily basis. To determine whether any significant influence of frequency or

coherence (or lack thereof) can be generalized to more typical reading situations should be a focus of further investigation.

Finally, the results of a replication study could lead to further investigations that involve individuals from clinical populations. For example, if the replication study indicates that the predictions for verb subcategorizations is significantly influenced by either coherence or by the frequency of previously occurring structures, the extent to which these factors influence predictions for verb subcategorizations made by individuals with language learning disabilities could also be investigated.

There is evidence that individuals with language learning disabilities tend to have difficulties integrating information across sentences within a text, and making inferences based on discourse-level information (Letts & Leinionen, 2001). This suggests that their predictions for upcoming syntactic structures might be influenced by discourse-level information to a lesser extent than might be the case for individuals with normal language ability. If this is found to be the case, perhaps reading remediation programs could incorporate instruction and practice in using these types of inter-sentential information to make predictions about upcoming syntactic structures.

This line of investigation could also focus on individuals who utilize augmentative and alternative communication (AAC) devices. As stated in the Review of the Literature, some recently developed AAC systems include prediction algorithms that allow these systems to make predictions about which words a user is likely to use next, based on the part of the message that has been constructed up to that point. These predictions are made based on sentence-level information. However, if the results of a replication study indicate that either the nature of the discourse context or syntactic priming might influence the prediction for upcoming syntactic

structures, incorporating this information into the algorithms used in these devices could improve their prediction accuracy as well.

Finally, the results of a better-controlled replication study or future investigation may suggest that there are gender-based differences in either the influence of syntactic priming or various types of discourse contexts on the prediction for upcoming syntactic structures. If this is the case, gender-based considerations could also be included in prediction algorithms for these devices, thereby making them more appropriate for users of different genders.

Appendix A: Summary tables for results of relevant research studies.

Table 1: The results of studies investigating the influence of verb subcategorization preference on the prediction for upcoming structures.

Study	Experiment	Method	Significant effect of verb subcategorization preference information on sentence parsing?		Significant effect of verb subcategorization information on first pass sentence parsing?	
Ferreira and Henderson (1990)	Experiment 1	Eye movements tracked during reading of temporarily ambiguous sentences.	NO	No significant effect of verb bias on total reading times in either the ambiguous ($p > .15$) or disambiguating regions ($p > .15$).	NO	No significant effect of verb bias on first pass reading times in either the ambiguous ($p > .20$) or disambiguating regions (Fs <1).
	Experiment 2	Word-by-word self-paced reading of temporarily ambiguous sentences; each word disappeared when the next word was presented.	YES	Significant interaction between verb type and complementizer presence in the postdisambiguating region ($p < .05$) and for total reading time ($p < .05$).	NO	No significant effect of verb bias on reading times in either the ambiguous (Fs < 1) or disambiguating regions (Fs < 1). There were also no significant interactions between verb bias and complementizer presence across these regions (Fs <1).
	Experiment 3	Word-by-word self-paced reading of temporarily ambiguous sentences; words remained on the screen until the entire sentence was visible.	YES	Significant effect of verb bias on the total reading time for sentences ($p < .05$). Sentences with nonminimal attachment verbs were read faster than those with minimal attachment verbs.	NO	No significant effect of verb bias on reading times in either the ambiguous (Fs < 1) or disambiguating regions (Fs < 1). No significant interaction between verb bias and complementizer presence (Fs <1).

Study	Experiment	Method	Significant effect of verb subcategorization preference information on sentence parsing?		Significant effect of verb subcategorization information on first pass sentence parsing?	
Boland (1993)	Experiment 1	Lexical decision and naming response times were measured following auditorily presented sentence fragments in which the final pronoun was either *congruent* or *incongruent* with the preferred subcategorization of the verb.	YES	Lexical decision was faster in the congruent than in the incongruent condition.	YES	Naming response times were faster in the congruent than in the incongruent condition.
Trueswell, Tanenhaus, and Kello (1993)	Experiment 1	Word naming task (naming pronoun) following auditorily presented sentence fragment that ended with either and NP-bias or S-bias verb. The pronoun was either a "good" or "bad" continuation of sentence. Fragment based on the subcategorization preference of the verb.	YES/ NO		YES/ NO	Significant interaction between verb type and case of pronoun ($p < .05$). Naming times for *him* were significantly faster following NP-bias verbs; naming times for *he* were significantly faster following S-bias verbs. Reliable simple effects for verb bias were found in trials were the accusative pronoun was named ($p < .01$). However, this effect was not seen in cases where the nominative pronoun was named ($Fs < 1$).
	Experiment 2	Word-by-word self-paced reading of temporarily ambiguous sentences; each word disappeared when the next word was presented.	YES	Main effect of verb bias across the noun (disambiguating) region was significant by subjects ($p < .05$), but not by items. In the final region, there was a significant interaction between verb bias and complementizer presence ($p < .01$). NP-bias verbs showed a large increase in reading times when the complementizer was present.	Did not test	

Study	Experiment	Method	Significant effect of verb subcategorization preference information on sentence parsing?		Significant effect of verb subcategorization information on first pass sentence parsing?	
Trueswell, Tanenhaus, and Kello (1993)	Experiment 3	Eye movements tracked during reading of temporarily ambiguous sentences	YES	For total reading time across all scoring regions, there was a significant main effect of verb bias ($p < .01$), as well as a significant interaction between verb bias and complementizer presence ($p < .01$). There was a significant main effect of verb type in the disambiguating region ($p < .05$).	NO	For first pass reading times (over the entire sentence), there was no significant effect of verb bias ($F < 1$). There was a significant effect of verb bias on first pass reading times across the final (post-disambiguating) region ($p < .05$).
Jennings, Randall, and Tyler (1997)		Word naming task (naming pronoun) following auditorily presented sentence fragment that ended with either and NP-bias or S-bias verb. The word was either a "good" or "bad" continuation of sentence fragment.	YES		YES	Naming times were significantly faster for both nominative and accusative pronouns when they were associated with the preferred subcategorization of the verb ($p < .05$).

Table 6: Summary of the results of studies investigating the influence of discourse context on syntactic parsing decisions.

Study	Experiment	Method	Type of Structural Ambiguity	Significant effect of discourse context?	
Altmann and Steedman (1988)	Experiment 1	Self-paced reading of sentences following visually-presented paragraph.	Prepositional phrases that were ambiguous between an NP and VP attachment	YES	Reading times for sentences in which the context supported the attachment bias of the sentence were significantly shorter than those that did not ($p < .005$).
	Experiment 2	Self-paced reading of sentences following visually-presented paragraph. Reading times were measured by region.	Prepositional phrases that were ambiguous between an NP and VP attachment	YES	Significant effect of context was found in the disambiguating region ($p = .0454$)
Britt, Perfetti, Garrod, and Rayner (1992)	Experiment 1	Self-paced reading of sentences following visually-presented paragraph.	Prepositional phrases that were ambiguous between an NP and VP attachment	YES	A significant difference in reading times between the NP and VP conditions that was found in the disambiguating region in the neutral context ($p < .001$). This was eliminated in the biasing context.
	Experiment 2	Monitored eye movements while participants read sentences that followed visually-presented paragraphs.	Prepositional phrases that were ambiguous between an NP and VP attachment	YES	A significant difference in first pass reading times between the NP and VP conditions that was found in the disambiguating region in the neutral context ($p < .01$). This was eliminated in the biasing context.
Britt, Perfetti, Garrod, and Rayner (1992)	Experiment 3	Self-paced reading of sentences either in isolation or following visually-presented Tparagraph.	Sentences were ambiguous between High VP attachment and reduced relative clause interpretation	NO	For the reduced relative clause/high VP comparison, reading times in the disambiguating region were significantly longer for the reduced relative conditions in both the isolation ($p < .006$) and context conditions ($p < .002$).

Study	Experiment	Method	Type of Structural Ambiguity	Significant effect of discourse context?	
Spivey and Tanenhaus	Experiment 1	Monitored eye movements while participants read target sentences.	Sentences were ambiguous between reduced relative clause and main clause reading. Reading times were compared to those of morphologically unambiguous reduced relative.	YES	Total reading times for the ambiguous region (the initial verb and *by* phrase regions) yield significant interactions between context and relative clause reduction ($p < .05$). First pass reading times for the *by* phrase region also showed a significant interaction between context and relative clause reduction ($p < .05$).
	Experiment 2	Monitored eye movements while participants read target sentences.	Sentences were ambiguous between reduced relative clause and main clause reading. Reading times were compared to those of syntactically unambiguous reduced relative.	YES	Total reading times for the ambiguous region (the initial verb and *by* phrase regions) yield significant interactions between context and relative clause reduction ($p < .05$). First pass reading times for the *by* phrase region also showed a significant interaction between context and relative clause reduction ($p < .05$).

Table 7: Summary of results of studies investigating the influence of discourse context on homonym activation.

Study	Experiment	Method	Significant Effect of Discourse Context?	
Vu, Kellas, Metcalf, and Herman (2000)	Experiment 1	Naming task following reading of paragraph	YES	Naming latencies for contextually appropriate targets were significantly faster for both dominant and subordinate passages than unrelated conditions ($p < .05$ for both). This was not true for the inappropriate targets ($p > .05$ for both).
	Experiment 2	Naming task following reading of paragraph; target word was only presented for 80 ms	YES	Naming latencies for contextually appropriate targets were significantly faster for both dominant and subordinate passages than unrelated conditions ($p < .05$ for both). This was not true for the inappropriate targets ($p > .05$ for both).
Hare, McRae, and Elman (2003)		Self-paced reading of sentences following visually-presented paragraph	YES	For the ambiguous condition, reading times were significantly slower following the DO-biasing context than after the SC-biasing context. This difference was not seen in the unambiguous condition.

Table 8: Studies investigating the influence of sentence and discourse context on predicted upcoming structures.

Study	Experiment	Methodology	Significant Effect of Discourse Context?		Effect Sizes	Significant Effect of Sentence-Level Context?		Effect Sizes
Schwanenflugel and White (1991)	Experiment 1	Lexical decision task following reading of paragraph	YES – but only when consistent with sentence-level context	Significant interaction of sentence and discourse-level context ($p < .05$)	Not available	YES	Significant main effect of sentence-level context ($p < .05$)	Not available
	Experiment 2	Lexical decision following reading of paragraph	YES (marginal)	Significant main effect of discourse context by subjects ($p < .05$), but not by items ($p > .10$)	Not available	YES (marginal)	Significant main effect of sentence-level context by subjects ($p < .05$), but not by items ($p > .10$)	Not available
	Experiment 3	Naming task following reading of paragraph	YES	Significant main effect of discourse context ($p < .05$).	Not available	YES (marginal)	Significant main effect of sentence-level context by subjects ($p < .05$), but not by items ($p > .10$)	Not available
Sharkey and Sharkey (1992)	Experiment 1	Lexical decision following first clause of visually presented sentence	Not tested			YES	Significant main effect for relatedness ($p < .025$)	Not available
	Experiment 2	Lexical decisions placed within visually presented paragraphs	YES	Significant main effect of context ($p < .01$)	Not available	Not tested		

Study	Experiment	Methodology	Significant Effect on Discourse Context?		Effect Sizes	Significant Effect of Sentence-Level Context?		Effect Sizes
Hess, Foss, and Carroll (1995)	Experiment 1	Naming task following auditorily presented sentences	Not tested			YES	Significant difference between the related condition and both unrelated ($p < .001$) and neutral ($p < .001$) conditions.	Difference between related and unrelated conditions ($\Delta = 1.988$); difference between related and neutral conditions ($\Delta = 2.203$)
	Experiment 2	Naming task following auditorily presented paragraphs	YES	Significant main effect for discourse-level context ($p < .001$).	Not available	YES	Significant main effect for sentence-level context ($p < .005$).	Not available
	Experiment 3	Naming task following auditorily presented paragraphs	YES	Significant differences between both discourse-level related conditions and control condition ($p < .05$ for both)	Δs: GRLR vs control = 3.784; GRLU vs control = 2.162	NO	No significant difference between the GULR condition and the control condition ($p > .05$).	$\Delta = 1.699$
	Experiment 4	Naming task following auditorily presented paragraphs	YES	Significant differences between both discourse-level related conditions and control condition ($p < .05$ for both)	Δs: GRLR vs control = 8.051; GRLU vs control = 6.215	NO	No significant difference between the GULR condition and the control condition ($p > .05$).	$\Delta = .424$

Study	Experiment	Methodology	Significant Effect on Discourse Context?		Effect Sizes	Significant Effect of Sentence-Level Context?		Effect Sizes
Hess, Foss, and Carroll (1995)	Experiment 5	Naming task following auditorily presented paragraphs	YES	Significant differences between both discourse-level related conditions and control condition ($p < .05$ for both).	Δs: GRLR v control = 4.462; GRLU v control = 2.716	NO	No significant difference between the GULR condition and the control condition ($p > .05$).	Δ = -.194
	Experiment 6	Naming task following auditorily presented paragraphs	YES	Significant differences between both discourse-level related conditions and control condition ($p < .05$ for both)	Δs: GRLR v control = 5.709; GRLU v control = 4.880	NO	No significant difference between the GULR condition and the control condition ($p > .05$).	Δ = -.184
	Experiment 7	Naming task following auditorily presented paragraphs	YES	Significant differences between both discourse-level related conditions and control condition ($p < .05$ for both)	Δs: GRLR v control = 5.657; GRLU v control = 4.326	NO	No significant difference between the GULR condition and the control condition ($p > .05$).	Δ = .333
	Experiment 9	Naming task following auditorily presented paragraphs	YES	Significant differences between both discourse-level related conditions and control condition ($p < .05$ for both)	(Δs: GRLR v control= 2.977; GRLU v control = 1.874	NO	No significant difference between the GULR condition and the control condition ($p > .05$).	Δ = -2.095

Study	Experiment	Methodology	Significant Effect on Discourse Context?		Effect Sizes	Significant Effect of Sentence-Level Context?		Effect Sizes
Cook and Myers (2004)	Experiment 1	Eye tracking for target words within visually-presented paragraphs	YES	Significant interaction between appropriateness and consistency with context in both target ($F(1, 30) = 5.37$) and post-target ($F(1, 30) = 5.96$) regions for first pass reading times during second encounter.	Not available	Not tested		
	Experiment 2	Eye tracking for target words within visually-presented paragraphs	YES	Significant interaction between appropriateness and justification in the target region ($F(1, 23) = 6.52$.	Not available	Not tested		

Table 9: Summary of studies on production to production syntactic priming among adjacent sentences.
* Luka and Barsalou (2005; Experiments 4 & 5) addressed production to production syntactic priming. However, within the text, they are included with the other experiments in their study in the section on comprehension to comprehension syntactic priming.

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect size
Bock (1986)	Experiment 1	Active and passive transitives; prepositional and double object datives	Participant repeated sentences of various target structure. They described a picture following each repeated sentence.	YES	Significant priming for both datives ($p <. 05$) and transitives ($p <.05$). The influence was greater for datives than transitives.	Not available
	Experiment 2	Active and passive transitives	Same as for Experiment 1	YES/NO	Significant priming for passive but not active constructions.	Not available
	Experiment 3	Active, passive transitives	Similar to Experiment 1, except that participants were encouraged to more deeply process picture through ongoing recognition task.	YES/NO	Significant priming for passive but not active constructions.	Not available
Bock and Loebell (1990)	Experiment 1	Prepositional and double object datives, and prepositional locatives	Same as for Bock (1986; Experiment 1)	YES	Both prepositional datives and prepositional locatives primed prepositional dative constructions. Double object sentences primed double object constructions. Differences that exceeded the 95% confidence interval were deemed as significant.	Not available
	Experiment 2	Active, passive, and locative transitives	Same as for Experiment 1	YES/NO	Both passive and locatives sentences primed passive picture descriptions. Active primes did not have a significant influence on the proportion of active sentences.	Not available

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect size
Bock and Loebell (1990)	Experiment 3	Prepositional and double object datives, and infinitive forms	Same as for Experiment 1	YES	Prepositional object primes yielded a greater number of prepositional object sentences relative to the other two types of primes.	Not available
Hartsuiker and Kolk (1998)	Experiment 1	Active and two types of passive sentences; prepositional, double object and medial datives	Participants read aloud sets of three priming sentences and then described a picture.	YES/NO	Significant priming for both types of passives ($p < .02$ for both), but not for actives. For the datives, significant priming for double object and medial datives ($p < .05$ for both), but not for prepositional datives.	Not available
	Experiment 2	Active and passive transitive sentences; prepositional, double object, and medial datives.	Similar to those from Experiment 1	YES/NO	No significant priming effect for passives. A negative priming effect for actives. Significant priming was found for prepositional object and medial datives ($p < .02$ for both), but not for direct object datives.	Not available
	Experiment 3	Active transitive sentences; prepositional object, double object datives	Participants listened to priming sentences and repeated them. Then, they described a picture.	YES/NO	There was a significant priming effect for both types of dative ($p < .05$ for both), but for neither of the transitive forms.	Not available
Pickering and Branigan (1998)	Experiment 1	Prepositional object or double object datives	Participants wrote completions to pairs of prime and target sentence fragments.	YES	Participants produced significantly more target completions following the primes of the same type ($p < .01$). This difference was more marked when the same verb	Not available

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect size
Pickering and Branigan (1998)	Experiment 2	Prepositional object or double object datives	Participants wrote completions to two prime fragments, and then to one target sentence fragment.	YES	Participants produced significantly more target completions following the primes of the same type ($p <$. 01), even when different verbs were used in the priming and target sentences. Analysis of simple effect revealed a when different verbs were used in the priming and target sentences. Analysis of simple effect revealed a significant priming effect for PO construction, but only a marginal effect for DO constructions	Not available
	Experiment 3	Prepositional object or double object datives	Participants wrote completions to pairs of prime and target sentence fragments. The tense of the prime and target verbs were either the same or different.	YES	Participants produced significantly more target completions following the primes of the same type ($p <$. 01). No significant main effect for tense or significant interactions with the tense variable ($p > .05$).	Not available
	Experiment 4	Prepositional object or double object datives	Participants wrote completions to pairs of prime and target sentence fragments. The aspect of the prime and target verbs were either the same or different.	YES	Participants produced significantly more target completions following the primes of the same type ($p <$. 01). The interaction of prime, target and aspect was not significant by participants or by items ($p >.05$).	Not available

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect size
Pickering and Branigan (1998)	Experiment 5	Prepositional object or double object datives	Participants wrote completions to pairs of prime and target sentence fragments. The number of the prime and target verbs were either the same or different.	YES	Participants produced significantly more target completions following the primes of the same type ($p < .01$). The interaction of prime, target and number was not significant by participants or by items ($p > .05$).	Not available
Potter and Lombardi (1998)	Experiment 1	Prepositional object and double object datives	Participants read sentence with target structure, and priming sentences, and then recalled target sentence.	YES	There was a significant main effect of the form of the priming sentences on the recall of the target sentences ($p < .001$).	Not available
* Luka and Barsalou (2005)	Experiment 4	Various structures that had been previously rated for grammaticality.	Participants rated repeated sentences in which varying numbers of types or tokens had been presented earlier.	YES	Participants' grammaticality ratings were significantly higher when either identical or structurally-related items had been presented earlier ($p < .01$). This influence was greater for moderately grammatical than for highly grammatical sentences.	For identical repetition, $\Delta = 2.615$; for structural repetition, $\Delta = .875$.
	Experiment 5	Various structures that had been previously rated for grammaticality.	Participants read aloud four identical tokens of each of several sentence types. Then, they rated identical, structurally-related, or unrelated sentences for grammaticality.	YES	Participants' grammaticality ratings were significantly higher when either identical or structurally-related items had been presented in earlier ($p < .01$). The ratings of highly and moderately grammatical sentences were similarly influenced by prior exposure.	For identical repetition, $\Delta = 1.846$, for structural repetition, $\Delta = 1.706$

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect size
Cleland and Pickering (2006)	Experiment 1	Prepositional object and double object datives	Participants wrote completions to pairs of prime and target sentence fragments that either were of the same or of a different modality (spoken vs written)	YES	Significant effect of syntactic priming of target structures across conditions ($p < .01$). No significant influence for same versus different modality.	Not available
	Experiment 2	Prepositional object and double object datives	Participants wrote completions to pairs of prime and target sentence fragments that either were of the same or of a different modality (spoken vs written).	YES	Significant effect of syntactic priming of target structures across conditions ($p < .01$). No significant influence for same versus different modality. Influence was of same versus different verb was also examined and found to be marginal ($p < .06$).	Not available
	Experiment 3	Prepositional object and double object datives	Participants wrote completions to pairs of prime and target sentence fragments that either were of the same or of a different modality (spoken vs written).	YES	Significant effect of syntactic priming of target structures across conditions ($p < .01$). No significant influence for same versus different modality.	Not available

Table 10: Summary of studies on local comprehension to comprehension syntactic priming

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect size
Frazier, Taft, Roeper, Clifton, & Ehrlich (1984)		Conjoined clause sentences with clause with the same or different constructions	Self-paced reading	YES	Significant main effect for parallel form.	Not available
Branigan, Pickering & Stewart (1995)		Early closure/late closure, reduced relative/main clause, complement clause/relative clause ambiguous sentences.	Self-paced reading	YES	Participants' reading times were faster for sentences when they followed a sentence with the same structure.	Not available
Noppeney and Price (2004)		Early closure/late closure, reduced relative/main clause ambiguous sentences	Self-paced reading	YES	Reading times were faster following sets of sentences with the same syntactic structure (p <.05).	Δ = .417
Luka and Barsalou (2005)	Experiment 1	Various structures that had been previously rated for grammaticality	Participants rated repeated and nonrepeated sentences for grammaticality.	YES	Participants gave significantly higher grammaticality ratings to sentences that they had read earlier in the experiment (p <.001).	Δ = 8.462
	Experiment 2	Various structures that had been previously rated for grammaticality	Participants rated sentences for grammaticality after being exposed to structural and positional variants of the sentences.	YES	Exposure to structural variants of a target sentence significantly influenced participants' grammaticality ratings (p <.05), but exposure to positional variants did not p > .05).	For influence of exposure to structural variants, Δ =1.667

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect size
Experiment 3	Various structures that had been previously rated for grammaticality	Participants rated repeated and nonrepeated sentences for grammaticality.	YES	Structural repetition was more effective than identical repetition in influencing grammaticality ratings ($p < .005$). Structural repetition was effective for moderately grammatical ($p < .0001$), but not for highly grammatical sentence.	For structural repetition (overall), $\Delta = 4.20$. For identical repetition (overall), $\Delta = 0.222$. Structural repetition following moderately grammatical sentences, $\Delta = 6.727$.	

Table 11: Summary of studies on local comprehension to production syntactic priming
* Potter and Lombardi (1998; Experiment 3) was a study of production to production syntactic priming, but was included here for ease of comparison with Experiment 2.

Study	Experiment	Structure(s) studied	Method	Significant priming effect?	
Potter and Lombardi (1998)	Experiment 2	Prepositional object and double objective datives.	Participants silently read sentences and then recalled them. Prime was heard, but not recalled before the target.	YES/NO	Significant effect of prime type was found on the production of the alternate form ($p < .01$).
	Experiment 3*	Prepositional object and double objective datives.	Participants silently read sentences and then recalled them. Prime was heard and recalled before the target.	YES/NO	Significant effect of prime type was found on the production of the alternate form ($p < .01$).
Branigan, Pickering, and Cleland (2000)		Prepositional object and double objective datives.	Confederate scripting	YES	Significantly more responses of each target structure following primes of the same type ($p < .01$ for each).
Cleland and Pickering (2003)	Experiment 1	Pre-nominal and relative clause constructions	Confederate scripting	YES	Significantly more responses of each syntactic structure occurred after primes of the same type ($p < .001$). Magnitude of priming was greater when the prime and target used the same noun ($p < .05$), but was only marginal between the same versus different adjective conditions ($p < .08$)
	Experiment 2	Pre-nominal and relative clause constructions	Confederate scripting	YES	Significantly more responses of each syntactic structure occurred after primes of the same type ($p < .001$). Magnitude of priming was greater when noun were semantically related than when they were unrelated ($p < .05$). There was only a marginal difference between the same noun and semantically related nouns conditions ($p < .08$).
	Experiment 3	Pre-nominal and relative clause constructions	Confederate scripting	YES	Significantly more responses of each syntactic structure occurred after primes of the same type ($p < .001$). No significant difference between phonologically related versus unrelated primes ($p > .05$), but a significant difference between same versus phonologically related primes $p < .05$).

Table 12: Summary of studies on cumulative syntactic priming

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect Size
Bock and Griffin (2000)	Experiment 1	Simple passive and prepositional dative	Participants read sets of sentences, and then described a picture. There were 0, 1, or 2 sentences between the prime sentence and target picture.	YES	A significant effect of priming was found for all three lag conditions.	Not available
	Experiment 2	Simple passive and prepositional dative	Same as Experiment 1, except that there were 0, 4, or 10 sentences between the prime sentence and target picture.	YES/NO	A significant effect of priming was found for the 0-lag and 10-lag conditions, but not for the 4-lag condition.	Not available
Kashak and Glenberg (2004)	Experiment 1	The *needs* construction	Participants read sets of sentences with either the *needs* construction or a standard construction, and reading times across trials was measured.	YES	Significant increase in reading speed across trials of the *needs* construction ($p < .05$).	For comparison between trials of *needs* construction 1-4 and trials, 5-8, $\Delta = 4.108$; for comparison between trials 1-4 and trials 9-12, $\Delta = 5.383$.
	Experiment 2	The *needs* and *wants* construction	First, participants read sets of sentences with either the *needs* construction or a standard construction. Then, they read series of sentences with a *wants* construction. Reading times across a disambiguating region were measured.	YES	Significant increase in reading speed across trials of the *needs* construction ($p < .05$). Participants in the *needs* condition read the *wants* sentences more quickly than those in the control condition ($p < .05$).	Not available

Study	Experiment	Structure(s) studied	Method	Significant priming effect?		Effect Size
Kashak and Glenberg (2004)	Experiment 3	A modifier construction	First, participants read sets of sentences with either the *needs* construction or a standard construction. Then, they continued to read the same construction, but also read the modifier construction.	YES	During the initial trials, reading times on the modifier construction were faster for participants who had been reading the *needs* construction ($p < .05$).	Not available
	Experiment 4	A modifier construction	Participants read the *needs* construction and then read sentences with the modifier construction. However, one group received instructions to decrease likelihood that they would consider the *needs* construction.	N/A	The group that received instruction on the needs construction slowed down (marginally, $p < .06$) when the modifier construction was introduced. The other group sped up slightly, but not significantly.	Not available
Kashak, Loney, and Borrenggine (2006)	Experiment 1	Prepositional and double object datives	Participants completed sets of sentence fragments that were intended to generate either or both of the target constructions as completions. The proportion of exposure to the construction types was varied, as well as the recency between the primes and the target sentence. Following each set, participants wrote a completion to sentence fragment.	YES	A significant difference was found in proportion s of target stem completions among frequency conditions ($p < .004$), but not among the two conditions that differed only in the recency of exposure to the target structure.	For comparison of DO completions between EE and UE conditions, $\Delta = .258$. For comparison of PO completions between EE and UE conditions, $\Delta = .563$.
	Experiment 2	Prepositional and double object datives	Same as Experiment 1, except for different frequency conditions and a greater number of priming fragments.	YES	There was a significant relationship between the frequency of target structures and proportion of target responses produced by the participants ($p < .03$).	For comparison of DO completions between EE-B and UE-75 conditions, $\Delta = .241$, between EE-B and UE-100 conditions, $\Delta = .241$. For comparison of PO completions between EE-B and UE-75 conditions, $\Delta = .538$, between EE-B and UE-100 conditions, $\Delta = .654$.

Appendix B: Recruitment Flyer

Language Comprehension Study

Are You?

- Between the ages of 20-35?
- A native speaker of English?
- Interested in trying an experimental task (requiring approximately 2 hours of your time)?

If So, Contact:

Jill L. Brady
Assistant Professor
Speech-Language Pathology Program
(412) 414-3098

Task requires approximately 2 hours

Appendix C: Script for Recruitment of Participants

I'm recruiting people between 20 and 35 years old for a research project. The purpose of this project is to examine participants' reading of sentences under varying conditions. Participants will be screened for normal hearing, vision, and reading ability. They must also have no history of language or learning disability and must be native speakers of English. They will read 60 short passages (between 2 and 20 sentences in length) and answer a total of 20 questions that will be asked randomly throughout the experiment. The entire session will take approximately 2 hours.

Testing will be done in the Department of Special Education and Clinical Services (in Davis Hall) at the Indiana University of Pennsylvania.

For addition information, please contact Jill Brady at: (412) 414-3098.

Appendix D: Background Questionnaire

Participant # _____

BACKGROUND QUESTIONNAIRE

Age:
Gender:
Circle one for each of the following:

Is American English your native language?
Yes No

Are you fluent in any language other than English?
Yes No

Have you ever been diagnosed with a learning disability (e.g., dyslexia, reading disability, language learning disability, central auditory processing disorder)?
Yes No

Did you ever receive special education or resource services, tutoring for language or reading difficulties, or speech-language therapy?
Yes No

Have you ever suffered a traumatic brain injury (TBI)?
Yes No

Do you have any physical limitations that might affect your ability to press buttons with your dominant hand (e.g., weakness or paralysis)?
Yes No

Appendix E: Selected verbs and their characteristics

Verb	Sense	# of websites surveyed	%NP	%S	%PP	% zero subcat.	% Inf "to"	%Other	%NP-%S	"that" proportion[1]	Classification
Admit	confess to be true or to be the case	100	35.38%	49.23%	6.15%	2.05%	7.18%	0.00%	-13.85%	63-74%	Equi-biased
Announce		100	64.73%	26.34%	0.00%	1.79%	0.00%	7.14%	38.39%	55-69%	NP-biased
Decide		50	17.70%	47.79%	4.42%	10.62%	19.47%	0.00%	-30.09%	67%	S-biased
Expect	Regard (something) as likely to happen	50	30.84%	41.12%	11.21%	0.00	15.87%	0.93%	-10.28%	50%	Equi-biased
Explain		100	59.31%	29.00%	6.49%	5.19%	0.00%	0.00%	-30.31%	58%	NP-biased
Guarantee		50	69.34%	23.36%	0.00%	3.65%	3.65%	0.00%	45.99%	62%	NP-biased
Indicate1	point out; show; be a sign or symptom of; strongly imply	50	32.89%	63.09%	0.00%	1.34%	0.00%	2.68%	-30.20%	64-69%	S-biased
Know1	be aware of through observation, inquiry, or information	50	36.07%	40.44%	6.01%	15.30%	1.64%	0.55%	-4.37%	37%	Equi-biased
Maintain3	state something strongly to be the case; assert	50	3.81%	96.19%	0.00%	0.00%	0.00%	0.00%	-92.38%	60-66%	S-biased
Predict		100	74.59%	21.31%	0.00%	1.64%	0.00%	2.46%	53.28%	50-63%	NP-biased
Prove		50	43.42%	48.03%	2.63%	3.95%	0.66%	1.32%	-4.61%	60%	Equi-biased
Realize1	to become fully aware of (something) as a fact	100	16.55%	81.30%	0.72%	1.44%	0.00%	0.00%	-64.75%	60-65%	S-biased
Say1	utter words so as to convey information, an opinion, a feeling or intention, or an instruction	50	34.94%	41.87%	0.006024	0.90%	0.00%	21.69%	-6.93%	71%	Equi-biased
Sense		100	32.22%	61.11%	0.56%	5.56%	0.00%	0.56%	-28.89%*	71%	S-biased
Show	display or allow to be perceived	50	55.56%	42.59%	0.93%	0.00%	0.00%	0.93%	12.97%	58%	Equi-biased
Suggest		50	59.80%	22.55%	0.058824	3.92%	6.86%	0.98%	37.25%	68%	NP-biased

[1]the percentage of sentential complement continuations that included a complementizer *that*, as determined by three studies: a.) Garnsey et al., 1997; b.) Jennings et al., 1997; Trueswell et al., 1993).

.

Appendix F: Websites and verb sense tokens that were not included in subcategorization frequency counts.
Table 13: Types of websites that were not included in subcategorization frequency counts

Types of websites that were excluded	Reason for exclusion
Online dictionaries	Online dictionaries typically provide examples that are chosen to clarifying the meaning and use of the verb, and not to communicate the information within the sentence. Therefore, it is unclear whether the frequency of various subcategorizations chosen for this purpose is the same as for those chosen for the purpose of conveying information.
Websites that provide information to English as a second language learners	These websites provide examples of English sentences to provide information to learners of English as a second language. Because these sentences are examples of English syntax and vocabulary, the verb subcategorization frequencies found within them may not be similar to those in "naturally occurring" English sentences.
Websites comprised of computer programming code	It is unclear whether the use of a verb and the proportion of various subcategorizations which follow it are similar in computer programming syntax as in English syntax.
Repeated websites	If a search resulted in a web page being retrieved more than once, repeated occurrences were not considered. This was intended to eliminate any overrepresentation of the subcategorization frequencies found on a particular website.

Table 14: Verb sense tokens that were not included in verb subcategorization frequency counts.

Verb sense tokens that were excluded	Reason for exclusion
Verbs in abandoned utterances	In abandoned utterances, it is not possible to determine with certainty the intended syntactic structure.
Verbs in repeated sentences and heading (such as in Really Simple Syndication (RSS) feeds and web logs)	Counting the same sentence more than once for repeated headings (which are common on the Internet) would greatly inflate the number of subcategorizations of a particular type.
Verbs in passive sentences	There is no known evidence that verb subcategorization frequency in passive sentences is the same as those in active sentences. Because all of the target sentences in the experiment will use an active construction, examining only active sentences eliminates this potential confound.
Verbs in sentences that were not written in modern English.	Because the subcategorization frequencies associated with particular verbs may have changed over time, verb subcategorizations found in texts representing non-modern English (e.g., excerpts from Chaucer's *The Canterbury Tales*) were not included in the frequency counts.

Appendix G: Experimental Stimuli

Set #1

6S/Coherent Discourse Context

Harvey noted a woman who entered his law office. She explained she needed a lawyer to take her case. She recalled that her table was ruined while in storage. According to the woman, the owner of the facility said it was damaged by water from a leaking pipe. However, the owner denied responsibility due to a clause in her contract. The woman insisted the clause did not cover this case. Harvey requested that she show him the contract. Reading it, Harvey noticed that the contract was clear on this issue.

Target: Harvey indicated the owner was not liable for the damage.

Comprehension Questions: Did the woman believe that the facility owner should pay for the damage?

2S/Coherent Discourse Context

Harvey noted a woman who entered his law office. Approaching Harvey, she explained her need for a lawyer. She recalled that her table was ruined by water damage while it was in storage. The owner of the storage facility found a leaking pipe that had caused the damage. Yet, he denied responsibility due to a clause in her contract. However, the woman doubted the relevance of the clause this case. Harvey requested a chance to view the contract. Reading it, Harvey noticed that the contract was clear on this issue.

Target: Harvey indicated the owner was not liable for the damage.

Comprehension Question: Did the woman believe the owner should pay for the damage?

6S/Scrambled Sentences

Harvey noted a woman who entered his law office. As Harvey approached, the woman asked him whether he liked the flowerbox. They decided they should find the cavern and go on a tour. Finally, the evening before the tournament, Harvey realized his friend had good reason to warn him. Angered, Harvey maintained his pie was better. They suggested the fish may have moved to another location. He guessed that they had plenty of time to get back that day. Reading it, Harvey noticed that the contract was clear on this issue.

Target: Harvey indicated the owner was not liable for the damage.

Comprehension Question: Did Harvey work in a law office?

2S/Scrambled Sentences

Harvey noted a woman who entered his law office. He believed that this would allow him to share his views. Harvey doubted the woman's theory. He informed the woman that he was auditioning for the same role. Then, Harvey explained his lack of willingness to buy the car without them. However, Harvey doubted the dog's ability to move the vase. However, Harvey mentioned his busy work schedule. Reading it, Harvey noticed that the contract was clear on this issue.

Target: Harvey indicated the owner was not liable for the damage.

Comprehension Question: Did Harvey work in a law office?

Set #2

6S/Coherent Discourse Context

Erin and Michelle, two young graphic artists, each hoped they would be hired for a job at a greeting card company. Michelle boasted that she had five years of experience. As a recent

graduate, Erin worried she could not get the job. She asked Michelle, if she would look at her portfolio, and give her feedback. Looking through it, Michelle indicated her approval. Then, Erin said she would like to see Michelle's portfolio. Michelle explained she did not bring one, since no one would question her talent. Coming through the door, the secretary advised that both of them have their portfolios ready.

Target: Michelle expected other interviews would require a portfolio.

2S/Coherent Discourse Context

Erin and Michelle, two young graphic artists, each hoped they would be hired for a job at a greeting card company. Waiting for their interviews, Michelle mentioned her five years of experience to Erin. At this, Erin doubted her chances of getting the job. She asked Michelle, if she would look at her portfolio, and give her feedback. Looking through it, Michelle indicated her approval. Then, Erin stated her wish to view Michelle's portfolio. Michelle doubted the need for them to see hers, since she felt her resume could stand on its own. Coming through the door, the secretary advised that both of them have their portfolios ready.

Target: Michelle expected other interviews would require a portfolio.

6S/Scrambled Sentences

Erin and Michelle, two young graphic artists, each hoped they would be hired for a job at a greeting card company. They decided they should stop to see what there was. As they looked at the elephants, she observed that Erin did not seem to be interested. However, Erin doubted his explanation. As she went into the room, she noticed that all three boys had pushed their desk beside hers. However, they maintained their confidence that the bus would arrive. She indicated the price shown on the bottom of the lamp was fifty dollars. Coming through the door, the secretary advised that both of them have their portfolios ready.

Target: Michelle expected other interviews would require a portfolio.

2S/Scrambled Sentences

Erin and Michelle, two young graphic artists, each hoped they would be hired for a job at a greeting card company. According to Michelle, she claimed the prize at every baking contest that she entered. Suddenly, from the audience, a man disputed this claim. He revealed his idea to camp in the woods that evening and to find their way back the next day. At eight o'clock, the friends noted the fact that they were still waiting for their bus. They promised him assistance in looking for the ski, if he would wait until the next morning. However, she doubted the relevance of the clause this case. Coming through the door, the secretary advised that both of them have their portfolios ready.

Target: Michelle expected other interviews would require a portfolio.

Set #3

6S/Coherent Discourse Context

As the grand finale for his show, a magician announced his plan to escape from a locked crate that was suspended in the air. Further, he claimed he would be the first person to do this. Suddenly, from the audience, a man said he had seen another magician do the same stunt. The magician argued that this was not possible. He suggested the other magician had only created an illusion. He insisted he was the only magician who could do this. However, the audience member doubted his explanation. He demanded the magician prove that no one else could perform this stunt.

Target: The magician maintained his act was unique.

Comprehension Question: Did the audience member believe the magician's claim?

2S/Coherent Discourse Context

As the grand finale for his show, a magician announced his plan to escape from a locked crate that was suspended in the air. Further, he mentioned that no other magician had ever done this stunt. Suddenly, from the audience, a man disputed this claim. He recalled another magician who had done the same stunt. The magician denied the truth of the man's statement. He explained the way in which some other magicians create an illusion of doing the stunt. However, the audience member doubted this explanation. He demanded the magician prove that no one else could perform this stunt.

Target: The magician maintained his act was unique.

Comprehension Question: Did the audience member believe the magician?

6S/Scrambled Sentences

As the grand finale for his show, a magician announced his plan to escape from a locked crate that was suspended in the air. He warned his father who approached the bed to stay by the door. He recalled that his table was ruined while in storage. Walking through the cave, the magician discovered that there was an attractive rock near him. However, the magician demanded they let him go. Then, the magician said he would like to see the audience member's portfolio. In addition, he predicted the magician would have many new friends by the end of the day. He demanded the magician prove that no one else could perform this stunt.

Target: The magician maintained his act was unique.

Comprehension Question: Was the magician's table in good condition after being in storage?

2S/Scrambled Sentences

As the grand finale for his show, a magician announced his plan to escape from a locked crate that was suspended in the air. When they asked him about this, he suggested they remain

patient. However, he mentioned his busy work schedule. But, he suspected his friend of being jealous of his ability. However, he explained his confidence that the bus would arrive. However, at the lions' cage, he mentioned his lack of interest. In response, he requested any evidence that he might have of the man's guilt. He demanded the magician prove that no one else could perform this stunt.

Target: The magician maintained his act was unique.

Comprehension Question: Was the magician interested in the lions?

Set #4

6S/Coherent Discourse Context

At the baking competition, a panel of chefs judged an array of pies, cakes, and other baked goods. Audrey boasted that her peach pie would certainly win first prize. Trying Audrey's pie, they confirmed that her pie was very good. However, they explained many of the other contestants could bake just as well. Angered, Audrey maintained her pie was better. The judges suggested she try another contestant's entry before making such a claim. Audrey accepted their suggestion, and tried a piece of apple pie. As she ate the pie, the judges noted the expression on Audrey's face changed.

Target: Audrey admitted the other contestant's entries were also good.

2S/Coherent Discourse Context

At the baking competition, a panel of chefs judged an array of pies, cakes, and other baked goods. According to Audrey, she claimed the prize at every baking contest that she entered. Trying Audrey's pie, they confirmed the quality of her baking. However, they explained the high quality of some of the other entries. Angered, Audrey maintained her pie was the best one. The judges advised her to try one of the other items before making such a claim. Audrey

accepted their advice, and tried a piece of apple pie. As she ate the pie, the judges noted the expression on Audrey's face changed.

Target: Audrey admitted the other contestant's entries were also good.

6S/Scrambled Sentences

At the baking competition, a panel of chefs judged an array of pies, cakes, and other baked goods. Audrey claimed an enormous fish had lived there for the past several years. Looking under the bed, the judges denied they saw a monster. Audrey's friends warned her that it would be dangerous to leave the cabin so late in the day. According to the Audrey, the judges said it was damaged by water from a leaking pipe. She suspected the judges were giving advice rather than a rule. A few minutes after the judges went to bed, Audrey discovered that her pencil was dull, and sharpened it. As she ate the pie, the judges noted the expression on Audrey's face changed.

Target: Audrey admitted the other contestant's entries were also good.

2S/Scrambled Sentences

At the baking competition, a panel of chefs judged an array of pies, cakes, and other baked goods. Stopping at one dealership, Audrey noticed a car that she liked. She recalled another magician who had done the same stunt. However, one of her friends warned Audrey that she was being too confident. Audrey explained her plan to try out for the role of the police officer. However, the judges guessed that they had been walking in circles. They noticed Audrey's disappointment. As she ate the pie, the judges noted the expression on Audrey's face changed.

Target: Audrey admitted the other contestant's entries were also good.

Set #5

6S/Coherent Discourse Context

One weekend, Wayne and Becky decided they would take a canoe trip. Traveling to the stream, Wayne confessed that he had never been in a canoe before this. He explained he was a bit concerned about the insects that they might encounter on the trip. Opening his bag, he revealed many bottles of insect repellant and sunscreen that he had bought for the trip. Becky said this might be more than they need for the trip. Arriving at the stream, Becky noted Wayne appeared be nervous. She asked him if there was a problem. He announced he was a bit afraid of deep water.

Target: Becky suggested the canoe trip might not be a good idea.

Comprehension Question: Did Becky have less canoeing experience than Wayne?

2S/Coherent Discourse Context

One weekend, Wayne and Becky decided they would take a canoe trip. Traveling to the stream, Wayne confessed his lack of canoeing experience. He explained his concern that there would be many insects in and around the stream. Opening his bag, he revealed many bottles of insect repellant and sun screen that he had bought for the trip. Becky advised Wayne that he did not need to pack so much for the trip. Arriving at the stream, Becky noted Wayne's nervousness. She asked him if there was a problem. He announced he was a bit afraid of deep water.

Target: Becky suggested the canoe trip might not be a good idea.

Comprehension Question: Did Becky have less canoeing experience than Wayne?

6S/Scrambled Sentences

One weekend, Wayne and Becky decided they would take a canoe trip. During one oboe lesson, Wayne mentioned that there was a girl who was in his class who wanted oboe classes. He

explained he made it larger than usual, to hold more flowers. However, Becky denied responsibility due to a clause in her contract. While purchasing their tickets, they asked the vendor about the bus schedule. He remembered that two of his friends were professional golfers. She said she thought that Wayne would enjoy kindergarten. He announced he was a bit afraid of deep water.

Target: Becky suggested the canoe trip might not be a good idea.

Comprehension Question: Was Wayne was afraid of deep water?

2S/Scrambled Sentences

One weekend, Wayne and Becky decided they would take a canoe trip. At this, Becky doubted her chances of getting the job. As he pulled out his camera, Becky advised him to follow the guide's rule about not taking pictures. Further, she recalled a blueprint that she gave him. Becky asked him if they could go the next weekend. As he gestured towards the book, Wayne forgot the drink he had his hand. She mentioned the many enjoyable things that children do in kindergarten. He announced he was a bit afraid of deep water.

Target: Becky suggested the canoe trip might not be a good idea.

Comprehension Question: Did Wayne have a drink in his hand?

Set #6

6S/Coherent Discourse Context

Returning home from work, Anne noticed that her favorite vase was broken into several pieces. Immediately, she suspected her young son, Matthew, of breaking it. Calling Matthew, Anne demanded he explain how the vase was broken. Matthew claimed he did not know what had happened to the vase. He proposed that the dog might have knocked over the vase with its tail. However, Anne doubted his explanation. Yet, Matthew maintained he was not guilty of

breaking the vase. Sensing Matthew's fear, Anne swore she would not be angry with him, if he would tell the truth.

Target: Matthew explained the broken vase was his fault.

2S/Coherent Discourse Context

Returning home from work, Anne noticed that her favorite vase was broken into several pieces. Calling her young son, Matthew, Anne demanded an explanation for the broken vase. Matthew denied knowledge of what happened to the vase. He proposed the possibility that the dog broke the vase. However, Anne doubted the dog's ability to move the vase. Matthew admitted the doubtfulness of his idea. Yet, he maintained his innocence. Anne swore she would not punish him, if he would tell the truth.

Target: Matthew explained the broken vase was his fault.

6S/Scrambled Sentences

Returning home from work, Anne noticed that her favorite vase was broken into several pieces. After an hour of waiting, she noticed that there was a lemonade stand nearby. Anne suggested they return the way the came. However, she mentioned that many of the other contestants could bake just as well. Anne asked Matthew, if she could keep a light on in the room to study. She asked Matthew, if he would look at her portfolio, and give her feedback. Anne argued that she could not buy the car without this information. Sensing Matthew's fear, Anne swore she would not be angry with him, if he would tell the truth.

Target: Matthew explained the broken vase was his fault.

2S/Scrambled Sentences

Returning home from work, Anne noticed that her favorite vase was broken into several pieces. During an oboe lesson, Matthew mentioned a girl in his class who wanted lessons.

However, Anne stated her idea that flower boxes should be small. Later that night, he remembered a competitor's history as a golf champion. Anne advised him to try one of the other items before making such a claim. Finding the seller, he stated his desire to buy the lamp. Anne informed Matthew that she was auditioning for the same role. Anne swore she would not punish him, if he would tell the truth.

Target: Matthew explained the broken vase was his fault.

Set #7

6S/Coherent Discourse Context

As he returned from skiing with friends, Don discovered that one of his skis was missing. Although it was getting dark and snowing heavily, he decided he should go back and find the ski before it was covered with snow. His friends warned him that it would be dangerous to leave the cabin so late in the day. They promised they would help him look for his ski in the morning. However, Don demanded they let him go. Walking through the snow, Don noticed a shape on the hillside that he thought was his ski. However, as he approached the item, he noticed that it was just a large stick. Turning around, he doubted he could find his way back.

Target: Don realized his mistake had been to ignore his friends' advice.

Comprehension Question: Did Don's friends agree with his decision?

2S/Coherent Discourse Context

As he returned from skiing with friends, Don discovered that one of his skis was missing. Although it was getting dark and snowing heavily, Don proposed a plan to go back and find the ski that evening. However, his friends warned him that it would be dangerous to leave the cabin so late in the day. They promised him assistance in looking for the ski, if he would wait until the next morning. However, he asserted his right to make his own decisions. After two hours of

walking through the snow, Don noticed a long, dark shape on the hillside. Digging the object from the snow, he revealed a large stick, not a ski. Turning around, he doubted he could find his way back.

Target: Don realized his mistake had been to ignore his friends' advice.

Comprehension Question: Did Don find the ski soon after beginning his search?

6S/Scrambled Sentences

As he returned from skiing with friends, Don discovered that one of his skis was missing. Walking toward the spot, one of the boys asked Don, if he would describe the fish. Within the first ten minutes of his shift, an angry woman said a green car had taken off her side mirror. However, his father warned him that the weather forecast predicted rain that day. As he greeted her, he saw that she held a book that he had just read. Don explained he could sleep with the light on, as long as the room was quiet. Sensing her disappointment, he proposed that he could rebuild it. Turning around, he doubted he could find his way back.

Target: Don realized his mistake had been to ignore his friends' advice.

Comprehension Question: Did the weather forecast predict rain?

2S/Scrambled Sentences

As he returned from skiing with friends, Don discovered that one of his skis was missing. Trying Don's pie, they confirmed the quality of his baking. Because it was late, they feared the possibility of being stranded in the woods. Don denied the truth of the man's statement. Pulling aside a curtain, Don revealed a large window hanging that produced a shadow. At eight o'clock, his friends noted the fact that they were still waiting for their bus. As the tour began, Don explained their policy of no photography.

Turning around, he doubted he could find his way back.

Target: Don realized his mistake had been to ignore his friends' advice.

Comprehension Question: Was Don concerned about being stranded in the woods?

Set #8

6S/Coherent Discourse Context

Walking to a lake with his two nephews, Ben advised them to select a particular spot to feed the fish. He claimed an enormous fish had lived there for the past several years. Walking toward the spot, one of the boys asked Ben to describe the fish. Ben explained the fish was gray with spots. Feeding the fish, the boys noted only small ones came into view. They suggested the fish may have moved to another location. However, Ben insisted they remain patient and that they would soon see the fish. After another hour of watching and waiting, the nephews indicated they had given up on seeing the fish.

Target: Ben knew his nephews were getting impatient.

2S/Coherent Discourse Context

Walking to a lake with his two nephews, Ben advised them to select a particular spot to feed the fish. He recalled an enormous fish that had lived there for the past several years. Walking toward the spot, one of the boys asked Ben, if he would describe the fish. Ben remembered the fish was gray and had spots. Feeding the fish, the boys observed small ones only. Disappointed, they doubted the fish's presence in the area. However, Ben demanded their patience in waiting for the fish. After another hour of watching and waiting, the nephews indicated they had given up on seeing the fish.

Target: Ben knew his nephews were getting impatient.

6S/Scrambled Sentences

Walking to a lake with his two nephews, Ben advised them to select a particular spot to feed the fish. For weeks before the tournament, Ben boasted that he would score lower than all of his friends. Ben asked him if they would be able to go another weekend. He explained he was a bit concerned about the insects that they might encounter on the trip. He insisted he would take the picture. They suggested he try another contestant's entry before making such a claim. Ben believed that his nephews were telling the truth. After another hour of watching and waiting, the nephews indicated they had given up on seeing the fish.

Target: Ben knew his nephews were getting impatient.

2S/Scrambled Sentences

Walking to a lake with his two nephews, Ben advised them to select a particular spot to feed the fish. Traveling to the stream, he confessed his lack of canoeing experience. Ben requested that she show him the contract. Seeing all three of their desks next to Ben's, they guessed their reasons for insisting on lesson for him. The man denied blame for the damage. However, he asserted his right to make his own decisions. Ben admitted the doubtfulness of his idea. After another hour of watching and waiting, the nephews indicated they had given up on seeing the fish.

Target: Ben knew his nephews were getting impatient.

Set #9

6S/Coherent Discourse Context

After three hours of hiking, Linda and Ed realized they were lost. Because it was late, they feared that they would not get back before dark. Linda suggested they return the way the came. However, Ed guessed that they had been walking in circles. He revealed his idea to camp

in the woods that evening and find their way back the next day. Then, Linda explained she did not want to camp there overnight. She guessed that they had plenty of time to get back that day. After another hour of walking, they discovered their car in the parking lot where they had left it.
Target: Linda announced her decision had been correct.

2S/Coherent Discourse Context

After three hours of hiking, Linda and Ed realized they were lost. Because it was late, they feared the possibility of being stranded in the woods. Linda suggested the idea of walking back the way they came. However, Ed guessed that they had been walking in circles. He revealed his idea to camp in the woods that evening and to find their way back the next day. Unhappy with Ed's idea, Linda indicated her dislike for camping. Linda maintained her preference for trying to return that day. After another hour of walking, they discovered their car in the parking lot where they had left it.
Target: Linda announced her decision had been correct.

6S/Scrambled Sentences

After three hours of hiking, Linda and Ed realized they were lost. Stopping at one dealership, Linda declared she had found the car of her dreams. As the tour began, the guide explained they were not to take pictures inside of the cavern. However, Ed insisted they remain patient and that they would soon see the fish. She recalled that her table was ruined while in storage. Each night, he observed the monster crawling across the floor. Suddenly, from the audience, a man said he had seen another magician do the same stunt. After another hour of walking, they discovered their car in the parking lot where they had left it.
Target: Linda announced her decision had been correct.

2S/Scrambled Sentences

After three hours of hiking, Linda and Ed realized they were lost. They explained she was a new student named Linda. She asked Linda, if she would look at her portfolio and give her feedback. Linda asked Ed about the role he wanted. However, they explained the high quality of some of the other entries. A few minutes later, Ed noticed the dull point of his pencil, and sharpened it. She advised them to be patient. After another hour of walking, they discovered their car, in the parking lot where they left it.

Target: Linda announced her decision had been correct.

Set #10

6S/Coherent Discourse Context

As a candidate for magistrate, Jason announced an information booth that he would open at a local fairground. After an hour of waiting, he noticed that there was a lemonade stand nearby. He decided he would take a short break, and get a drink. Returning with his lemonade, he noted there was a woman at his stand. As he greeted her, he saw that she held a book that he had just read. Gesturing towards the book, he forgot that he had a drink in his hand. Spilling the drink on woman's book and dress, he declared he was very sorry. Grudgingly, she accepted his apology, and walked away.

Target: Jason decided his next move would be to dispose of the lemonade.

2S/Coherent Discourse Context

As a candidate for magistrate, Jason announced an information booth that he would open at a local fairground. He believed that this would allow him to share his views. Noticing a lemonade stand, he decided he would take a short break, and get a drink. As he returned, he noted a woman who was at his stand. Greeting her, he saw a book that she was carrying. As he

gestured towards the book, he forgot the drink he had his hand. Spilling the drink on woman's book and dress, he declared his apology. Grudgingly, she accepted his apology, and walked away.

Target: Jason decided his next move would be to dispose of the lemonade.

6S/Scrambled Sentences

As a candidate for magistrate, Jason announced an information booth that he would open at a local fairground. Although it was getting dark and snowing heavily, he decided he should go back and find the ski before it was covered with snow. However, Jason feared that the monster would attack him. He said he would like to see the maintenance records for the car. He boasted that he had five years of experience. He guessed that they had their own reasons for insisting on oboe lessons for their friend. At eight o'clock, Jason and his friends noticed that the bus had not yet arrived. Grudgingly, she accepted his apology, and walked away.

Target: Jason decided his next move would be to dispose of the lemonade.

2S/Scrambled Sentences

As a candidate for magistrate, Jason announced an information booth that he would open at a local fairground. As she approached, Jason asked her whether she liked the flower box. They explained he was a new boy in school, named Jason. Around midnight, Jason announced his plan to go to bed. They decided they should stop to see what there was. Then, he stated his wish to view her portfolio. Suddenly, he noticed a chipmunk that ran across the sidewalk. Grudgingly, she accepted his apology, and walked away.

Target: Jason decided his next move would be to dispose of the lemonade.

Set #11

6S/Coherent Discourse Context

On her sixteenth birthday, Lynne and her father decided they should begin looking for a used car for her. Stopping at one dealership, Lynne declared she had found the car of her dreams. Approaching them, a dealer asked them, if they needed any help. Lynne's father said he would like to see the maintenance records for the car. After searching, the dealer explained he could not find any records. The father argued that he could not buy the car without this information. However, he noticed the disappointment on his daughter's face. The father indicated he would like a guarantee that he could return the car, if it failed.

Target: The dealer guaranteed the car was returnable for a refund.

2S/Coherent Discourse Context

On her sixteenth birthday, Lynne and her father decided they should begin looking for a used car for her. Stopping at one dealership, Lynne noticed a car that she liked. Approaching them, a dealer asked them if they needed any help. The father requested a chance to see the records for the car. The dealer denied the existence of any such records. Then, the father explained his lack of willingness to buy the car without them. However, he noticed the disappointment on his daughter's face. The father indicated he would like a guarantee that he could return the car, if it failed.

Target: The dealer guaranteed the car was returnable for a refund.

6S/Scrambled Sentences

On her sixteenth birthday, Lynne and her father decided they should begin looking for a used car for her. Her father asserted he did not damage her car. He revealed his idea to camp in the woods that evening and find their way back the next day. As a recent graduate, Lynne

worried she could not get the job. She asked him if there was a problem. Lynne insisted the clause did not cover this case. Gesturing towards the book, he forgot that he had a drink in his hand. The father indicated he would like a guarantee that he could return the car, if it failed.
Target: The dealer guaranteed the car was returnable for a refund.

2S/Scrambled Sentences

On her sixteenth birthday, Lynne and her father decided they should begin looking for a used car for her. Her father advised her to stay up all night to study for the exam. As they looked at the elephants, Lynne observed the bored look on her father's face. Lynne observed a small lamp on one of the tables that appeared to be very old. Greeting her, he saw a book that she was carrying. Opening his bag, he revealed many bottles of insect repellant and sun screen that he had bought for the trip. Unhappy with Lynne's idea, her father indicated his dislike for camping. The father indicated he would like a guarantee that he could return the car, if it failed.
Target: The dealer guaranteed the car was returnable for a refund.

Set #12

6S/Coherent Discourse Context

As an acting student, Adrian decided he would try out for a small role in a play at a community theater. As he waited for his chance to audition, he noticed that there was a young man approaching him. The young man stated that his name was Nathan. Nathan asked Adrian about the role he wanted. Adrian said he was auditioning for the role of the police officer. Nathan informed Adrian that he was auditioning for the same role. He boasted that he had a great amount of experience. He doubted Adrian had a chance of being chosen over him.
Target: Adrian showed his ability was as strong as Nathan's.

2S/Coherent Discourse Context

As an acting student, Adrian decided he would try out for a small role in a play at a community theater. As he waited for his audition, he noticed a young man who approached him. Smiling, the young man stated his name, which was Nathan. Nathan asked Adrian about the role he wanted. Adrian explained his plan to try out for the role of the police officer. Nathan informed Adrian that he was auditioning for the same role. He mentioned his great amount of experience. He doubted Adrian had a chance of being chosen over him.

Target: Adrian showed his ability was as strong as Nathan's.

6S/Scrambled Sentences

As an acting student, Adrian decided he would try out for a small role in a play at a community theater. Having no preschool experience, Adrian feared the unknown.
Adrian said this might be more than they need for such a short trip. Looking through it, Adrian indicated his approval. Trying Adrian's pie, Nathan confirmed that his pie was very good. Soon, however, he thought he might regret this decision. He explained he needed a lawyer to handle his case. He doubted Adrian had a chance of being chosen over him.

Target: Adrian showed his ability was as strong as Nathan's.

2S/Scrambled Sentences

As an acting student, Adrian decided he would try out for a small role in a play at a community theater. While purchasing his ticket, he asked the vendor about the bus schedule. Nathan maintained his plan to take a picture. He explained his idea for making it larger, to hold more flowers. Adrian requested a chance to see the records for the car. He maintained his preference for trying to return that day. Adrian requested a chance to view the contract. He doubted Adrian had a chance of being chosen over him.

Target: Adrian showed his ability was as strong as Nathan's.

Set #13

6S/Coherent Discourse Context

Jacob insisted there was a monster under his bed. He warned his father who approached the bed to stay by the door. Looking under the bed, his father denied he saw a monster. However, according to the boy, the unclaimed shadow on the floor proved the monster was real. Each night, he observed the monster crawling across the floor. The father proposed that Jacob look under the bed with him, to show him the monster. However, Jacob feared that the monster would attack him. His father promised he would protect Jacob, if there was a monster.

Target: Jacob predicted the situation was one that his father could handle.

2S/Coherent Discourse Context

Jacob insisted there was a monster under his bed. He warned his father to stay by the door. Looking under the bed, the father denied any sign of a monster. However, according to the boy, the shadow on the floor indicated the monster's presence. Pulling aside a curtain, the father revealed a large window hanging that produced a shadow. The boy doubted his father's theory. His father proposed the idea that he and the boy look under the bed together. His father promised he would protect Jacob, if there was a monster.

Target: Jacob predicted the situation was one that his father could handle.

6S/Scrambled Sentences

Jacob insisted there was a monster under his bed. However, his father explained they wouldn't be able to go, if it rained. Jacob asked his father about the role he wanted. Returning with his lemonade, he noted there was a woman at his stand. Then, Jacob and his father observed

a man who was coming to them. He proposed that the dog might have knocked over the vase with its tail. Arriving at the stream, Jacob noted his father appeared to be nervous. His father promised he would protect Jacob, if there was a monster.

Target: Jacob predicted the situation was one that his father could handle.

2S/Scrambled Sentences

Jacob insisted there was a monster under his bed. Calling his young son, Jacob, his father demanded an explanation for the broken vase. Approaching them, a dealer asked them if they needed any help. Jacob remembered the fish as being gray and having spots. He explained the way in which some other magicians create an illusion of doing the stunt. The following week, Jacob maintained his schedule of picking up the students from class. Digging the object from the snow, he revealed a large stick, not a ski. His father promised he would protect Jacob, if there was a monster.

Target: Jacob predicted the situation was one that his father could handle.

Set #14

6S/Coherent Discourse Context

As a newly hired police officer, Lionel accepted an assignment to patrol a busy section of a highway. Soon, however, he thought he might regret this decision. Within the first ten minutes of his shift, an angry woman said a green car had taken off her side mirror. Then, Lionel and the woman observed a man who was coming to them. The man asserted he did not damage her car. The woman insisted Lionel should believe her. Lionel requested the woman give him evidence of the man's guilt. She noted the paint on the man's car was newly scratched.

Target: The scratched paint proved the woman's case was strong.

2S/Coherent Discourse Context

As a newly hired police officer, Lionel accepted an assignment to patrol a busy section of a highway. Soon, however, he doubted the wisdom of this decision. Within the first ten minutes of his shift, an angry woman claimed a man's car had taken off her side mirror. Soon, the man approached Lionel and the woman. The man denied blame for the damage. He disputed the woman's claim, saying that she had wrecked her own car. In response, Lionel requested any evidence that she might have of the man's guilt. She noted the paint on the man's car was newly scratched.

Target: The scratched paint proved the woman's case was strong.

6S/Scrambled Sentences

As a newly hired police officer, Lionel accepted an assignment to patrol a busy section of a highway. Lionel requested that she show him the contract. Yet, he maintained his innocence. The woman insisted Lionel should be able to join their class. Lionel explained the fish was gray with spots. He claimed he did not know what had happened to the vase. He admitted he had seen the blueprint. She noted the paint on the man's car was newly scratched.

Target: The scratched paint proved the woman's case was strong.

2S/Scrambled Sentences

As a newly hired police officer, Lionel accepted an assignment to patrol a busy section of a highway. He denied knowledge of what happened to the vase. Walking toward the spot, the woman asked him to describe the fish. However, the others confirmed his story. He recalled that his table was ruined by water damage while it was in storage. Suddenly, Lionel noticed a chipmunk that ran across the sidewalk. After two hours of walking through the snow, he noticed a long, dark shape on the hillside. She noted the paint on the man's car was newly scratched.

Target: The scratched paint proved the woman's case was strong.

Set #15

6S/Coherent Discourse Context

Jake and three of his friends decided they would enter a golf tournament. For weeks before the tournament, Jake boasted that he would score lower than all of his friends. However, one of his friends warned Jake that he was being too confident. In response, Jake claimed his friend was simply jealous of his superior ability. Finally, the evening before the tournament, Jake realized his friend had good reason to warn him. He remembered that two of his friends were professional golfers. In fact, they taught him how to golf. He realized there was a good chance that the next day would be an embarrassing one.

Target: Jake's friends sensed his confidence had disappeared.

Comprehension Question: Did Jake confidence improve just before the competition?

2S/Coherent Discourse Context

Jake and three of his friends decided they would enter a golf tournament. For weeks before the competition, Jake advised his friends to not be too disappointed when they lost. According to Jake, he claimed the prize in every tournament he entered. However, one of his friends warned Jake that he was being too confident. But, Jake suspected his friend of being jealous of his ability. Later that night, he remembered a competitor's history as a golf champion. In fact, this competitor taught him how to golf. He realized there was a good chance that the next day would be an embarrassing one.

Target: Jake's friends sensed his confidence had disappeared.

Comprehension Question: Did Jake feel more confident just before the tournament?

6S/Scrambled Sentences

Jake and three of his friends decided they would enter a golf tournament. Jake predicted the weather would be fine. Getting no response, He proposed that they see the elephants. As the woman pulled out her camera, Jake advised that she follow the guide's rule about not taking pictures. He indicated flowerboxes were usually smaller than the one he built. However, he noticed the disappointment on his friends' faces. A few minutes later, he heard the loud ring of his cell phone, and answered it. He realized there was a good chance that the next day would be an embarrassing one.

Target: Jake's friends sensed his confidence had disappeared.

Comprehension Question: Did Jake ignore his cell phone when it rang?

2S/Scrambled Sentences

Jake and three of his friends decided they would enter a golf tournament. Jake warned him to be sure that the lamp was actually an antique. His friends found a leaking pipe that had caused the damage. His friends advised him not to be afraid. Disappointed, they doubted the fish's presence in the area. Smiling, the young man stated his name, which was Jake. Jake recalled his own good experiences in kindergarten. He realized there was a good chance that the next day would be an embarrassing one.

Target: Jake's friends sensed his confidence had disappeared.

Comprehension Question: Did Jake's friends find the cause of the damage?

Set #16

6S/Coherent Discourse Context

One summer, Jennifer decided she would take her son, Craig to the zoo. Arriving at the zoo, Jennifer asked Craig what he would like to see first. Getting no response, Jennifer proposed

that they see the elephants. As they looked at the elephants, Jennifer observed that Craig did not seem to be interested. She guessed that he would prefer the lions. However, at the lions' cage, he explained he was bored. Suddenly, Jennifer and Craig noticed that a chipmunk ran across the sidewalk. Craig announced his delight at seeing it.

Target: Jennifer said a better idea would have been to take him to a park.

2S/Coherent Discourse Context

One summer, Jennifer decided she would take her son, Craig to the zoo. Arriving at the zoo, Jennifer asked Craig what he would like to see first. Since Craig did not answer, Jennifer suggested the idea that they see the elephants first. As they looked at the elephants, Jennifer observed the bored look on Craig's face. She guessed that he would find the lions more exciting. However, at the lions' cage, he mentioned his lack of interest. Suddenly, Jennifer and Craig noticed a chipmunk that ran across the sidewalk. Craig announced his delight at seeing it.

Target: Jennifer said a better idea would have been to take him to a park.

6S/Scrambled Sentences

One summer, Jennifer decided she would take her son, Craig to the zoo. In response, Craig claimed his friend was simply jealous of his superior ability. Further, he claimed he would be the first person to do this. Craig's mother sensed his nervousness. Jennifer boasted that she had five years of experience. Having bought the lamp, Craig and Jennifer sensed they had made a good purchase. She explained their bus was scheduled to depart at eight o'clock. Craig announced his delight at seeing it.

Target: Jennifer said a better idea would have been to take him to a park.

2S/Scrambled Sentences

One summer, Jennifer decided she would take her son, Craig to the zoo. She asked him if there was a problem. He decided he would take a short break, and get a drink. Jennifer accepted their advice, and tried a piece of apple pie. He disputed the woman's claim, saying that she had wrecked her own car. However, Craig demanded their patience in waiting for the fish. Jennifer suggested the idea of walking back the way they came. Craig announced his delight at seeing it.
Target: Jennifer said a better idea would have been to take him to a park.

Appendix H: Filler Items

Item #1

Travis was a waiter who worked the morning shift. One morning, a man came into the restaurant. He suggested to Travis that he was thinking about ordering an omelet. He asked what ingredients were in the omelet. Travis described to him everything that was included in the omelet. The man explained to Travis that he would like the omelet, but without ham or mushrooms. Travis wrote down the order. Then, the man decided to change his order. He said that he would also like to have the omelet with no green peppers or onions. Travis asked if he would like the omelet to have cheese. The man did not want to have cheese either. In response, Travis wrote down that the man would have a scrambled egg.

Target: Travis was used to customers making changes to the menu.

Comprehension Question: Did Travis' customer make his decision quickly?

Item #2

As a project for their fifth grade art class, the students were assigned to make a piece of fruit out of paper mache. The students were given two weeks to complete the project. Andy, who did not enjoy art class, delayed working on the project until the night before it was due. However, that night, his mother told him that he had not left enough time for the paper mache to dry. At first, Andy worried about not finishing it. But then, a solution finally occurred to him. He decided to make a grape.

Target: His teacher was not pleased with his project.

Item #3

One evening, Lindsay was riding a bus home from work, when she became occupied with a novel. The man did not want to have cheese either. However, he wanted to continue to see her,

and decided to ask her for a date. One evening, a teenage couple came into the coffee shop. She explained to her that she was planning to add them to her collection. Lindsay shrugged in response. She realized that she could have been home much sooner.

Target: She apologized to her mother for being late.

Item #4

One day, a woman was shoveling snow in her driveway, when her favorite pin fell into the snow. Before she realized what had happened, it became buried under the snow. Despite an hour of searching, she was unable to find it. That night, she explained to her husband that she had lost the pin. A few days later, he was able to find the pin in the snow, but it was badly damaged. He took it to a jeweler to have it repaired. Meanwhile, the woman, unaware of her husband's discovery, decided to buy a new pin, just like the one she had lost. Wearing her new pin home, she soon found, beside a note from her husband, a badly repaired version of the same pin. The old pin was in poor condition, and she hoped that her husband would not be disappointed, if she wore the new pin instead. However, after reading the note, she decided to return the new pin to the store.

Target sentence: She didn't want to make him feel that his effort was unappreciated.

Comprehension Question: Was the woman's husband concerned about the lost pin?

Item #5

For weeks, Carrie had been trying to come up with an idea for an art project for class. Approaching her roommate, she wanted to know whether she was behind the plan. She brought out the dog's cage. She showed Carrie the coupon the she had for saving one dollar on two cans of soups. A week later, she found out that she had received an A on her drawing.

Target: She was surprised that the teacher liked her project.

Comprehension Question: Was Carrie shown a coupon for soup?

Item #6

Several students were enrolled in an American History course at a small university. Since the course was difficult, many of them were doing poorly. They feared that they would not pass. Just before the final exam, they were informed that their worst fears might become reality. Their instructor told them that the final exam would be more difficult than the other exams. She suggested to them that they form study groups for this exam. They decided that they should take her advice, so that they would do better this time.

Target: All of the students did better than they expected on the exam.

Item #7

Martin and Hazel enjoyed telling others about the story of how they met. He had to admit to them that his advisors had warned him. Each week, he talked to boy scouts troops about recycling and keeping the park free of litter. Hazel asked whether he was sure that he wanted pure espresso. She asked how she could have the time, if her watch was broken. Following her audition, she was discouraged by the look on the director's face. Next, as the dog watched, he put the dog's favorite blanket into the cage. The man did not want to have cheese either. Hazel and her mother went to the pet store to pick out a kitten. As he moved to get a better view, his head hit against the glass. She remembered how the seashells that she had just sat in a cardboard box. She said that she had been buying for two years, and had two encyclopedias already.

Target: She was glad to be finished with buying encyclopedia volumes.

Comprehension Question: Did Hazel go to the pet store to buy a puppy?

Item #8

A few years ago, a national park was having problems with bears. He called upstairs to his wife, and asked her to help him find them. She had hoped for more enthusiasm from them. She wanted to have a chance to begin a career in the ballet. He described to him everything that was included in the omelet. They were together ever since that time. He explained to the press that there must be overlap between the intelligence of the smartest bears and the dumbest tourists.

Target: They would continue to work on designing better lids.

Item #9

As a high school guidance counselor, Jane believed that it was her duty to help students deal with conflict. After years of training, Jane was aware of many strategies for helping students resolve conflict. She was convinced that students should meet in groups to discuss their disagreements. At these meetings, students were given a chance to explain their viewpoints to Jane and to the other students. Then, each student was given a chance to state the first speaker's viewpoint. Jane explained to the students various ways to come to an agreement. At first, some of the students were resistant to her ideas. However, once the conflict was resolved, they began to see the value of her strategies. They were happy to learn better ways to interact with their peers. Many students remembered her strategies for years to come.

Target: She frequently heard from students years after their training.

Item #10

One day, Scott became concerned about his dog's behavior. He decided to take his dog to the veterinarian. He brought out the dog's cage. Next, as the dog watched, he put the dog's favorite blanket into the cage. Finally, the dog was coaxed into the cage. Arriving at the

veterinarian's office, Scott explained to the veterinarian all of the dog's symptoms. The veterinarian gave the dog a shot, and suggested to Scott that they return for another shot the following week. The following week, Scott did as the veterinarian suggested. He brought out the dog's cage, and put into it the dog's favorite blanket. The dog began to seem worried. Then, he tried to coax the dog into the cage. At this, the dog ran upstairs and hid.

Target: Scott quickly ran upstairs and looked for the dog.

Item #11

Two swimmers, each convinced that she was faster than the other, decided to race to the shore. They decided to bet each other fifty dollars that they would win. Within a half hour of taking off, one swimmer had made it to shore, while the other finished a few minutes later. The woman who had won collected her prize, and spent the rest of the day

bragging to her friends. The woman who had lost admitted to being defeated, but then explained to her friends that the other swimmer must have cheated.

Target: Her friends agreed with her that she was the faster swimmer.

Item #12

One morning, Adam decided to get the mail. Many people from the community came to the dinner to see the band. However, he did not want to appear to change his mind. Adam mentioned that many diseases have not been cured. He said that he would also like to have the omelet with no green peppers or onions. He hoped to convince the man the he already had many customers. Suddenly, the bird's accident did not seem to be so ridiculous.

Target: He went outside to get a better view of the bird.

Comprehension Questions: Did Adam want green peppers and onions in his omelet?

Item #13

One evening, Marie decided to order a book. Sitting there, she noticed that there was a man in the waiting room. When the concert was over, Marie's family met came up to her in the lobby. After trying the espresso, the boy explained to Marie that the espresso was stronger than usual. Three days later, she was surprised to find two identical books in her mailbox.

Target: She decided to give one of the books to a friend.

Item #14

George, a young lawyer, was eager to tell his friends how he had won a difficult case. Since there was time, George began to look through his coupons. He asked what ingredients were in the omelet. When he was young, George worked as an encyclopedia salesman. He said that the only thing that he regretted was waiving the fee.

Target: His was told by his friends that they were happy for him.

Item #15

Mrs. Johns, a third-grade teacher, said her students could do an extra credit assignment. She proposed to them two possible assignments. They could choose to learn the multiplication table or memorize a poem. As she looked around the room, she noticed that none of the students seemed to be interested in the assignments. Disappointed, she returned to her seat. She had hoped for more enthusiasm from them. She assumed that they were satisfied with the grades that they had already earned. However, she was mistaken. One student, Jordan, said that he would like to do one of the assignments. However, he explained to her that he not yet decided which assignment to do. He was told that he should choose the assignment that he would find more enjoyable. He decided to memorize a poem, since he thought it would be easier for him.

Target: Mrs. Johns was pleased that he had chosen to memorize a poem.

Item #16

Angela worked at an art gallery, and frequently led group tours. Each Saturday, her job was to lead groups of high school students though the gallery. One Saturday, she noticed that one of the students was carefully examining each of the paintings. She was pleased to see that a young person was so interested in the artwork. At the end of the tour, he seemed to be interested in one of the paintings. So, she asked whether he had any questions about that painting. Looking confused, he asked where the price tags were.

Target: He was informed that the gallery was not a store.

Item #17

Walter was a retired boxer who lifted weights each day to maintain his strength. He wanted to sit down and have some tea. Walter hoped that other areas of space may have the materials that are needed for these reasons. As Walter's parents watched his play, his three-year-old sister stood on a chair, so that she could see. Walter explained to her that he never found time to do that. He proposed to them two possible assignments. The veterinarian gave the dog a shot, and suggested to Walter that they return for another shot the following week. Still curious, Walter wondered whom he could ask next. Suddenly, he felt better about his chances of being selected. He asked the man for a rematch.

Target: He wanted to have one last chance to prove himself.

Comprehension Question: Did the veterinarian talk to Walter about a dog?

Item #18

Kevin, who is an astronomer, was interviewed by a reporter. He said to Kevin that the repair would be a simple one. Her roommate refused to tell her anything about the plans. One morning, a man came into the restaurant. He decided to rent an office space, and furnished it

with antiques. The girl began to sip at her drink. However, they burned only a few minutes after she continued baking them. Her mother asked why she had not gotten off right away and walked back. When he arrived, Kevin was greeted by the host, and introduced to several other men who had been athletes. He said to the

reporter that he tells this to people who ask him about his field.

Target: He hoped to convince others of the value of astronomy in daily life.

Item #19

Anita's parents and younger sister went to her junior high school concert. She thought how ridiculous it was that the bird did not realize that a window was there. Anita explained to them that she had to make many calls to talk to her classmates about school. Frustrated, she decided to take a break. An hour later, she finally made it home. However, he did not want to appear to change his mind. Anita mentioned that many diseases have not been cured. Nodding, he explained to her that he ordered it all of the time. Still pretending to be busy, he asked if he could help him. Finally, the dog was coaxed into the cage. He was told that he should choose the assignment that he would find more enjoyable. Hearing this, Leslie thanked him.

Target: Anita's father said to her that she had done well also

Comprehension Question: Was he given a choice of which assignment to do?

Item #20

Alex was on his way to the library, when he realized that has forgotten his glasses. At this, Alex looked confused. However, Alex hoped that he would regain his reputation. She was reminded that she needed to choose just one of them. However, he explained to her that he not yet decided which assignment to do. Following her audition, she was discouraged by the look on the director's face. Next, as the dog watched, he put the dog's favorite blanket into the cage. The

man did not want to have cheese either. As they played, she waved her hands in the air. She told him that he was wearing them.

Target: He was embarrassed that he had been so absent-minded.

Item # 21

One afternoon, an old woman was busy planting in her flower garden. As she planted, she noticed that the flowers at the far end of the garden were much more colorful than the flowers that were next to her. So, she went over to that end of the garden to see if she could find a reason for this. Looking around, her eyes fell upon a pile of shiny white stones. So, she gathered into her arms as many white stones as she could before placing them all around the garden.

Target: She wanted to make all of her flowers more beautiful.

Item # 22

Robert was hired to train horses for a wealthy man, who enjoyed racing horses but did not have time to train them himself. For his work, Robert was promised a small salary, and any horse that did not do well in the previous season of racing. All successful horses would stay with their owner. Although the salary was small, Robert decided that he would take the job to obtain some horses of his own. After the first season, three horses had done well, and two had not.

Target: Robert was pleased with his decision to take the job.

Comprehension Question: Did Robert take his job for the large salary?

Item #23

A young businessman started his own firm. However, they burned only a few minutes after he continued baking them. Disappointed, he returned to his seat. He hoped to convince the director to give him another chance. However, he was reluctant to tell her anything else. He brought out the dog's cage. Looking through them, the cashier informed him that several of his

coupons were expired. After he made their drinks, he brought them over to the table. He asked why his sister never collected seashells. However, he was mistaken. This friend also refused to tell him anything. The man said that he was there to activate his phone lines.

Target: He was sure that the workman thought he was foolish.

Item #24

As a forest ranger, Mike believed that it was his responsibility to educate young people about preserving the environment. Each week, he talked to boy scouts troops about recycling and keeping the park free of litter. He explained to them that it is every person's duty to keep the parks free of litter. His goal was to convince each of the boys that they should make an effort to recycle. Mike acknowledged that not all of the boys were going to follow his instructions. Some would forget about his advice as soon they got home. Others would be afraid of criticism by their peers, if they were concerned about the environment. Also, Mike was aware that his ideas would not be reinforced by all parents. However, he hoped to influence some of the boys. He wanted to think that most of the boys would be influenced by his message.

Target: He was convinced that each person that he influenced made a difference.

BIBLIOGRAPHY

Abe, J.A. & Izard, C.E. (1999). Compliance, noncompliance strategies, and the correlates of compliance in 5-year-old Japanese and American children. *Social Development, 8*(1), 1-20.

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

Balogh, J., Zurif, E., Prather, P., Swinney, D., Finkel, L. (1998). Gap-filling and end-of-sentence effects in real-time language processing: Implications for modeling sentence comprehension in Aphasia. *Brain and Language, 61,* 169-182.

Balota, D. A., Cortese, M.J., Hutchinson, K.A., Neely, J. H., Nelson, D., Simpson, G.B., & Treiman, R. (2002). The English lexicon project: A web-based repository of descriptive and behavioral measures for 40, 481 English words and nonwords. HTTP://elexicon.wustl.edu

Bernieri, F.J., Davis, J.M., Rosenthal, R., Knee, C.R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of of sound and facial affect. *Personality and Social Psychology Bulletin, 20*(3), 303-311.

Biber, D. (1998). The study of discourse characteristics. In *Corpus Linguistics: Investigating Language Structure and Use.* D. Biber, S. Conrad, & R. Reppen (Eds.). New York: Cambridge University Press.

Bock, J.K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355-378.

Bock, J.K. & Griffin, Z.M. (2000). The persistence of structural priming: Transient activation or implicit learning. *Journal of Experimental Psychology: General, 129*(2), 177-192.

Bock, J.K. & Loebell, H. (1990). Framing sentences. *Cognition, 35*, 1-39.

Boland, J. (1993). The role of verb argument structure in sentence processing: Distinguishing between syntactic and semantic effects. *Journal of Psycholinguistic Research, 22*(2), 133-152.

Boland, J. & Blodgett, A. (2001). Understanding the constraints on syntactic generation: Lexical bias and discourse congruency effects on eye movements. *Journal of Memory and Language, 45*, 391-411.

Branigan, H.P., Pickering, M.J., Cleland, A.A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*, B13-25.

Branigan, H.P., Pickering, M.J. Liversedge, S.P., Stewart, A.J., & Urbach, T.P. (1995a). Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research, 24*(6), 489-506.

Branigan, H.P., Pickering, M.J., & Stewart, A.J. (1995b). Syntactic priming in language comprehension. Unpublished manuscript.

Brantmeier, C. (2003). Does gender make a difference? Passage content and comprehension in second language reading. *Reading in a Foreign Language, 15*(1), 1-27.

Britt, M.A., Perfetti, C.A., Garrod, S., & Rayner, K. (1992). Parsing in discourse: Context effects and their limits. *Journal of Memory and Language, 31*, 293-314.

Bügel, K, & Buunk, B.P. (1996). Sex differences in foreign language text comprehension: The role of interest and prior knowledge. *Modern Language Journal, 80*, 15-31.

Cave, C.B. (1997). Very long-lasting priming in picture naming. *Psychological Science, 8*(4), 322-325.

Chen, E., Gibson, E., & Wolf, F. (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language, 52*, 144-169.

Chow, D. & Skuy, M. (1999). Simultaneous and successive cognitive processing in children with nonverbal learning disabilities. *School Psychology International, 20*(2), 219-231.

Cleland, A.A. & Pickering, M.J. (2006). Do writing and speaking employ the same syntactic representations? *Journal of Memory and Language, 54*, 185-198.

Cleland, A.A. & Pickering, M.J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language, 49,* 214-230.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Coltheart, M.N., Curtis, B., Atkins, P. & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review, 100*, 589-608.

Coltheart, M.N., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual-route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Connine, C., Ferreira, F., Jones, C., Clifton, C., & Frazier L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research, 13*(4), 307-319.

Cook, A.E. & Myers, J.L. (2004). Processing discourse roles in scripted narratives: The influence of context and world knowledge. *Journal of Memory and Language, 50*, 268-288.

Doolittle, A. & Welch, C. (1989). Gender differences on performance on a college-level achievement test (ACT Research Rep, Series 89-9). Iowa City: Iowa: American College Testing Program.

Dunn, L.M. & Dunn, D.M. (2007). *Peabody Picture Vocabulary Test – 4*. Bloomington, MN: Pearson Assessments.

Eagly, A.H. & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of social psychology literature. *Psychological Bulletin, 100*(3), 283-308.

Eakin, D.K. (2005). Illusions of knowing: Metamemory and memory under conditions of retroactive interference. *Journal of Memory and Language, 52*(4), 526-534.

Elman, J.L. (1992). Grammatical structure and distributed representations. In S. Davis (Ed.). Connectionism: Theory and practice (pp. 138-178). New York: Oxford University Press.

Elman, J.L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning, 7*, 195-224.

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*(4), 555-568.

Ford, M., Bresnan, J.W., & Kaplan, R.M. (1982). A competence based theory of syntactic closure. In J.W. Bresnan (Ed.). *The mental representation of grammatical relations* (pp. 727-796). Cambridge, MA: The MIT Press.

Frazier, L. (1987). Linguistic complexity. In A. Davidson (Ed.), *Linguistic complexity and text comprehension: A re- examination of readability with alternative views*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Frazier, L., Taft, L., Roeper, T., Clifton, C., & Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition, 12*, 421-430.

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.

Frost, R. (1994). Prelexical and postlexical strategies in reading: Evidence from a deep and a shallow orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 116-129.

Garnsey, S. (1994). Percentage of completions in a sentence completion task on 107 subjects. Unpublished data.

Garnsey, S. M., Pearlmutter, N. P., Myers, E., & Lotocky, M. (1997), The Contributions of Verb Bias and Plausibility to the Comprehension of Temporarily Ambiguous Sentences, *Journal of Memory and Language, 37*(1), 58-93.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In *Image, Language, and Brain: Papers from the First Mind Articulation Project Symposium*. A. Marantz, Y. Miyashita, & W. O'Neil, (Eds.). (pp. 95-126) Cambridge, MA: The MIT Press.

Gibson, E. (1998). Linguistic complexity: Locality of linguistic dependencies. *Cognition, 68*, 1-76.

Graesser, A.C., Millis, K.K., & Zwaan, R.A. (1997). Discourse comprehension. *Annual Reviews in Psychology, 48*, 163-189.

Graesser, A. C., Olde, B., & Klettke, B. (2002). How does the mind construct and represent stories? In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), Narrative Impact: Social and Cognitive (pp. 229-262). Mahwah NJ: Lawrence Erlbaum Associates.

Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395.

Grammer, K., Kruck, K.B., & Magnusson, M.S. (1998). The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior, 22*(1), 3-29.

Grodner, D., Gibson, E. & Watson, D. (2005). The influence of contextual contrast on syntactic processing: Evidence for strong-interaction in sentence comprehension, *Cognition, 95*, 275-296

Hare, M., McRae, K., & Elman, J.L (2004). Admitting that admitting verb sense into corpus analysis makes sense. *Language and Cognitive Processes, 19*(2), 181-224

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

Hartsuiker, R.J. & Kolk, H.J. (1998). Syntactic persistence in Dutch. *Language and Speech, 41*(8), 143-184.

Hess, D.J., Foss, D.J., & Carroll, P. (1995). Effects of global and local contexts on lexical processing during language comprehension. *Journal of Experimental Psychology: General, 124*(1), 62-82.

Holmes, V. M., Stowe, L. & Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language*, *28*, 501–522.

Howell, D.C. (1992). *Statistical methods for psychology* (3rd Edition). Belmont, CA: Wadsworth Publishing Company.

Humphreys, G.W., Besner, D., & Quinlan, P.T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General, 117*(1), 51-67.

Hunnicutt, S. & Carlberger, J. (2001). Improving word prediction using Markov models and heuristic methods. *AAC Augmentative and Alternative Communication, 17*, 255- 264.

Hyde, J.S. & Linn, M.C. (1988). Gender differences in verbal activity: A meta-analysis. *Psychological Bulletin, 104*, 53-69.

Jennings, F., Randall, B., & Tyler, L.K. (1997). Graded effects of verb subcategory preferences on parsing: Support for constraint-satisfaction models. *Language and Cognitive Processes, 12*(4), 485-504.

Jurafsky, D. (2002). Pragmatics and Computational Linguistics. In Laurence R. Horn & Gregory Ward (Eds.) *Handbook of Pragmatics*. Oxford: Blackwell.

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

Kashak, M.P. & Glenberg, A.M. (2004). This construction needs learned. *Journal of Experimental Psychology: General, 133*(3), 450-467.

Kashak, M.P., Loney, R.A., & Borrenggine, K.L. (2006). Recent experience affects the strength of structural priming. *Cognition, 99*, B73-B82.

Kintsch, W. (1998). Comprehension: A Paradigm for Comprehension. New York: Cambridge University Press, pp. 13-48.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182.

Kochanska, G., Trebkes, T.L., & Forman, D.R. (1998). Children's emerging regulation of conduct: Restraint, compliance, and internalization from infancy to the second year. *Child Development, 69*(5), 1378-1389.

La France, M., & Ickes, W. (1981). Posture mirroring and interactional involvement: Sex and sex-typing effects. *Journal of Nonverbal Behavior, 5*(3), 139-154.

Langston, M.C., & Trabasso, T. (1999). Modeling causal integration and availability of information during comprehension of narrative texts. In H. van Oostendorp & S.R. Goldman (Eds.), *The construction of mental representations during reading*. Mahwah, NJ: Erlbaum.

Lapata, M., Keller, F., & Schulte im Walde, S. (2001). Verb frame frequency as a predictor of verb bias. *Journal of Psycholinguistic Research* 30(4), 419-435.

Letts, C., & Leinonen, E. (2001). Comprehension of inferential meaning in language-impaired and language-normal children. *International Journal of Language and Communication Disorders, 36*, 307-328.

Luka, B.J. & Barsalou, L.W. (2005). Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language, 52*, 436-459.

Lund, K. & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*, 203-208.

MacDonald, M.C. & Christiansen, M.H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109(*1), 35-54.

Markwardt, F.C. (1997). *Peabody Individual Achievement Test – Revised/Normative Update*: Bloomington, MN: Pearson Assessments

Martens, S. & Wolters, G. (2002). Interference in implicit memory caused by processing of interpolated material. *The American Journal of Psychology, 115*(2), 169-185.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*, 440-466.

Merlo, P. (1994). A corpus-based analysis of verb continuation frequencies, *Journal of Psycholinguistic Research, 23*(6), 435-457.

Meyers-Levy, J. & Maheswaran, D. (1991). Exploring differences in males' and females' processing strategies. *Journal of Consumer Research, 18*, 63-70.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research, 24*, 469-488

Monahan, J.L., Murphy, S.T., & Zajonc, R.B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science, 11*(1), 462-466.

Nation, K. & Snowling, M.J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*, B1-B13.

Nicholas, L.E. & Brookshire, R.H. (1995). Comprehension of spoken narrative discourse by adults with aphasia, right-hemisphere brain damage, or traumatic brain injury. *American Journal of Speech-Language Pathology, 4*(3), 69-81.

Nippold, M. (1998). *Later language development: The school age and adolescent years.* Austin,TX: Pro-Ed.

Nixon, S.M. (2006). Effects of orthographic, phonologic, and semantic information sources on visual and auditory lexical decision. Unpublished doctoral dissertation: University of Pittsburgh.

Noppeney, U., & Price, C.J. (2004). An fMRI study of syntactic adaptation. *Journal of Cognitive Neuroscience, 16*(4), 702-713.

Oldfield, R.C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia. 9*(1), 97-113.

Perfetti, C.A., Landi, N. & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M.J. Snowling & C. Hulme (Eds.). *The Science of Reading: A Handbook* (pp. 227-247). Oxford: Blackwell.

Perfetti, C.A. & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24-33.

Perfetti, C.A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In R. Frost & L. Katz (Eds.). *Orthography, Phonology, Morphology, and Meaning* (pp. 227-248). Amsterdam: North-Holland.

Pickering, M.J. & Branigan, H.P. (1999). Syntactic priming in language production. *Trends in Cognitive Science, 3*(4), 136-141.

Pickering, M.J. & Branigan, H.P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*, 633-651.

Pickering, M.J., Traxler, M.J., & Crocker, M.W. (2000). Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language, 43*, 447-475.

Potter, M.C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language, 38,* 265-282.

Rodd, J., Gaskell, G., & Marslen-Wilson, W.D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46*, 245-266.

Roland, D. (2001). Verb sense and verb subcategorization probabilities. Unpublished doctoral dissertation. University of Colorado.

Roland, D., Elman, J.L., & Ferreira, V.S. (2006). What is that? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition, 98*, 245-272.

Roland, D. & Jurafsky, D. (2002). Verb sense and verb subcategorization probabilities. In P. Merlo and S. Stevenson (Eds.), *The Lexical Basis of Sentence Processing: Formal, Computational, and Experimental Issues* (pp. 325-245).Amsterdam: John Benjamins.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 1.1) [Computer software]. Pittsburgh, PA: Psychology Software Tools, Inc.

Schwanenflugel, P.J. & White, C.R. (1991). The influence of paragraph information on the processing of upcoming words. *Reading Research Quarterly, 26*(2), 160-177.

Seidenberg, M.S. (2005). Connectionist models of word naming. *Current Directions in Psychological Science, 14*, 238-242.

Seidenberg, M.S. & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.

Shapiro, L.P., Zurif, E.B., & Grimshaw, J. (1989). Verb processing during sentence comprehension: Contextual impenetrability. *Journal of Psycholinguistic Research, 18*(2), 223-243.

Sharkey, A.J.C. & Sharkey, N.E. (1992). Weak contextual constraints in text and word priming. *Journal of Memory and Language, 31*, 543-572.

Shaywitz, S.E. & Shaywitz, B.A. (2004). Reading disability and the brain. *Educational Leadership,* 7-11.

Slosson, R.L. & Nicholson, C.L. (2002). *Slosson Oral Reading Test –R3*. East Aurora, NY: Slosson Educational Publications, Inc.

Snellen, H. (1862). Eye examination chart originally created by Dutch ophthalmologist Hermann Snellen, M.D.

Spivey, M.J. & Tanenhaus, M.K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6), 1521-1543.

Tabossi, P. & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition, 20*, 303-313.

The New Oxford American dictionary (2nd ed.). (2005). New York: Oxford University Press.

Trabasso, T. & Bartolone, J. (2003). Story understanding and counterfactual reasoning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*(5), 904-923.

Trabasso, T. & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes, 15*(3), 249-75.

Trabasso, T. & Sperry, L.L. (1985). Causal relatedness and importance in story events. *Journal of Memory and Language, 24*, 595-611.

Trabasso. T. & Wiley, J. (2005). Goal plans of action and inferences during comprehension of narratives. *Discourse Processes, 39*(2&3), 129-164.

Traxler, M.J, Foss, D.J., Seely, R.E., Kaup, B., & Morris, R.K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research, 29*(6), 581-595.

Trueswell, J.C., Tanenhaus, M.K. & Kello, C. (1993). Verb-specific constraints on sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 528-553.

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

Warren, T. & McConnell, K. (2006). Effects of building syntactic predictions on eye-fixations in reading. Poster session submitted to the Annual CUNY Conference on Human Sentence Processing, New York, NY.

Wirth, M., Horn, H., Koenig, T., Stein, M., Federspiel, A., Meier, B., Michel, C.M., Strik, W. (2007). Sex differences in semantic processing: Event-related brain potentials distinguish between lower and higher order semantic analysis during word reading. *Cerebral Cortex, 17*, 1987-1997.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Riverside Publishing: Rolling Meadows, IL.

Young, D. & Oxford, R. (1997). A gender-related analysis of strategies used to process written input in the native language and a foreign language. *Applied Language Learning, 8*(1), 43-73.

Zwaan, R.A., Magliano, J.P., & Graesser, A.C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 386-397.